CONTENTS

CHAPTER ONE : **INTRODUCTION TO RADIO COMMUNICATION**

The spectrum of electromagnetic wave, 1, Electronic devices at high frequency ranges 2, Need of high frequencies 3, Modulation Methods 5, Reception of Radio signals 6, General picture of radio transmission and reception 6.

CHAPTER TWO : **RESONANT CIRCUITS**

Selective circuits 9, Series resonant circuit 9, Series circuit as a selector 12, The Q of a series circuit 14, Parallel resonant circuit 15, Discussion of Q values 18, Selectivity of parallel resonant circuit 19, Further discussion of sharpness of resonance in parallel circuit 20, Effect of source impedance on the sharpness of resonance curves 21, Design of an anti-resonant circuit 22, Comparison between series and parallel resonance 23, Coupled circuits 24, Terms used in coupled circuits 24, The air core transformer 28, Behaviour of resonant circuits tuned to the same frequency and coupled together 33, Fidelity of receivers 36, Numericals 37.

CHAPTER THREE : **FOUR TERMINAL, NETWORK AND NETWORK THEOREM**

Four terminal network 41, Impedance matching 48, Insertion loss 49, Superposition theorem 52, Reciprocity Theorem 57 Thevenin's theorem 61, Norton's theorem 67, The maximu. power transfer theorem 73.

CHAPTER FOUR : **FILTER CIRCUITS**

Filter circuits 79, Elementary filter theory 79, Characteristic impedance of T and π sections 82, Characteristic impedance as a function of open circuit and short circuit impedances 83, Low pass filter 84, High pass filter 85, Band pass filter 86, Band elimination filter 90, Disadvantages of simple types of filter 93, m-derived T section 93, m-derived π section filter 96, m-derived band pass filter 98, Attenuators 100, Numerical examples 103.

CHAPTER FIVE : **THERMIONIC TUBES**
VACUUM TUBES

Vacuum diode 108, Vacuum triode 110, Tube parameters 112, Vacuum tetrode 115, Pentode tube 117, Beam

power tube 121, Coefficient of screen grid, Beam and Pentode tubes 123, Design consideration of power and high frequency tubes 124, Cathode ray tube, Oscilloscope 125, Motion of an electron in a parallel plate diode 125, Motion of an electron in a constant magnetic field 133, Motion of an electron in electric and magnetic fields at right angle to each other 136, Cathode ray tube 138, Electrostatic deflection sensitivity 141, Magnetic deflection sensitivity 144, Cathode ray tube connections 146, Klystron 153, Magnetron 154.

GAS TUBES

Effect of gas in a thermionic diode 155, Gas diodes with thermionic cathodes 156, Effect of gas pressure on the characteristics of a gas diode 158, Thyratrons 159, Cold cathode gas diode (V.R. Tube) 162, Tubes with mercury pool cathods 164, Ignitron 165.

PHOTO ELECTRIC DEVICES

Classification of Photo-electric devices 166, Photo-emissive cathodes 166, Photo tubes 168, Photo electric relay 172, Photo multipliers 172, Photo-conductive devices 174, Photo voltaic cells 176, Exp. No. 1. To measure the inter-electrode capacitance of triode 177, Exp. No. 2. To measure the dynamic constants of a triode valve 178, Exp. No. 3. To draw the characteristic curves of gas photo electric cell 180, Exp. No. 4. To determine the value of Planck's constant and the photoelectric work-function of the metal with the help of a photocell 181, Expt. No. 5. To determine (*a*) striking potential of a gas filled triode (b) ionisation potential of a gas-field triode 184, Expt. No. 6. Measurement of *e/m* by Helical Method 186, Exp. No. 7. To determine *e/m* for cathode rays by Thomson method 189, Expt. No. 8. Measurement of *e/m* by Magnetron method 191, Numerical Examples 193.

CHAPTER SIX : **SEMICONDUCTORS AND TRANSISTORS**

Chemical bonds in semiconductors 205, Band gap 206, Intrinsic and extrinsic semiconductor 209, Mobility, drift velocity and conductivity of intrinsic semiconductor 210, Carrier concentration of intrinsic semiconductor 212, Law of mass action 218, Impurity semiconductors 219, Impurity states 221, P-N junction 224, Barrier voltage 226, Biassing of P-N junction 230, Current across P-N junction 233, Rectifier equation 235, Diode resistance 237, Depletion region 238, Junction transistor 242, Characteristics of transistors 246, Importance of characteristics 248,

Load line 249, Manufacture of Grown junction diodes and transistors 252, Manufacture of fused junction or alloy junction diodes and transistor 253, Point contact diode or transistors 255, The tunnel Diode 256, Zener diode or breakdown diode 257, The field effect transistor 258, JFET 259, MOSFET 264, The unijunction transistor 267, Gunn effects and diodes 267, Impatt (Impact Avalanche and transit time) diodes 273, Pin diodes 275, Schottky barrier diode 276, Backward diodes 277, Silicon controlled rectifier 278, Triac 279, Diac 280, Transistor as a four [illegible] 280, Impedance parameters or Z parameters 281, Admittance parameters or Y parameters 284, Hybrid parameters or *h* parameters 285, Load line of diode circuit 287.

PTER SEVEN : ELECTRON EMISSION

Crystal structure 292, Bound and free electrons in the crystal lattice 293, Surface Barrier 294, Potential barrier diagram of metal 295, Fermi Dirac Energy distribution curve 296, Electron emission 297, Thermionic emission 299, Richardson equation 299, Derivation of Richardson equation by Fermi-Dirac statistics 302, Experimental verification 304, Explanation for the discrepancies in theoretical and experimental results 305, Practical emitters 306, Comparative-study 307, Photoelectric emission 308, Secondary emission 308, Field emission 310, Schottky effect 310, Space charge and temperature limited currents 315, Child Langmuir Law 318, Child Langmuir equation for cylindrical electrodes 321, Experiment :—To determine the work function of the emitter (pure metal) of vacuum tube 324, Numerical examples 326.

CHAPTER EIGHT : RECTIFIERS AND POWER SUPPLY

Vacuum tube half wave rectifier 334, Vacuum tube full wave rectifiers 342, Comparison, of half and full wave rectifiers 348, Gas diode half wave rectifier 348, Gas diode full wave rectifier 350, Comparison of vacuum and gas diode rectifiers and uses 351, Full wave bridge rectifier circuit 352, Metal rectifiers or barrier layer rectifiers 353, Multiphase rectifier circuits 356, Three phase half wave system 356, Six phase half wave system 358, Three phase full wave system 359, Controlled rectifiers 360, Thyratron controlled rectifier 360, Phase shift control 364, On off control 366, Ignitron rectifiers and Ignitron control circuit 366, Powers supply 368, Series inductor filter 369, Shunt capacitor filter 372, Choke input filter 378, Capacitor input filter 381, Regulation of [illegible] supply output voltage 383, Voltage regulator (V.R.

tube) 384, Electronic potential regulator 386, Electronic voltage stabilizer 387, Analysis of electronic voltage stabiliser 388, Modified form of electronic voltage stabliser 391, Voltage doubler 392, Experiment :—To design the power supply unit and to determine (i) voltage regulation (ii) percentage of ripple 392, Numerical examples 39[illegible].

CHAPTER NINE . **BASIC AMPLIFIER PRINCIPLES**

Symbols for voltages and currents 408, Basic voltage amplifier 409, Vacuum tube amplifier as a linear device 410, Zero signal or quiescent operation of vacuum tube amplifier 410, Working of vacuum tube amplifier with small a.c. grid signal 412, Dynamic characteristic 414, Equivalent circuit representation for triode amplifier 415, Voltage gain of a signal stage untuned linear A_1 amplifier 418, Classification of amplifiers 423, Interelectrode capacitances in a triode 427, Input admittance of a triode amplifier 428, Method of biasing the control grid 431, Distortion in amplifiers 432, Waveform distortion due to non-linearity of the tube characteristics 438, Numerical problems 440.

CHAPTER TEN : **UNTUNED POTENTIAL AMPLIFIERS**

Resistance capacitance coupled amplifier 447, Study of the frequency and phase response of R-C coupled amplifier 453, Gain bandwidth product of R-C coupled amplifier 456, Effect of cascading on the bondwidth 457, Effect of cathode resistance and condenser on the frequency response of an amplifier 458, Low frequency compensation of R-C coupled amplifier 459, High frequency compensation of R-C coupled amplifier 461, Inductance capacitance (L-C) coupled amplifier 463, Trasformer coupled amplifier 465, Numerical examples 471.

CHAPTER ELEVEN : **FEED-BACK AMPLIFIERS**

Gain with feed-back 485, Increase in bandwidth of an amplifier due to negative feed back 488, Feed back circuits 490, Current feed back 490, Voltage feed back 492, Compound feed back 493, Effect of feed back on the input terminal impedance 494, Effect of feed back on the output terminal impedance 496, Cathode follower 498. R-C coupled cathode follower 500, Grounded grid amplifier 504, Anode follower 506, Numericals 508.

CHAPTER TWELVE : **AUDIO POWER AMPLIFIERS**

Class A_1 untuned power amplifier 523, Maximum power

sensitivity of series fed A_1 power amplifiers 525, Shunt- and transformer coupled linear A_1 power amplifier 527, Maximum undistorted power 530, Theoretical plate circuit efficiency of triode power amplifier 533, Class A push pull amplifier 534, Class B audio frequency power amplifiers 539, Numericals 541.

CHAPTER THIRTEEN : **RADIO FREQUENCY AMPLIFIERS**

Valve performance at high frequencies 552, Consideration of circuit design and methods of coupling 553, R-C coupling 553, Impedance coupling 553, Transformer coupling 558, Gain bandwidth product of tuned amplifiers 569, Cascaded tuned amplifiers 570, Staggered tuned amplifier 572, Tuned power amplifiers 574, Basic circuit of tuned power amplifier 575, Analytical analysis of class B tuned amplifier 576, Tuned class C amplifier 581, Analysis of class C power amplifier 583, Self oscillations in amplifiers 588, Noise in amplifiers 590, Numericals 591.

CHAPTER FOURTEEN : **SPECIAL AMPLIFIERS**

Uncompensated wideband amplifier 599, High frequency compensation in wideband amplifier 605, Direct coupled amplifiers, Loftin-White form 608, Drift in D-C amplifiers 612.

EXPERIMENTS ON AMPLIFIERS

Expt. (1) To study the frequency response of a single stage audio frequency amplifier with different type of coupling 616, Expt. 2. To study the effect of negative feed back on the performance of an audio frequency amplifier 619, Expt. 3. To study the cathode follower amplifier 620, Expt. 4. To design and study the gain and phase shift of a speech amplifier 621.

CHAPTER FIFTEEN : **OSCILLATORS**

Classification of oscillators 626, Barkhausen criterion for oscillations 627, Analysis of the fundamental oscillatory circuit 629, Tuned plate oscillator 631, Tuned grid oscillator 634, Hartley oscillator 637, Colpitt's oscillator 639, Dynatron oscillator 642, R-C oscillators 645, Phase shift oscillator 645, Wien Bridge oscillator 648, Frequency stability of oscillators 649, Electron coupled oscillator 650, Crystal oscillator 651, Relaxation oscillators : multi-vibrators 654, Electron motion in time varying field 659, Ultra-high frequency effects in conventional tubes 661, Circuit reactance limitations 661, Circuit and tube loss limitations 663, Electron transit time limitations 663, Positive grid

oscillator 668, Klystron oscillator 675, The reflex klystron 684, Magnetron 689, Travelling wave tube 703, Backward wave oscillator (carcinotron) 712, Electron wave tube 714, Experiment : To verify the conditions of oscillations of a Hartley circuit and to determine the dielectric constant at various frequencies 715, Numerical examples 718.

CHAPTER SIXTEEN : **TRANSISTOR AMPLIFIERS**

Equivalent circuit of a transistor 726, Ground emitter circuit 727, Common base amplifier 735, Grounded collector circuit 742, Comparative study of three types of amplifier circuits 750, Emitter bias 750, Bias stabilisation 752, Bias instability in common base amplifier 752, Bias instability in common emitter amplifier 753, Stability factors 753, Stabilisation by collector base resistance 754, Stabilisation by potential divider and emitter resistance 755, Two battery bias stabilisation 759, Bias compensation 759, Current and voltage amplifiers 761, The comm[illegible] emitter amplifier with emitter resistor 762, Simpli[illegible] common emitter hybrid model 764, Effect of an emitter bypass capacitor in low frequency response 765, High frequency effects and hybrid Pi model 766, Relation between hybrid π circuit components and low frequency h-parameters 768, Common emitter short circuit current gain using hybrid pi-model 771, High frequency CE current gain with resistive load 774, Effect of source resistance on frequency response of CE amplifier at high frequencies 776, R. C. coupled amplifiers 778, Transformer coupled amplifiers 783, Negative feed back amplifiers 783, Effect of feed back on bandwidth 791, A. F. power amplifiers 792, Class A amplifier 792 Class B amplifier 794, Tuned amplifiers 797, Single tuned inductively coupled transistor amplifier 798, Double tuned transistor amplifier 803, Experiments : Expt. No. 1. To draw the characteristic curves of PNP junction transistor 810, Expt. No. 2. To draw frequency response curves of a common emitter amplifier circuit and to determine (i) Input impedance Z_i. (ii) Output impedance Z_0. (iii) Power gain 812.

CHAPTER SEVENTEEN : **FIELD EFFECT TRANSISTOR AMPLIFIERS**

Biasing the FET 815, FET small signal model 816. Common source AC amplifier 818, The common drain - source follower 819, Common gate amplifier 820, General treatment of low frequency common source and common drain amplifier 821, Common source amplifier at high frequencies 824, Common drain amplifier (source follower) at high frequencies 827.

CHAPTER EIGHTEEN : **TRANSISTOR OSCILLATORS**

Feed back oscillators : Tuned collector oscillator 831, Tuned emitter oscillator 833, Hartley oscillator 834, Colpitt oscillator 836, Phase shift oscillator 834, Wein bridge oscillator 841.

CHAPTER NINETEEN : **MULTIVIBRATORS**

Switching characteristic of a transistor 844, Switching time in a transistor 845, Multivibrators 846, Astable multivibrator 847, Monostable multivibrator 851, Bistable multivibrator 855.

CHAPTER TWENTY : **PULSE GENERATION AND SCALING VACUUM TUBE CIRCUITS**

Pulse generators : Free running multivibrators 862, M stable multivibrator 862, Bistable multivibrator locking oscillator 867, Saw tooth wave generator 86?, Neon time base circuit 870, Thyratron sweep generators 870, Vacuum tube sweep circuit 871, Circuits to improve the linearity of the vacuum tube sweeps 872, Bootstrap circuit 872, Miller integrating sweep 874, Scaling circuits counting to the base 2; 876, Decade counter 880.

CHAPTER TWENTY ONE : **WAVE SHAPING CIRCUITS**

Linear wave shaping 883, High pass RC circuit 883, High pass RC circuit as differentiator 890, Low pass RC circuit 892, Low pass RC circuit as an integrator 898, Non-linear waveshaping 899, Clipping circuits 900, Peak limiter 900, Base clipper 901, Slicer 902, Clamping circuits 903, PN junction clippers 906, PN junction clamping circuit 9 . Circuits to produce sharp pulses 909, Ringing circuit 910.

CHAPTER TWENTY TWO : **VOLTAGE AND CURRENT SWEEP GENERATORS**

Terminology 912, Fundamental sweep voltage generators 916, Transistor constant current sweep generators 920, UJT sweep waveform 922, Miller integrator sweep circuit 924, Bootstrap sweep generator 929, Current time base generator 932.

CHAPTER TWENTY THREE : **INTEGRATED CIRCUIT FABRICATION**

Integrated circuit technology 937, Advantages and limita-

tions of integrated circuits 938, Basic monolithic inegrated circuit technology 939, Basic processes used in monolithic technology 941, monolithic integrated components 946, Transistors of monolithic integrated circuits 947, Monolithic diodes 950.

CHAPTER TWENTY FOUR : **LINEAR INTEGRATED CIRCUITS : OPERATIONAL AMPLIFIER**

Basic concepts 953, Ideal operational amplifier 953, Practical inverting OP-AMP 957, Practical noninverting OP-AMP 957, Differential amplifier 959, A general purpose OP-AMP 966, Some operational amplifier parameters, Effects of offset 970, Frequency response and stability 972, OP-AMP as scale changer, phase shifter and summing amplifier 973, OP-AMP as voltage follower 974, as current to voltage converter 975, Integrator 976, Differentiator 976, Analog computation 977, Logarithmic amplifier 979, Antilog amplifier 979, Astable multivibrator 980, Bridge amplifier 983, Voltage comparator 984.

CHAPTER TWENTY FIVE : **NON-LINEAR INTEGRATED CIRCUITS : LOGIC FUNDAMENTALS : (Combinational and Sequential Circuits)**

Digital and analong methods 987, number systems used in digital electronics 987, Binary arithmetic in computers 990, Binary codes 994, Boolean alge[illegible] and logic gates 995, The AND operation [illegible], The OR operator 998, The NOT operation 1000, Universal building blocks 1002, Laws of boolean algebra and their conversion into a logic diagram 1005, Reducing boolean expressions 1010, Arithmatic circuits, exclusive OR (XOR) gates 1011, Half adder 1014, Full adder 1015, Parallel adder 1017, Simplification using Karnaugh maps 1019, Logic gate parameters 1024, Logic families 1025, Resistor transistor logic (RTL) 1025, Diode transistor logic (DTL) 1027, Transistor transistor logic (TTL) 1029, Emitter coupled logic (ECL 1030), Sequential circuits 1031, Flip-flop, S-R 1031, Clock pulse 1034, RS/T flip-flop 1035, JK flip-flop 1038, Master slave JK flip-flop 1040, D-type flip-flop 1041, Counters 1041, BCD decade counter 1044.

CHAPTER TWENTY SIX : **TRANSMISSION LINES**

Distributed parameters 1048, Types of transmission lines 1049, Calculation of line parameters 1050, Voltage and current relations on radio frequency transmission line 1059,

1 INTRODUCTION TO RADIO COMMUNICATION

Radio communication, within recent decades, has become a popular way of conveying the intelligence between places that may be separated by thousands of miles. The 'wireless' type of electrical communication utilising the radio wave transmission is called radio communication. The radio waves are the electromagnetic waves produced due to the escape of electrical energy into the free space. They travel with the velocity of light and consist of electric and magnetic fields at right angles to each other. These fields are perpendicular to the direction of travel of the wave. *Therefore radio communication can be defined as the interchange of symbols, signals, intelligence between two or more points, employing electromagnetic waves as the medium of transmission.*

1·0. THE SPECTRUM OF ELECTROMAGNETIC WAVES :

The electromagnetic waves cover a wide range of frequencies or wavelengths. They are classified as follows :

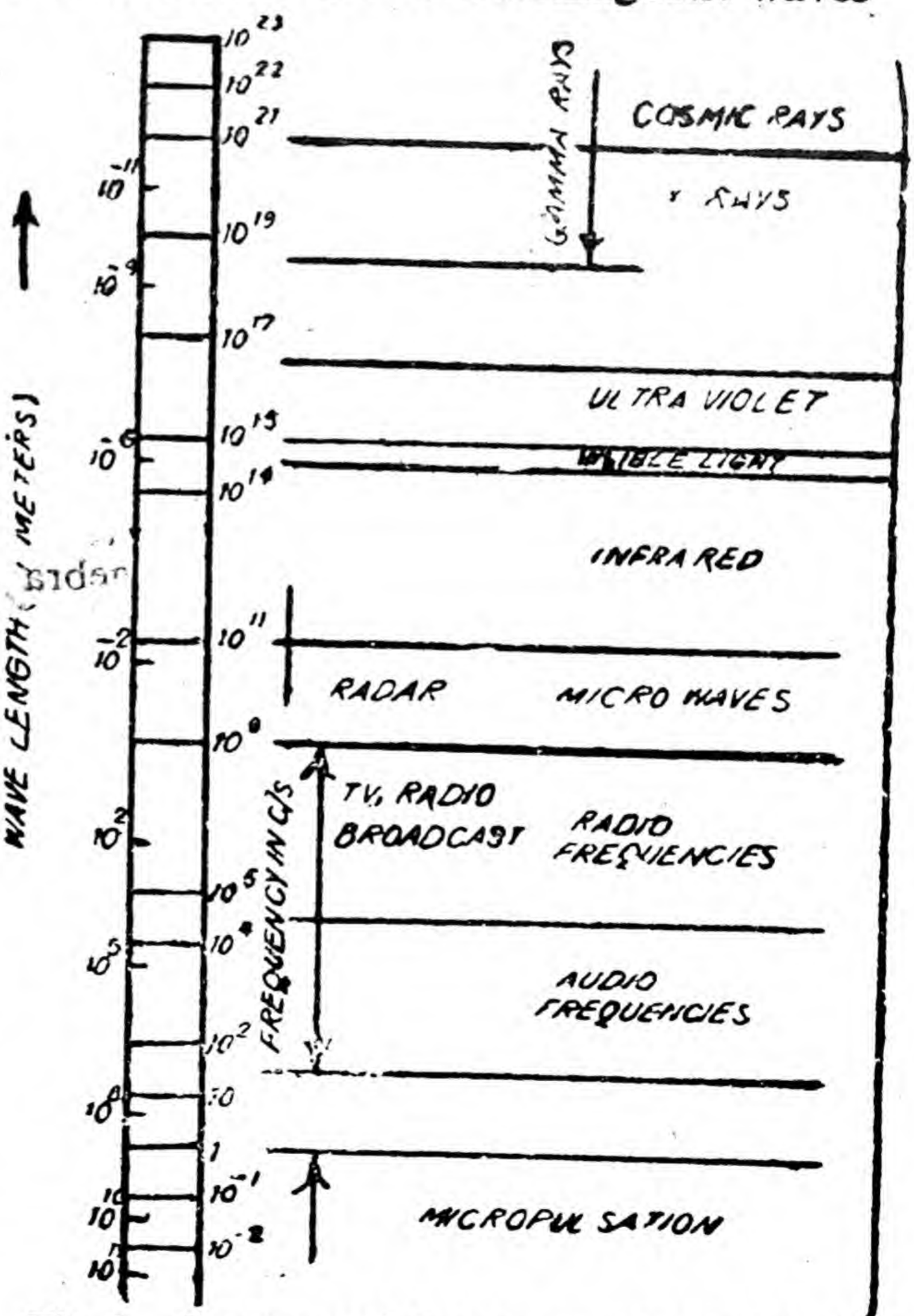


Fig. 1. Spectrum of electromagnetic radiation.

(1) *Radio frequency waves* : They have frequency range from a few c/s upto 10^9 c/s and are generated by electronic devices,

mainly oscillating circuits. The waves are used in television and radio broadcast systems. Classification of electromagnetic spectrum used in radio communication is depicted in the table separately.

(2) *Microwaves* : The frequency range is from 10^9 c/s to 3×10^{11} c/s and are generated by electronic devices such as klystron and magnetron. These waves are used in radar and masers. They are also used to reveal the finer details of atomic and molecular structure.

(3) *Infra red spectrum* : The frequeney range is from 3×10^{11} c/s upto 4×10^{14} c/s and are produced by molecules and hot bodies. These waves find various applications in industry, astronomy, medicine etc.

(4) *Light or visible spectrum* : The frequency range is from 4×10^{14} c/s upto 8×10^{14} c/s and are produced by atoms and molecules as a result of internal adjustment in the motion of their components, principally that of the electrons. This region is used in laser technology.

(5) *Ultraviolet rays* : The frequency range is from 8×10^{14} c/s to about 3×10^{17} c/s and are produced by atoms and molecules in electric discharges. These ultra violet radiation of sun interacts with atoms in upper atmosphere, producing a large number of ions. This result in the formation of ionosphere that extends from 50 Km. to about 1000 Km. from earth's surface. Ultraviolet rays are also used in some medical applications.

(6) *X-rays* : The frequency range is from 3×10^{17} c/s upto 5×10^{19} c/s and are produced by the inner, or more tightly bound electrons in atoms. They are found quite helpful in medical diagnosis.

(7) *Gamma rays* : The frequency range is from 3×10^{18} c/s upto more that 3×10^{22} c/s and are of nuclear origin. They are produced by many radioactive substances. The waves having still higher frequencies are cosmic rays and are of special interest in astronomical research.

1·1. ELECTRONIC DEVICES AT HIGH FREQUENCY RANGES :

Whole of the study of electronics is divided according to the different frequency ranges required for the various particular uses *i.e.* the choice of electronic circuitry for generation, amplification, transmission and reception depends largely upon the frequency of electromagnetic waves in operation.

At high frequencies wavelength being sufficiently small, usual inductance coils and condensers cease to become useful circuit elements. With the rise in frequency, the distributed capacitance existing between the turns of a coil offers a low impedance. Consequently, the coil acts more as a complex transmission line

rather than as a lumped element with constant inductance. Further, at high frequencies, losses are increased to an appreciable proportion due to skin effect, proximity effects and radiation of energy. Condensers also meet the same fate because of distributed inductance that arises due to linkage of magnetic flux.

At high frequencies, conventional tubes do not operate favourably and efficiently. In case of amplifier gain decreases ; input impedance and the maximum impedance that can be realised in the plate circuit also decreases. When valves operate as oscillators, even output decreases : they become less efficient and output obtained is limited by both plate dissipation and grid dissipation at ultra high frequencies. The circuit reactance limitations, circuit and tube loss limitations and electron transit time limitations are responsible for above effects in conventional tubes.

For transmission purposes, parallel wire lines are used at frequencies of the order of 100 mc/s. Above 200 mc/s power loss due to radiation and induction becomes excessive. Coaxial lines are best suited above 200 mc/s upto 3000 mc/s because their attenuation even at very high frequencies is very low. In microwave region waveguides are most suited. Their peak power capacity is greater than of coaxial lines.

For radiation purposes, at radio frequency range short vertical grounded or remote from ground aerials are used. Upto 1000 mc s arrays of antennas are commonly employed. Reflectors and lenses are most common above 1000 mc/s. although parabolic reflectors are fairly common down to atleast 1000 mc/s. Lenses are basically microwave devices not ordinarily used below about 3000 mc/s. Above 10,000 mc/s lenses are most commonly used.

For propagation purposes, at high frequencies upto 30 mc/s sky waves are used for long distance communication. At such frequencies ground wave attenuation is very large and therefore ground wave propagation is limited to very short range. At 300 mc/s ground waves because of their high attenuation and sky waves because of their failure to be reflected almost fail to propagate the signals. Space wave or line of sight propagation is practically the only mode of propagation in this frequency range. Due to refraction in earth's atmosphere propagation beyond the line of sight is possible. Above 300 mc/s only space wave propagation is possible.

1·2. NEED OF HIGH FREQUENCIES :

For radio communication, a wide range of frequencies is required in order to impart appreciable amount of intelligence to the electrical system in the communication process.

For better reception of energy flux at the receiving area, the radiation efficiency of the conductor should be sufficiently high A conductor will radiate sufficient amount of energy, if its dimension is comparable to wavelength of energy wave. The size of the

Table : *Classification of Electromagnetic Spectrum used in Radio Communication* :

Carrier frequency (KC/S)	Free Space wavelength (meters)	Class	Principal communication uses	Propagation characteristics
10—30	3×10^4—10^4	Very low freq. (V. L. F)	Long distance point to point communication.	At all times of day and of year, attenuation is quite low.
30—300	10^4—10^3	Low freq. (L. F.)	Navigational services, point to point (long distance) services.	At day time absorption is greater as compared to V L.F. The night propagation is quite similar to V.L.F.
300—3000	10^3-10^2	Medium freq. (M. F.)	Broadcasting, ship-harbour police.	At day time attenuation is high while low at night.
3×10^3—3×10^4	10^2—10	High freq. (H. F.)	International broadcasting, government, aviation.	Transmission over considerable distances depends upon the ionosphere. Therefore propagation characteristics will vary as the density of ionised region changes *i.e.*, at day and night times.
3×10^4—3×10^5	10—1	Very high frequency (V. H. F.)	Radar, television, short distance communication.	These are unaffected by the ionosphere and are used for straight line propagation.
3×10^5—3×10^6	1·0—0·1	Ultra high frequency (U. H. F.)	Facsimile, air navigation, television, relay system.	Same.
3×10^6—3×10^7	0·1—0·01	Super high frequency (S. H. F.)	Radio relay, radar navigation.	Same.

conductor is inversely proportional to the wave frequency. High frequency waves can therefore, be radiated by a small radiator, while low frequency waves require a large antenna system for effective radiation.

Another reason for high frequency transmission is to have better signal selection at the receiving point. In higher frequency range, the percentage with which the components carrying intelligence extend above and below the mean frequency of the wave is relatively small. The selectivity, therefore, increases because individual narrow bands of frequency spectrum will be assigned to different transmissions and a desired transmission may be selected for reception on the basis of discrete band occupied by its frequency components.

1·3. MODULATION METHODS :

These are the methods used to transmit the intelligence at radio frequencies. Some characteristic of the wave is varied inaccordance with the information to be transmitted. The proper combination of the carrier and the signal waves, such that the carrier is able to convey the intelligence through the transmission system is called a "modulated" carrier wave.

One way of modulation is termed as *amplitude modulation*. In this process the amplitude of the carrier is varied inaccordance with the amplitude and frequency of the modulating signal. Frequency of carrier remains constant. The amplitude modulated wave is shown in figure 2 (c). The resultant wave consists of two side bands appearing above and below the carrier frequency. Each of these side bands consists complete information. The side band frequencies can be considered as being generated as a result of varying the wave. The magnitude and frequency are determined by the character of modulation.

Another way of modulation is to alter the frequency of the carrier for the purpose of intelligence transmission, termed as *frequency modulation*. In this way, the amplitude for the carrier is kept essentially constant and the frequency is varied above and

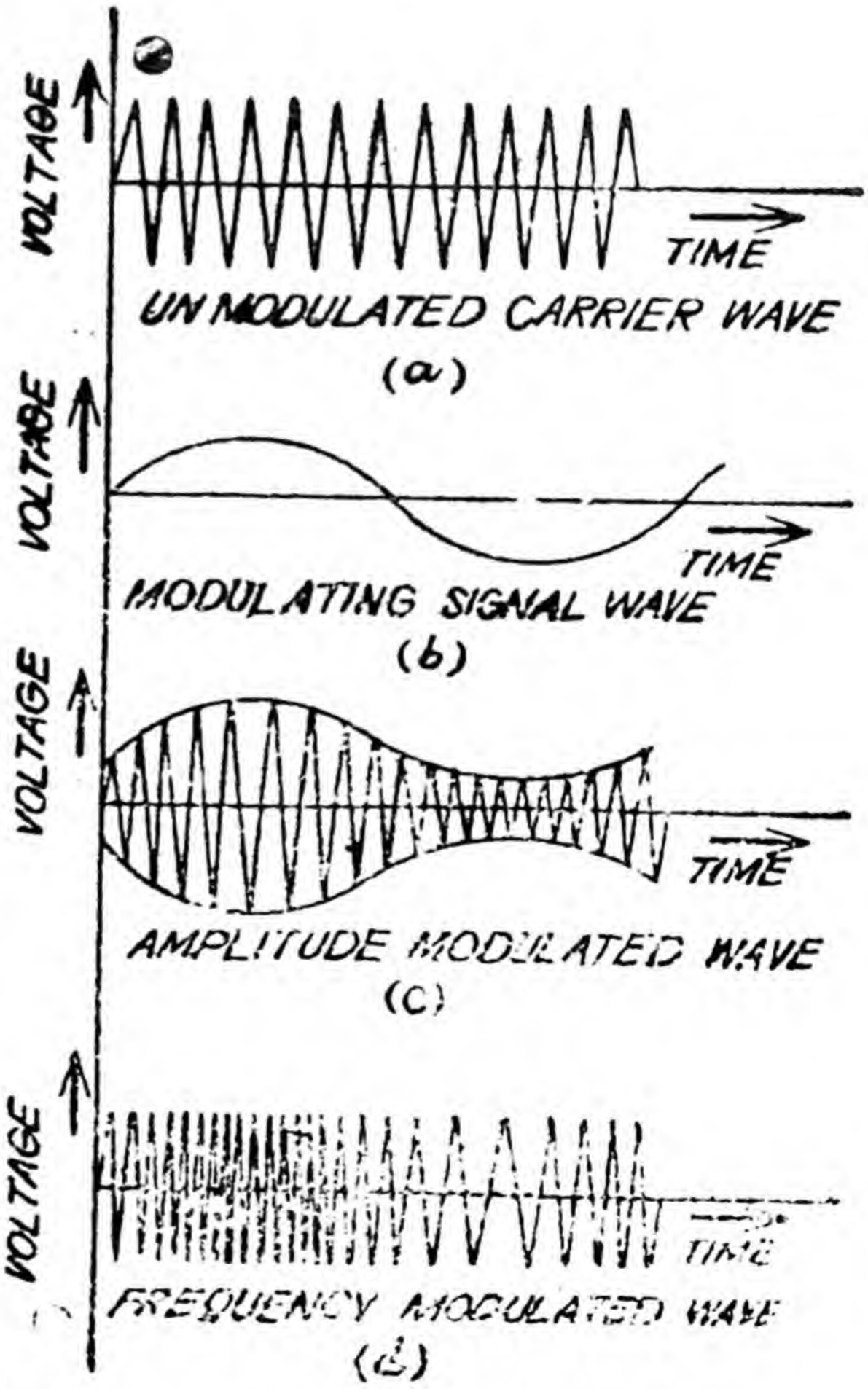


Fig. 2. Amplitude and frequency modulated waves along with carrier wave and modulating signal

below its original value inaccordance with the vibration of the signal. A frequency modulated wave is shown in fig. 2 (d). The number of times per second that the carrier wave goes through complete cycle of frequency variation is equal to the frequency of modulating signal and the range of frequency deviation is proportional to the amplitude of the modulating signal and is independent of signal frequency. For example, if we consider that a 100 mega cycle carrier being modulated by a sinusoidal signal and if there is 2 k.c. frequency shift per volt of signal amplitude, then for a modulating signal having an amplitude of 25 volts and a frequency of 1 k.c., the instantaneous frequency of the modulated wave will vary between 99·95 megacycles and 100·05 mega-cycles at rate of 1000 times per second.

1·4. RECEPTION OF RADIO SIGNALS :

The first step in the reception of radio signals is to abstract energy from the radio wave passing the receiving point. When antenna conductor cuts the electromagnetic flux of the wave, a voltage is induced This induced voltage develops a current in the conductor having its time variations exactly same as radiating antenna.

The second step is to select the wanted signal and discriminate the signal of other frequencies. The ability is called *selectivity* and is achieved by tuning of resonant circuit used in the receiving system. Since energy may be received after travelling thousands of miles, it is preferrable to amplify the received energy for satisfactory reception.

The third step is to recover the original modulating signal from the modulated wave. The process is called *detection*. In the case of amplitude modulation, a simple rectification of modulated wave serves the purpose. For detection ot f. m. wave, it is first transmitted through a circuit in which the relative response depends upon frequency. The wave thus emerging from this circuit is amplitude modulated which is then rectified.

1·5. GENERAL PICTURE OF RADIO TRANSMISSION AND RECEPTION :

Figure 3 gives a clear picture of the steps involved at the transmitting and receiving stages. The next chapters of the book deal with details of each of these steps.

The first step at the transmitter is to convert audible sound waves into electrical impulses by microphone. The impulses are then amplified by audio amplifier. For the transmission of these impulses, modulation is performed. The carrier wave 550 k.c. is generated by r.f. oscillator. The output modulator is applied to a transmitting antenna. The modulated carrier wave is thus sent out into space by the antenna.

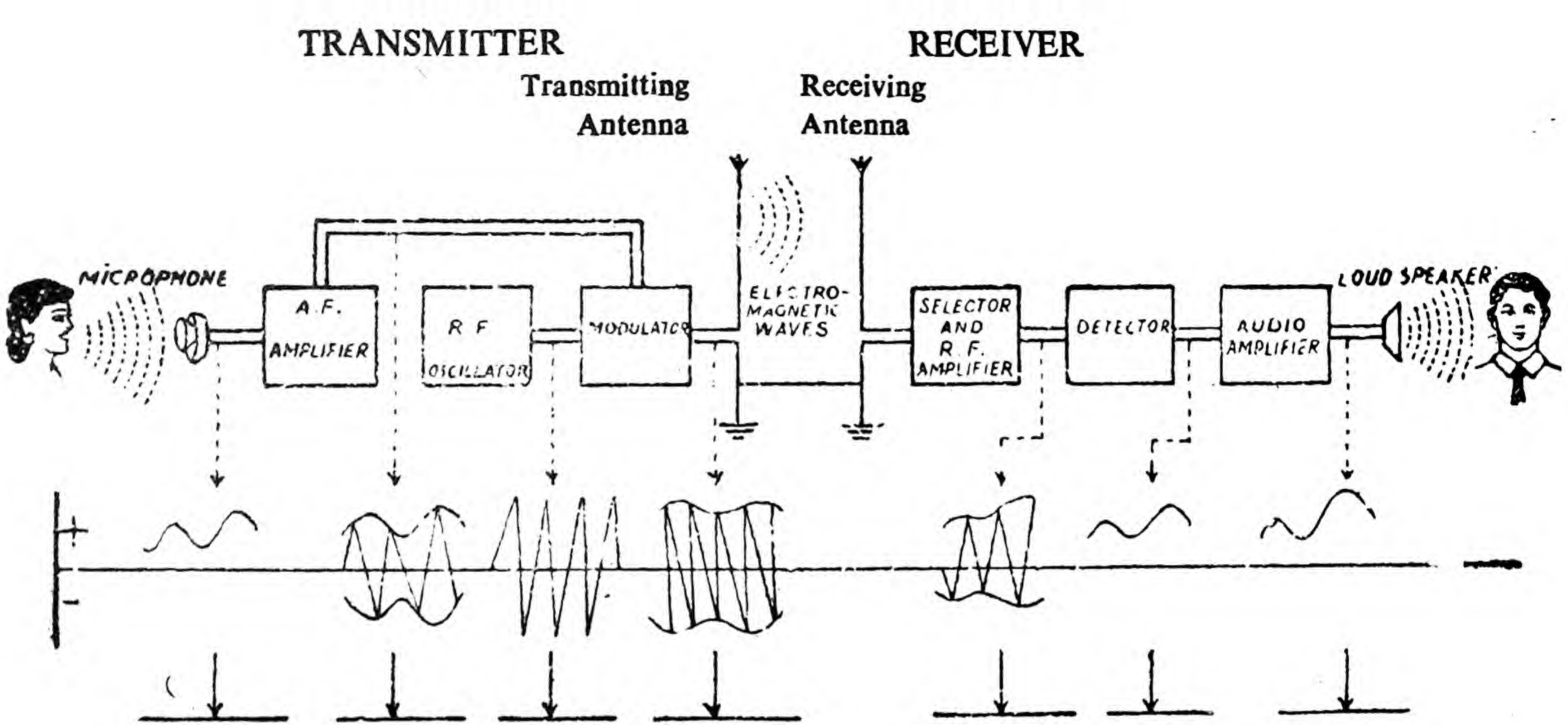


Fig. 3. Illustration of various steps in transmitting and receiving stages.

At the receiver, a voltage is induced in the antenna when the latter cuts these electromagnetic waves radiated by the transmitting antenna. The proper frequency is chosn by means of tuner. The so received modulated wave being weak, is amplified by r.f. amplifier. The output of r.f. amplifier is coupled to the detector stage at which audio signal is separated from the carrier. Audio signal is then amplified so as to operate the loudspeaker.

2 RESONANT CIRCUITS

In radio system several electrical circuits are used to perform various functions. The property of the system depends upon the operational characteristics of these electrical circuits. In this section we shall discuss the selective circuits.

2·1. SELECTIVE CIRCUITS :

To achieve desired selectivity, resonant circuits are employed. The operational characteristics of these circuits depend upon whether the inductance and capacitance are in series or in parallel. We, therefore, classify the resonant circuits into the following categories :

(*i*) *Series resonant circuit.*

(*ii*) *Parallel resonant circuit.*

2·2. SERIES RESONANT CIRCUIT :

A series circuit containing *LCR* is shown in fig. 1. When the resultant reactance is zero, the circuit is said to be in resonance. The voltage drop across the inductance leads the current by 90°, whereas across condenser lags the current by 90°. If the two drops are made equal, as shown in fig. 2, then the total impedance of the

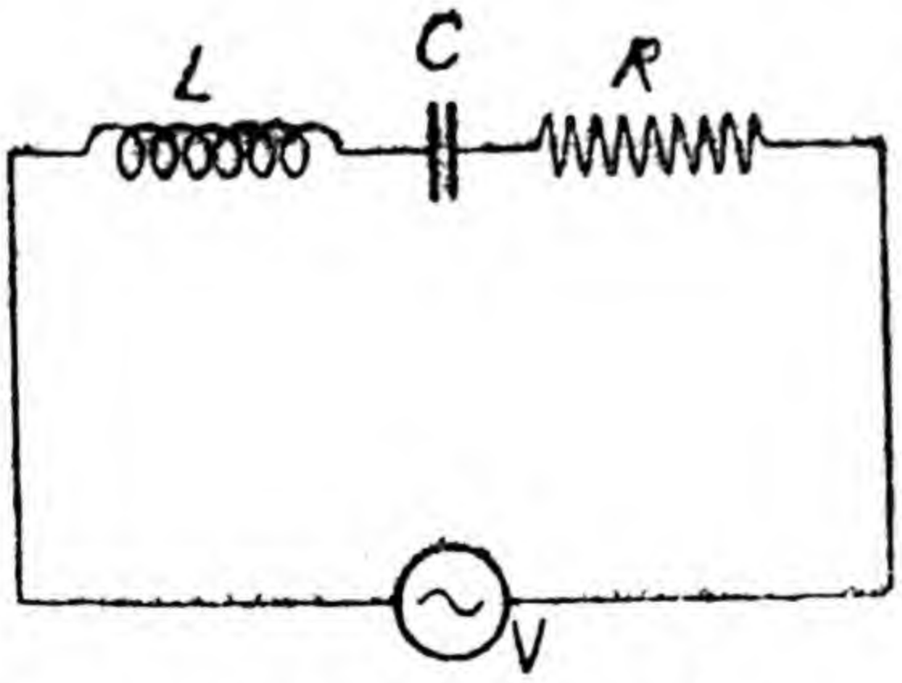


Fig. 1. Series *LCR* circuit.

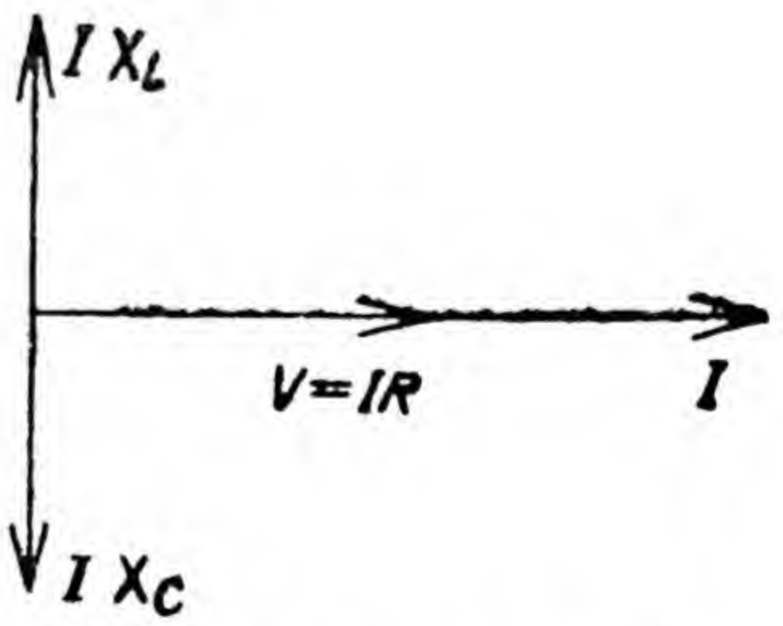


Fig. 2. Vector diagram of series circuit in resonance.

circuit will be only resistive and the impressed voltage is equal to resistive drop. This condition is called *series resonance*. Thus for resonance.

$$I \cdot X_L = I \cdot X_C \quad \text{or} \quad X_L = X_C.$$

where X_L and X_C are the reactances of inductance L and capacitance C. Also

$$2\pi f.L=\frac{1}{2\pi fC} \quad \therefore \quad f=\frac{1}{2\pi\sqrt{(LC)}}=f_0, \qquad \ldots(1)$$

f_0 is resonant frequency in c/s., where L is in henrys, and C in farad.

From above relation, it is obvious that the resonance can be obtained by varying either of L, C and f in the series circuit.

Current I in series LCR circuit can be written as :

$$I=\frac{V}{Z}=\frac{V}{\sqrt{\{R^2+(X_L-X_C)^2\}}}. \qquad \ldots(2)$$

The voltage drop across three components will be :

drop across resistance

$$V_R=I\ R=\frac{VR}{\sqrt{\{R^2+(X_L-X_C)^2\}}} \qquad \ldots(3)$$

drop across inductance

$$V_L=I\ X_L=\frac{VX_L}{\sqrt{\{R^2+(X_L-X_C)^2\}}} \qquad \ldots(4)$$

drop across capacitance

$$V_C=I\ X_C=\frac{VX_C}{\sqrt{\{R^2+(X_L-X_C)^2\}}} \qquad \ldots(5)$$

Resonance curves : At resonance the power factor is unity, and current $I_0=\frac{V}{R}$. In figure 3, the shape of current curve before, at, and after resonance is shown. It is obvious from relation for I_0 that if at resonance resistance is doubled, current will be half the previous value. We see that with small resistance, current peak is sharper than the current peak when resistance is large. *Therefore, small resistance is said to give sharp tuning whereas large resistance gives broad tuning.*

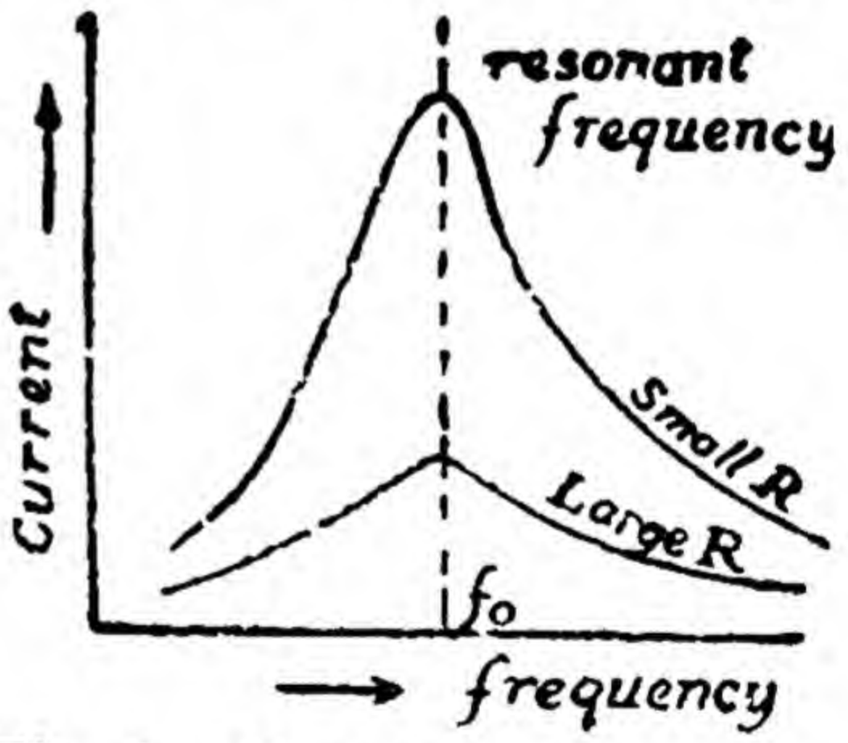


Fig. 3. Graph representing variation of current with frequency.

Now we shall consider various ways of securing resonance.

(a) Varying inductance : When L is varied to produce resonance, a series is obtained as shown in fig. 4.

Figure shows that V_C is maximum at resonance while V_L, the drop across inductance, rises to its maximum value after resonance. This behaviour is inaccordance with our expectation because $V_C=IX_C$, X_C is constant, therefore V_C will be maximum when

I is maximum, *i.e.*, at resonance. In the case of $V_L = I.X_L$ both I and X_L are increasing before resonance, therefore V_L is increasing. At resonance, I is not changing, but X_L is increasing ; hence drop is increasing. The drop continues to increase until the reduction in current affects the increase in L, that is X_L. This point can be obtained by differentiating V_L with respect to X_L and equating it to zero,

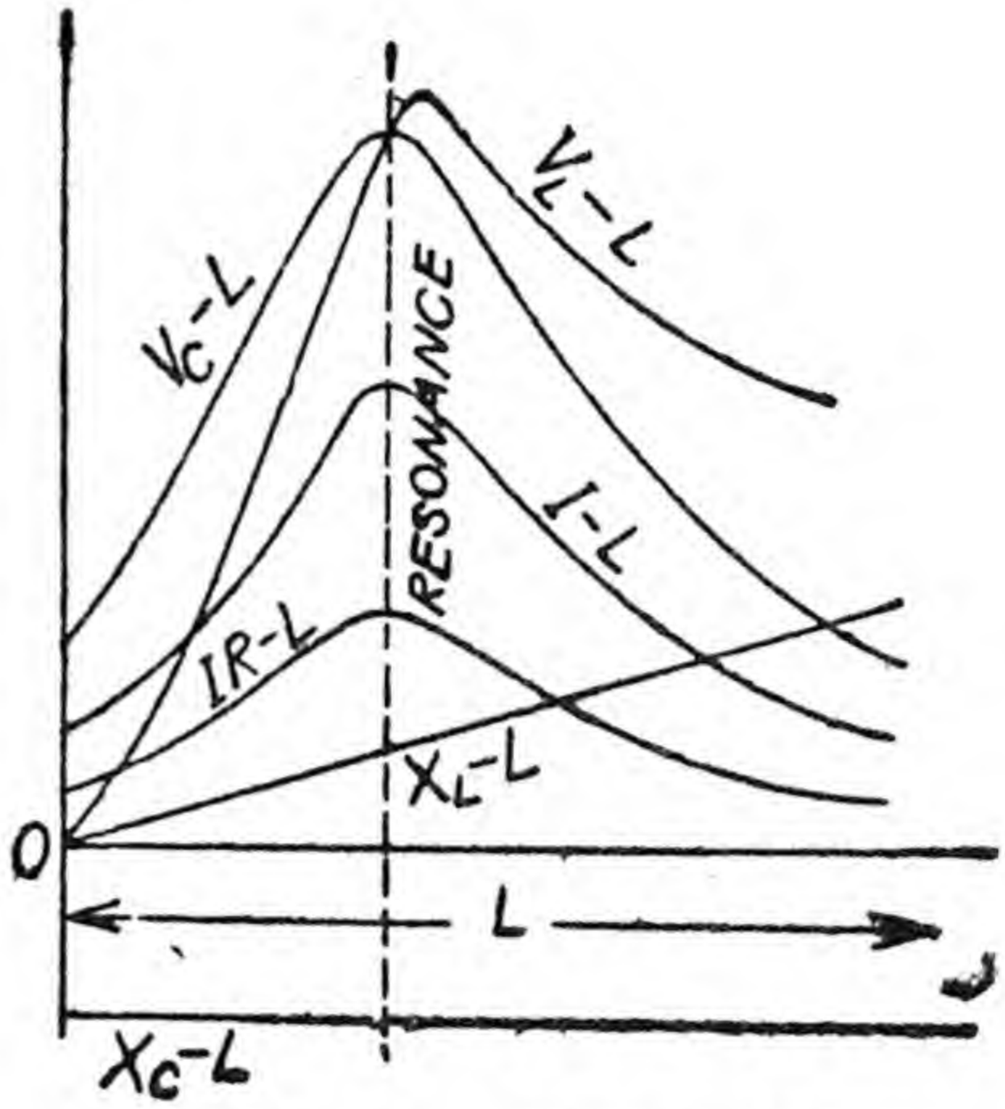


Fig. 4. Series resonance curves by varying L.

$$\text{i.e. } \frac{dV_L}{dX_L} = \frac{d}{dX_L}\left[\frac{VX_L}{\{R^2+(X_L-X_C)^2\}^{1/2}}\right]=0$$

$$= \frac{[R^2+(X_L-X_C)^2]^{1/2}\, V - VX_L \cdot \frac{1}{2}\,[R^2+(X_L-X_C)^2]^{-1/2}.2\,(X_L-X_C)}{R^2+(X_L-X_C)^2}$$

$$=0$$

or
$$X_L = \frac{R^2+X_C^2}{X_C},$$

therefore
$$L = \frac{1}{2\pi f}\left[\frac{R^2+X_C^2}{X_C}\right] = C\,(R^2+X_C^2) \qquad \ldots(6)$$

(b) Varying capacitance. When C is varied to produce resonance, curves as shown in fig. 5, are obtained.

In this case V_L will be maximum at resonance, since X_L is constant. V_C attains maximum value before resonance because at resonance, X_C is decreasing whereas the current is not changing. Therefore drop $I \cdot X_C$ must be decreasing ; consequently, the drop must have been maximum before resonance. At resonance the drops across L and C are equal and opposite. The condition for maximum V_C can be determined by differentiating equation (5) with respect to X_C and then equating it to zero :

$$\frac{dV_C}{dX_C} = \frac{d}{dX_C}\left[\frac{VX_C}{[R^2+(X_L-X_C)^2]^{1/2}}\right]$$

$$=0$$

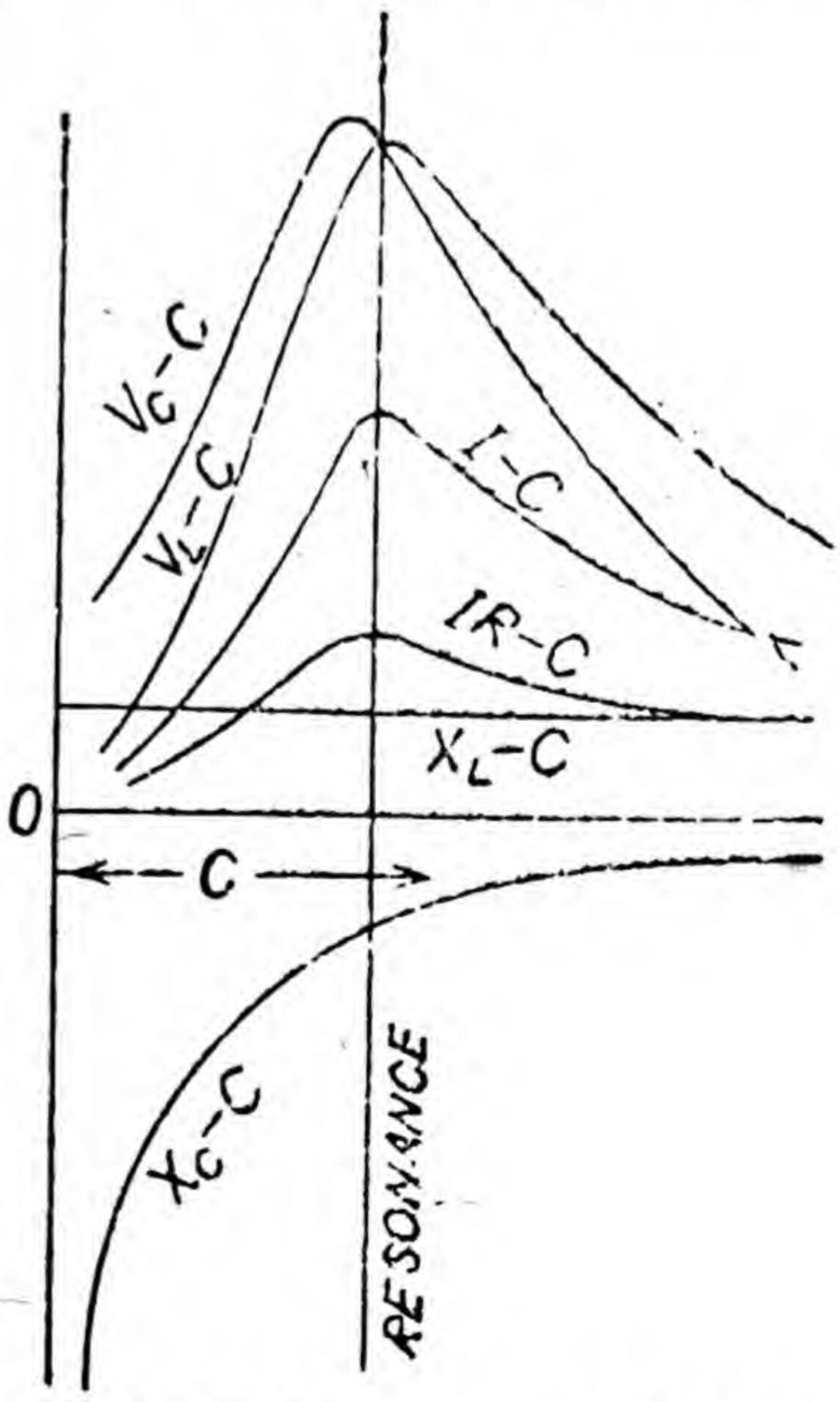


Fig. 5. Series resonance curves by varying capacitance.

$$=\frac{[R^2+(X_L-X_C)^2]^{1/2}.V+VX_C.\frac{1}{2}[R^2+(X_L-X_C)^2]^{1/2}.2(X_L-X_C)}{R^2+(X_L-X_C)^2}$$

$=0$

or $$X_C=\frac{R^2+X_L^2}{X_L}$$

or $$\frac{1}{C}=\frac{1}{L}(R^2+X_L^2) \quad \therefore \quad C=\frac{L}{(R^2+X_L^2)} \qquad ...(7)$$

The impressed voltage equals IR drop, *the power factor is unity and the current is maximum at the resonance.* For zero capacitance, the capacitive reactance is infinite and the current is, therefore, zero. The power factor varies from $\frac{R}{\sqrt{(R^2+X_L^2)}}$ when C is infinite, to zero when C is zero.

Since it is quite easy to vary capacitance gradually and smoothly, resonance is usually obtained by varying capacitance C.

(c) Varying frequency : When frequency is varied then the resonance curves obtained are shown in fig. 6.

This case includes both the above cases of varying inductance and capacitance. Neither V_C nor V_L is maximum at resonance. The current is zero for both zero frequency and infinite frequency.

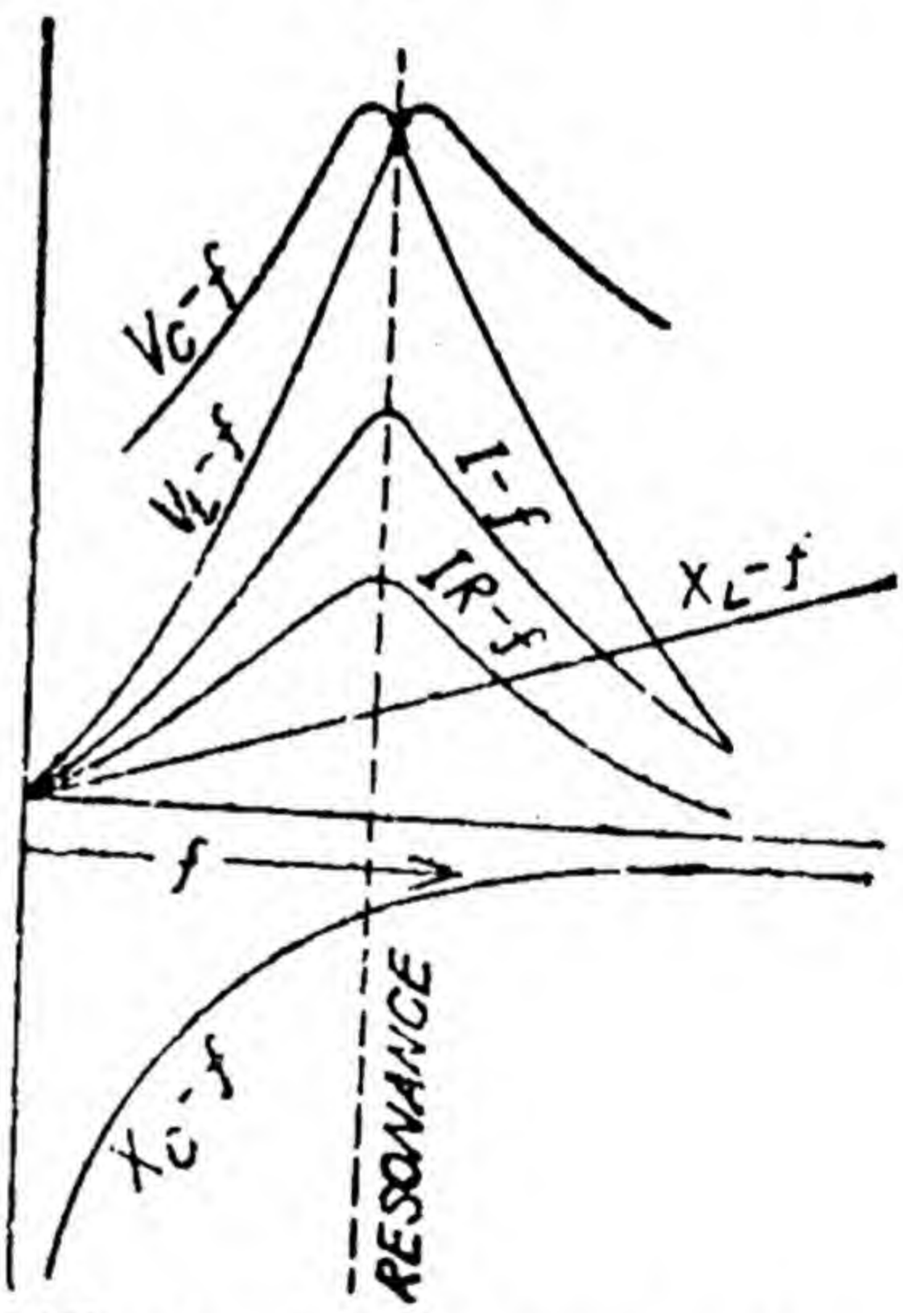


Fig. 6. Series resonance by varying frequency.

It will be observed that for all methods of producing resonance :

(*i*) The current is maximum and dependent only upon the impressed voltage and the resistance of the circuit,

(*ii*) the power factor is unity,

(*iii*) the power is maximum and equal to the voltampere (VI) at the point of resonance.

2·2. SERIES CIRCUIT AS A SELECTOR :

The passage of frequency through a circuit depends upon the magnitude of impedance offered to it by the circuit. As shown in figure 7, the circuit has lower impedance for the resonant frequency. Therefore, LCR circuit passes frequencies near the resonance frequency more readily than other frequencies. This shows its selective nature. The band of frequencies which has passed quite readily is called the *pass band.* The pass band is sometimes arbitrarily considered to be the range of frequency over which the current is equal to or greater than $\frac{V}{\sqrt{2}R}$ as shown in figure 7.

The relation for current is

$$I=\frac{V}{\sqrt{\{R^2+(\omega L-1/\omega C)^2\}}}$$

If ω_P be the angular frequency corresponding to

$$*I=\frac{V}{\sqrt{2}.R},$$

then

$$\frac{V}{\sqrt{2}R}=I=\frac{V}{\sqrt{\left\{R^2+\left(\omega_P L-\frac{1}{\omega_P C}\right)^2\right\}}}$$

or $$\sqrt{2}R=\sqrt{\left\{R^2+\left(\omega_P L-\frac{1}{\omega_P C}\right)^2\right\}}$$

or $$\left(\omega_P L-\frac{1}{\omega_P C}\right)^2=R^2 \quad \text{or} \quad \omega_P L-\frac{1}{\omega_P C}=\pm R$$

or $$\omega_P{}^2\pm\frac{R}{L}\omega_P-\frac{1}{LC}=0.$$

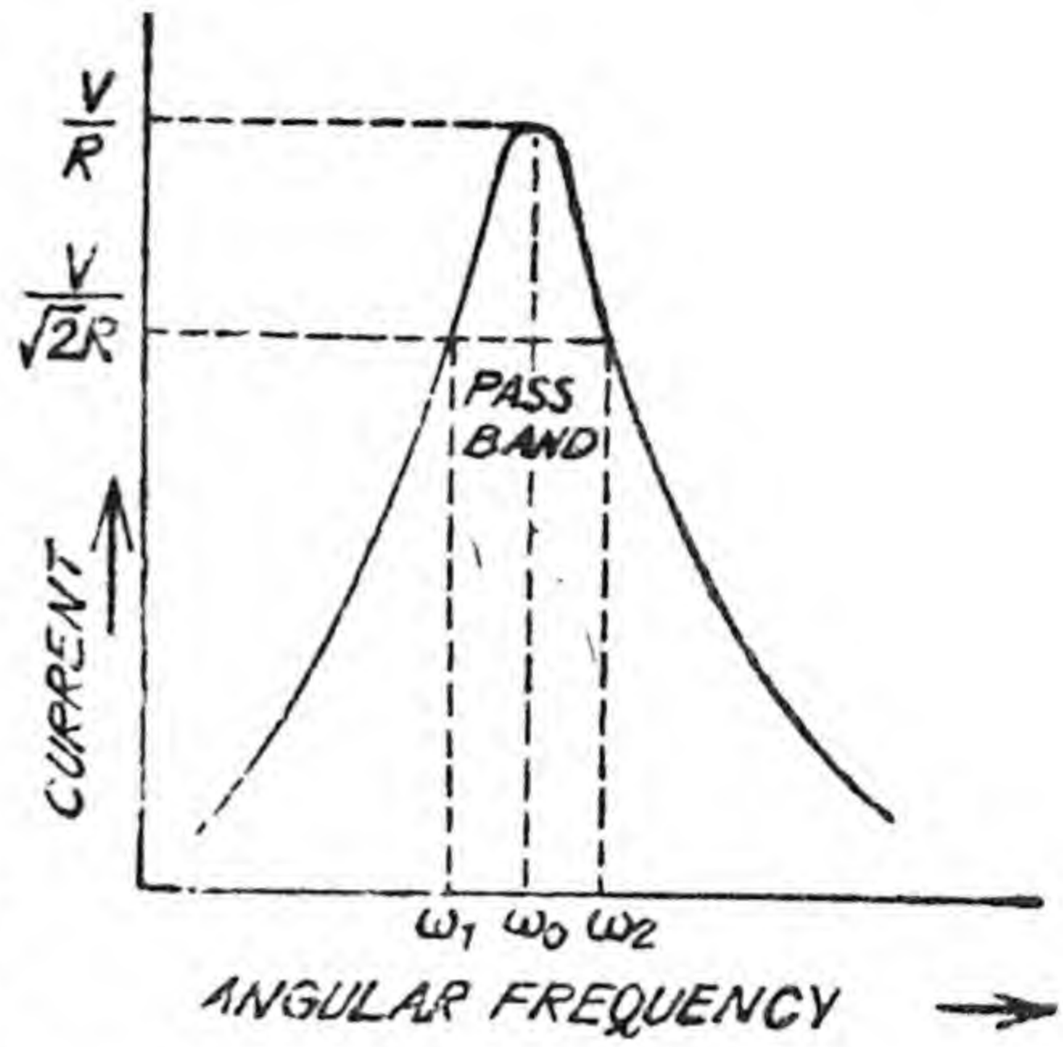


Fig. 7. Series *RLC* circuit as a band selector.

When solved for ω_P above equation yields

$$\omega_P=\frac{\pm R/L\pm\sqrt{(R^2/L^2+4/LC)}}{2},$$

$$\omega_P=\pm\frac{R}{2L}\pm\sqrt{\left(\frac{R^2}{4L^2}+\frac{1}{LC}\right)}$$

In a selective *RLC* branch $\left(\frac{R}{2L}\right)^2$ is usually much smaller than $1/LC$. Hence the above relation can be written as

$$\omega_P=\pm\frac{R}{2L}\pm\sqrt{\left(\frac{1}{LC}\right)}.$$

but $$\frac{1}{LC}=\omega_0{}^2=\text{resonant frequency.}$$

Therefore $$\omega_P=\pm\frac{R}{2L}\pm\omega_0$$

If only positive values of ω_0 are considered, then

$$\omega_P=\omega_0\pm\frac{R}{2L}.$$

Let $$\omega_1=\omega_0-\frac{R}{2L}, \qquad \omega_2=\omega_0+\frac{R}{2L}.$$

Then the width of pass band will be

---

*The points at which the current assumes the value are called half power points because, at these points, the power is exactly one half the maximum power which occurs at resonance.

$$\Delta\omega=\omega_2-\omega_1=\frac{R}{L}\text{rad./sec.}$$

or
$$\Delta f=f_2-f_1=\frac{R}{2\pi L}\text{ c/s.} \qquad \ldots(1)$$

Therefore the frequency range for the pass band as defined here is $\Delta f=f_2-f_1=R/2\pi L$ c/s.

2·4. THE Q OF SERIES CIRCUIT :

The degree of selectivity of a circuit, that is, the narrowness of the bandwidth is usually expressed in terms of the symbol Q, and is expressed as

$$Q=\frac{\omega_0}{\omega_2-\omega_1}=\frac{\omega_0}{\Delta\omega}=\frac{f_0}{\Delta f} \qquad \ldots(1)$$

In the case of series circuit

$$Q_{series}=\frac{\omega_0}{\Delta\omega}=\frac{\omega_0}{R_s/L}=\frac{\omega_0 L}{R_s}$$

$$=\frac{1}{\omega_0 C R_s}=\frac{1}{\frac{1}{\sqrt{(LC)}}\cdot R_s\cdot C}=\frac{1}{R_s}\sqrt{\left(\frac{L}{C}\right)}, \qquad \ldots(2)$$

where R_s is the total equivalent series resistance of the circuit.

$$Q_{series}=\frac{\omega_0 L}{R_s}=\frac{\omega_0 L I_0}{R_s I}$$

$$=\frac{\text{Voltage drop across } L}{\text{Voltage drop across } R_s \text{ (impressed voltage)}}$$

Also
$$Q_{series}=\frac{1}{\omega_0 C R_s}=\frac{1}{R_s.I_0}\times\frac{I_0}{\omega_0 C}=\frac{I_0/\omega_0 C}{R_s.I_0}$$

$$=\frac{\text{Voltage drop across } C}{\text{Applied voltage}}.$$

This Q_{series} is multiple of the applied circuit voltage that will exist across each of the reactive components at resonance Q_s, bears its own significance in radio frequency circuits where Q_s is essentially constant than in low frequency circuits where R_s is essentially constant.

Example. *To calculate the width of passband of a series LCR circuit with*

$Q=120,$
$L=150.10^{-6}$ henry
$C=200$ pf.

Width of passband $=f_2-f_1=\frac{R}{2\pi L}$ c/s.

and
$$Q=\frac{L\omega_0}{R}=\frac{L}{R}\frac{1}{\sqrt{(LC)}}$$

$$\therefore \quad R=\frac{L}{Q}\cdot\frac{1}{\sqrt{(LC)}}$$

$$=\frac{150.10^{-6}}{120}\cdot\frac{1}{\sqrt{(150.10^{-6}.200.10^{-12})}}\ \text{ohms}=\frac{150}{12\sqrt{(3)}}\ \text{ohms.}$$

$$\therefore \quad f_2-f_1=\frac{150}{12\sqrt{3.2.3\cdot14.150.10^{-6}}}=7\cdot663\ \text{kc./s.}$$

2·5. PARALLEL RESONANT CIRCUIT :

A parallel circuit consisting of a coil associated with a resistance R in parallel with a capacitance is shown in figure 8. Resistance associated with capacitance is assumed to be negligible.

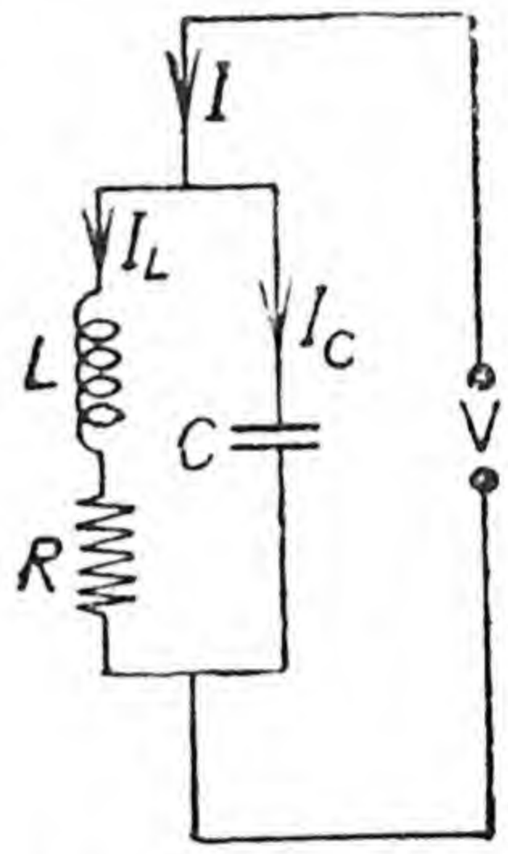


Fig. 8. Parallel resonant circuit.

An expression for the importance of this parallel circuit can be derived as follows :

Let Z be total impedance of the circuit, Z_R, Z_L, Z_C be the impedances of resistive, inductive and capacitive branches ; then we can write

$$Z=\frac{(Z_R+Z_L)\,Z_C}{Z_R+Z_L+Z_C}=\frac{(R+j\omega L)\,(-j/\omega C)}{R+j\,(\omega L-1/\omega C)}$$

$$=\frac{(R+j\omega L)\,(-j/\omega C)[R-j\,(\omega L-1/\omega C)]}{R^2+\left(\omega L-\frac{1}{\omega C}\right)^2}$$

$$=\frac{\frac{R}{\omega^2C^2}+\frac{j\omega}{\omega^2C^2}\{L-C\,(\omega^2L^2+R^2)\}}{R^2+\left(\omega L-\frac{1}{\omega C}\right)}$$

$$=\frac{R+j\omega\,[L-C\,(\omega^2L^2+R^2)]}{\omega^2C^2R^2+(\omega^2LC-1)^2}.$$

Separating the real and imaginary parts, we get

effective resistance $R_e=\dfrac{R}{\omega^2C^2R^2+(\omega^2LC-1)^2}$.

effective reactance $X_e=\dfrac{\{L-C\,(\omega^2L^2+R^2)\}\ \omega}{\omega^2C^2R^2+(\omega^2LC-1)^2}$.

The magnitude of total impedance Z can be written as

$$Z=\sqrt{(R_e^2+X_e^2)}=\frac{\sqrt{[R^2+\omega^2\,\{L-C(\omega^2L^2+R^2)\}^2]}}{\omega^2C^2R^2+(\omega^2LC-1)^2}$$

$$=\sqrt{\left[\frac{R^2+\omega^2L^2}{\omega^2C^2R^2+(\omega^2LC-1)^2}\right]} \qquad \ldots(1)$$

$$=\frac{R^2+\omega^2L^2}{\sqrt{[\{\omega^2C^2R^2+\ \omega^2LC-1)^2\}\ (R^2+\omega^2L^2)]}}$$

$$=\frac{R^2+\omega^2L^2}{\sqrt{\{R^2+\omega^2\,(CR^2+\omega^2L^2C-L)^2\}}}, \qquad \ldots(2)$$

The admittance $Y=1/Z$ will be minimum when in equation (2)

$$CR^2+\omega^2L^2C-L=0.$$

This corresponds to the reactive or quadrature component of admittance vanishing. At such a minimum admittance *i.e.*, maximum impedance, the circuit current will be minimum for an applied e.m.f. This will occur at a resonant frequency

$$f_0=\frac{\omega_0}{2\pi}$$

where $CR^2+\omega_0^2L^2C-L=0,$

$$\therefore\ \omega_0=\sqrt{\left(\frac{1}{LC}-\frac{R^2}{L^2}\right)} \qquad \therefore\ f_0=\frac{1}{2\pi}\sqrt{\left(\frac{1}{LC}-\frac{R^2}{L^2}\right)} \qquad \ldots(3)$$

This type of circuit is known as a *rejector circuit.* Since the reactive component of the admittance has been made zero, the impedance of the circuit at resonance will be purely resistive in its effect. Equation (1) at resonance takes the form,

$$Z_{resonance}=\frac{R^2+\omega_0^2L^2}{R}.$$

This impedance at resonance is called *dynamic resistance.*

Therefore dynamic resistance $=\frac{R^2+\omega_0^2L^2}{R}=R+\frac{L^2}{R}\omega_0^2$

$$=R+\frac{L^2}{R}\left(\frac{1}{LC}-\frac{R^2}{L^2}\right)=\frac{L}{CR}.$$

Peak current from supply source at resonance called the *make up current* will, therefore, be equal to $\frac{V}{L/CR}=\frac{VCR}{L}$. The current in capacitance and inductance branch called *forced oscillatory current* will be equal to the peak current in the capacitance, I_C. But $I_C=V\omega_0C$, so that

$$\frac{\text{Oscillatory current}}{\text{Make up current from supply}}=\frac{V\omega_0C}{VCR}=\frac{L\omega_0}{R}=Q \text{ factor.}$$

Thus the parallel rejector circuit at resonance exhibits current magnification of $\frac{\omega_0L}{R}$, *while series acceptor circuit a voltage magnification of the same ratio.*

The dynamic resistance can also be expressed in terms of Q, *i.e.*, dynamic resistance

$$=\frac{L}{CR}=\frac{L\omega_0}{R}\cdot\frac{1}{\omega_0C}=Q\cdot\frac{1}{\omega_0C}$$

$$=L\omega_0\cdot\frac{1}{\omega_0CR}=L\omega_0Q.$$

Thus at resonance, the impedance of circuit becomes Q times

mpedance of the branch line. Figure 9 shows the resonance curve for the parallel resonant circuit representing a minimum current at the resonance.

Figure 10 shows the magnitude of a parallel circuit impedance as a function of frequency for different circuit Q's.

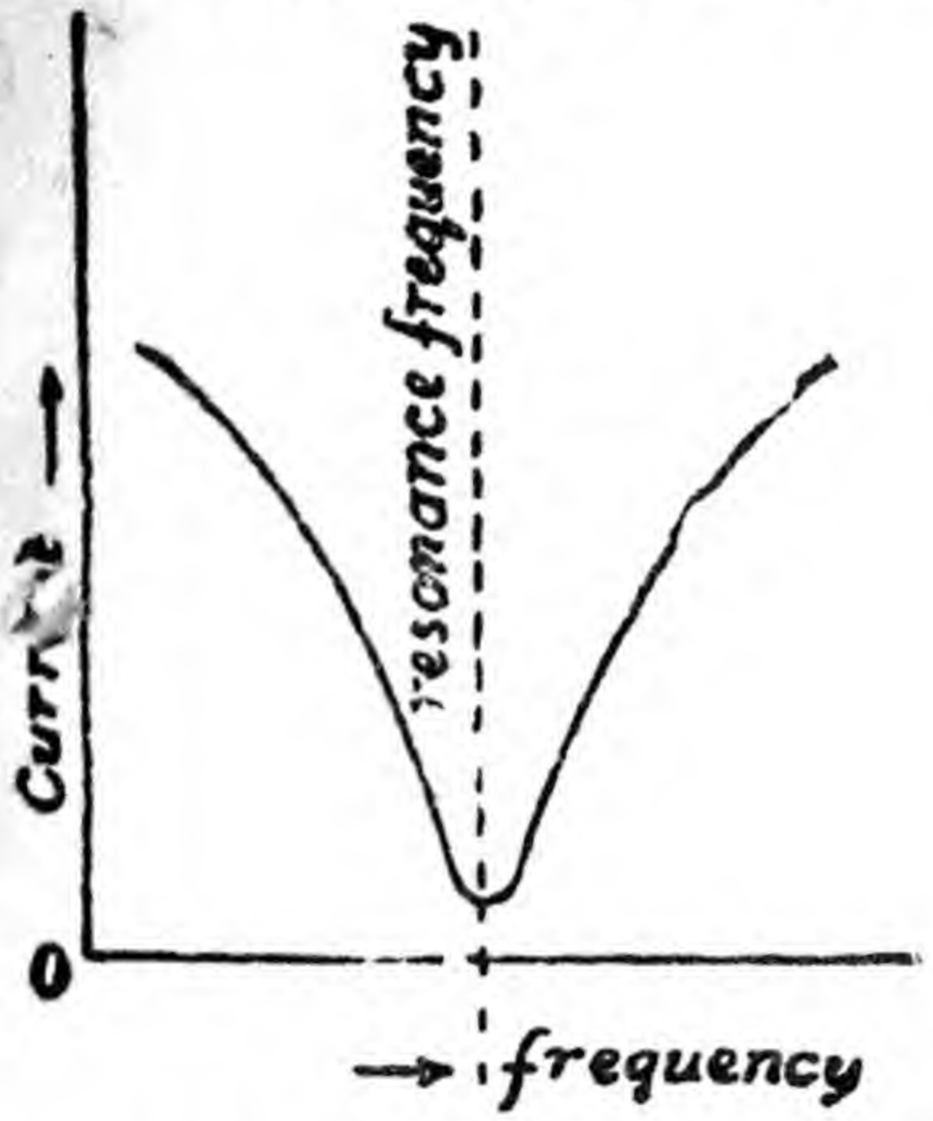


Fig. 9. Resonance curve of parallel circuit.

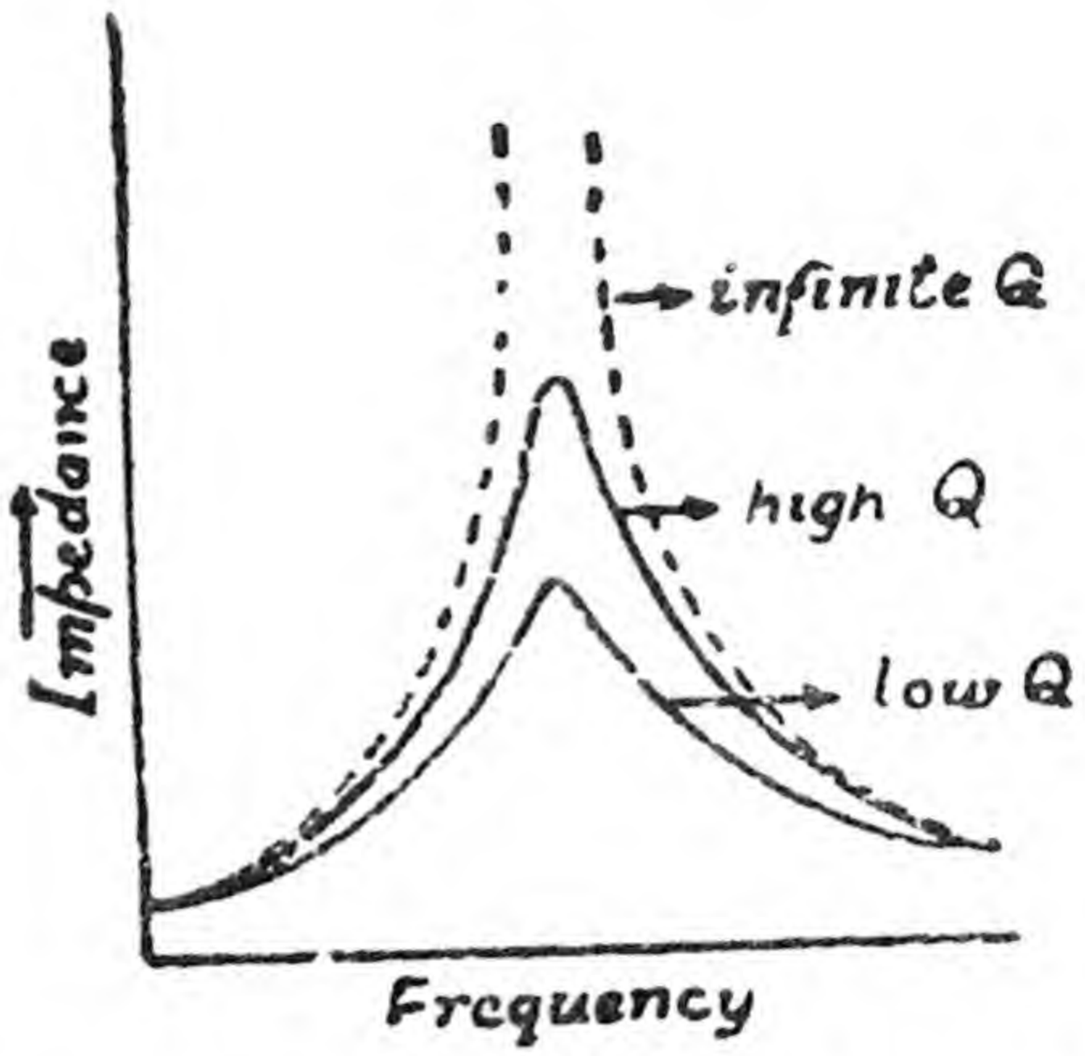


Fig. 10. Variation of impedance of parallel circuit with frequency for different Q values.

We see from equation (3) that the resonance can be obtained by varying either L, C and ω. . As we have discussed with series circuit, the resonance condition can here be found for all the three cases as follows:

(*a*) **Varying inductance :** The frequency for parallel resonance when inductance is varied, can be found by differentiating admittance with respect to inductance and equating it to zero for minimum condition. From equation (1) the square of admittance is

$$|Y|^2=\frac{\omega^2C^2R^2+(\omega^2LC-1)^2}{(R^2+\omega^2L^2)},$$

$$\frac{d|Y|^2}{dL}=\frac{(R^2+\omega^2L^2)\,[0+2(\omega^2LC-1)\,\omega^2C]-2\omega^2L\,[\omega^2C^2R^2+(\omega^2LC-1)^2]}{(R^2+\omega^2L^2)^2}=0$$

$$\text{or}\quad (R^2+\omega^2L^2)(\omega^2LC-1)C-L\,[\omega^2C^2R^2+(\omega^2LC-1)^2]=0$$

$$\text{or}\quad (R^2C+L)(\omega^2LC-1)-L\omega^2C^2R^2=0$$

$$\text{or}\quad -\omega^2L^2C=L+R^2C.$$

$$\therefore\ \omega^2=\frac{1}{LC}+\frac{R^2}{L^2},\quad \omega=\omega_0'=\sqrt{\left(\frac{1}{LC}+\frac{R^2}{L^2}\right)}.$$

We see that at this frequency, though impedance of the parallel circuit is maximum, but the power factor is not unity. Consequently if L is adjusted until the circuit has maximum impedance, the power factor will no longer be unity.

(*b*) **Varying capacitance :** Similar to above case, the condition for maximum impedance can be found as

$$\frac{d\,|\,Y\,|^2}{dC}=\frac{(R^2+\omega^2L^2)\,[2\omega^2CR^2+2\,(\omega^2LC-1)\,\omega^2L]}{(R^2+\omega^2L^2)^2}=0$$

or
$$CR^2+L\,(\omega^2LC-1)=0$$

$$\omega^2L^2C=L-CR^2$$

$$\omega=\omega_0=\sqrt{\left(\frac{1}{LC}-\frac{R^2}{L^2}\right)},$$

which is the condition for *unity power factor resonance*. Thus when capacitor is adjusted for maximum impedance, the condition for unity power factor is itself satisfied. This metod of tuning is, therefore, preferred in radio experiments.

(*c*) **Varying frequency :** By a procedure, adopted in above two cases, it can be shown that when frequency is varied to obtain maximum impedance, the frequency of parallel resonance is

$$\omega''_0=\left[\frac{1}{LC}\sqrt{\left(1+\frac{2R^2C}{L}\right)}-\frac{R^2}{L^2}\right]^{1/2}$$

and

$$f_0''=\frac{1}{2\pi\sqrt{(LC)}}\left[\sqrt{\left(1+\frac{2R^2C}{L}\right)}-\frac{R^2C}{L}\right]^{1/2}$$

The condition is not consistent with the condition of unity power factor. Resonance curves, when frequency is made variable, are shown in figure 11.

It is to be noted here that impedance varsus frequency curve for parallel circuit resembles with current versus frequency curve of series circuit. We shall discuss this in the next article.

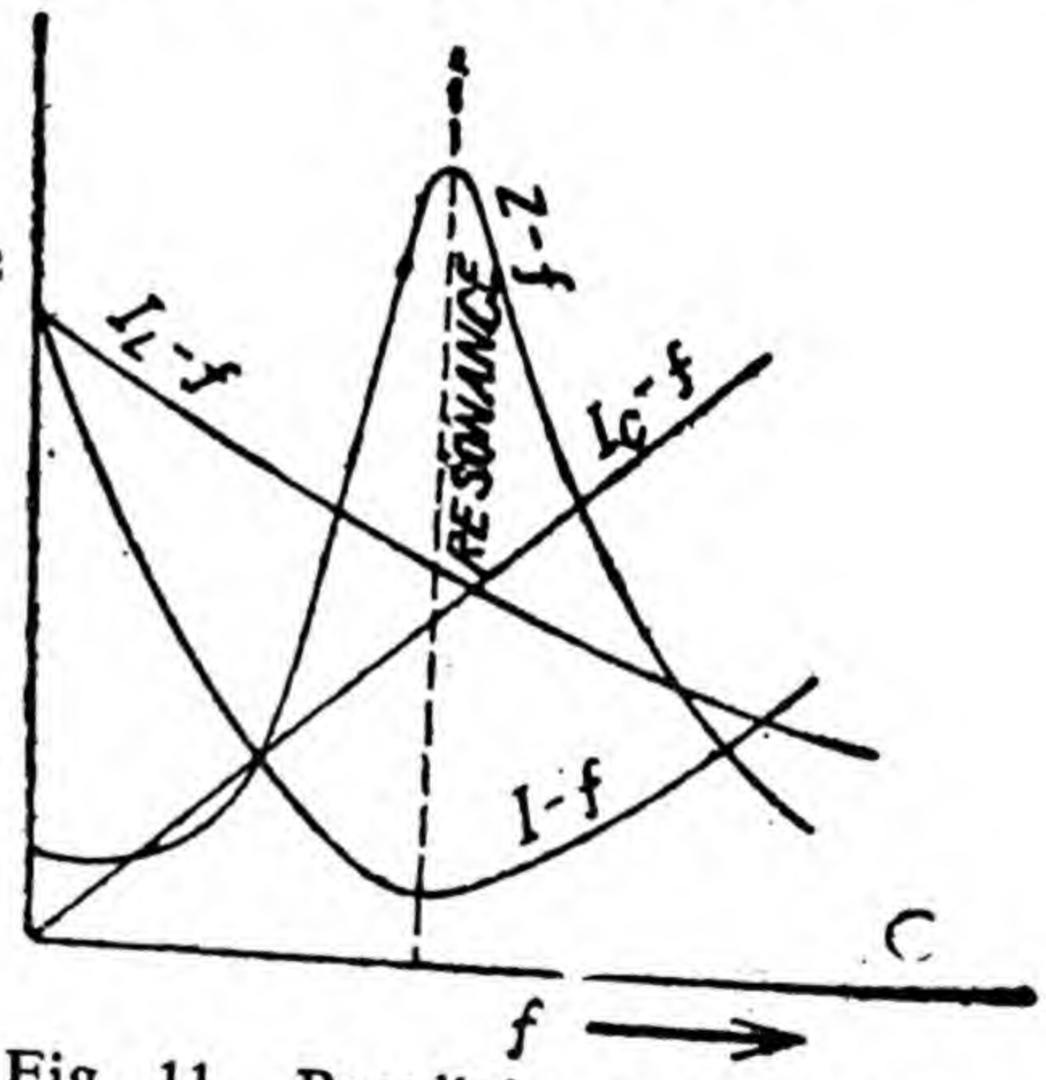


Fig. 11. Parallel resonance by varying frequency.

2·6. DISCUSSION OF Q-VALUES :

In above three cases, we have described the resonance condition for the circuit when Q value is low. Radio circuits mostly employ the parallel circuits which have sufficiently high Q values. When the value of Q is high then $L \triangleright > R$ and we can neglect terms involving R/L. When we apply this condition to above three cases then the three possible conditions for parallel resonance are satisfied essentially at the same frequency, *i.e.*

$$\omega_0=\omega_0'=\omega_0''=\frac{1}{\sqrt{(LC)}},$$

which is same as for series circuit.

2·7. SELECTIVITY OF PARALLEL RESONANT CIRCUIT :

If for the high Q value we write the expression for impedance of parallel circuit, it will come out to be

$$Z=\frac{Z_L Z_C}{Z_R+Z_L+Z_C}=-\frac{j\omega L.\ j/\omega C}{R+j\left(\omega L-\frac{1}{\omega C}\right)}=\frac{L/C}{R+j\left(\omega L-\frac{1}{\omega C}\right)}.$$

Multiplying by ω_0 to L and C in numerator,

$$Z=\frac{\omega_0 L/\omega_0 C}{R+j\left(\omega L-\frac{1}{\omega C}\right)}$$

$$Z_{parallel}=\frac{(\omega_0 L)^2}{R+j\left(\omega L-\frac{1}{\omega C}\right)}, \text{ since } \omega_0 L=\frac{1}{\omega_0 C}$$

$$Z_{parallel}=\frac{\text{constant}}{R+j\left(\omega L-\frac{1}{\omega C}\right)}, \qquad \ldots(1)$$

If we write equation (2) § 2·2 for the current in series circuit,

then $$I_{series}=\frac{V}{R+j\left(\omega L-\frac{1}{\omega C}\right)}=\frac{\text{constant}}{R+j\left(\omega L-\frac{1}{\omega L}\right)}. \qquad \ldots(2)$$

It follows from equations (1) and (2) that both parallel impedance and series current are equal to a constant divided by the same impedance. Therefore, *the resonance curve of the parallel impedance of a circuit can be considered to have the same shape as the resonance curve of the series current in a circuit consisting the same inductance, capacitance and resistance connected in series instead of a parallel arrangement.* Consequently the working rules that were applied for estimating the sharpness of resonance of the series circuit also apply to the case of parallel resonance when the circuit Q is high or moderate.

The parallel circuit's response curve is shown in fig. 12. The peak of the curve will correspond to impedance at resonance *i. e.* L/CR. The sharpness of the curve depends upon the circuit Q and may be increased by increasing the value of L/C or decreasing the value of R.

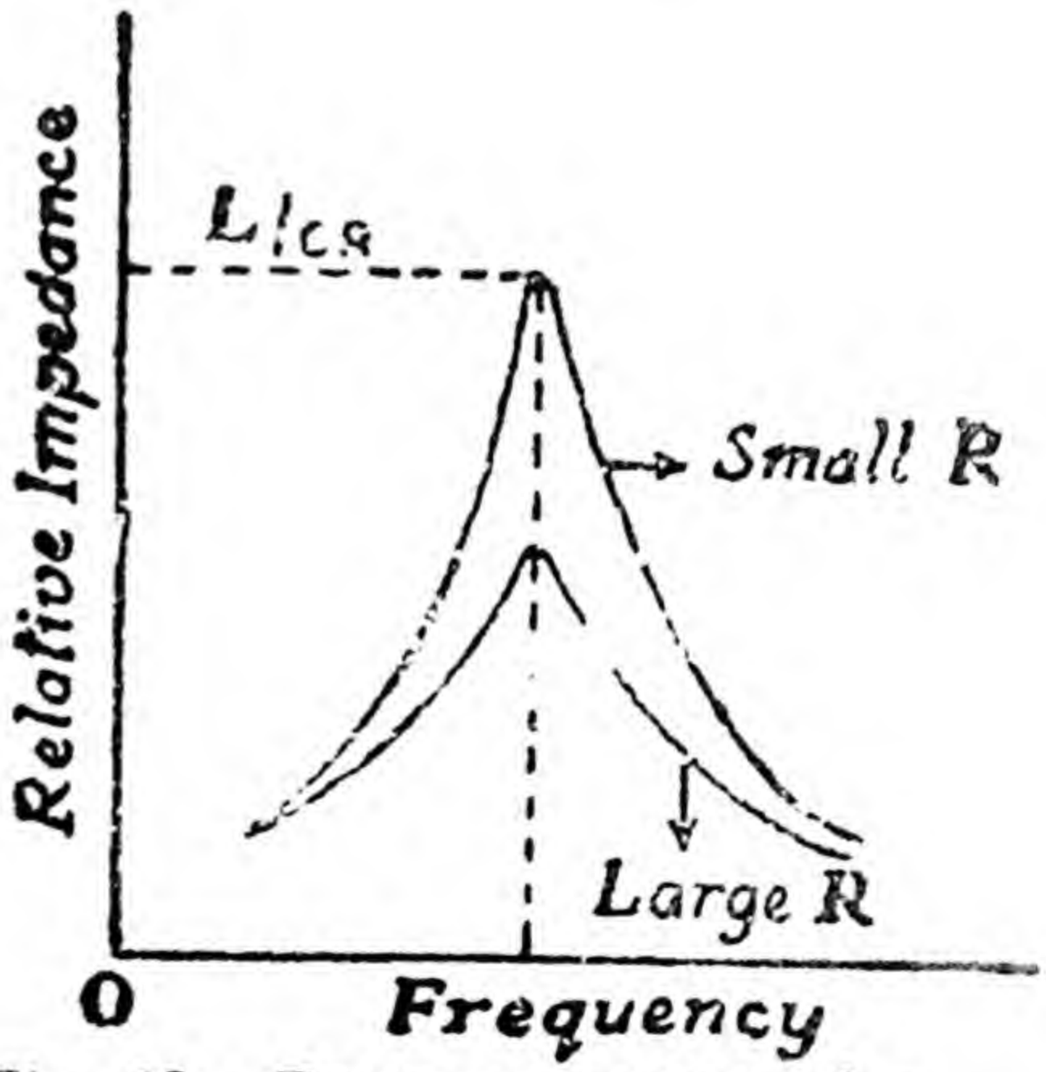


Fig. 12. Resonance curves for *RLC* parallel circuit.

For high R values the pass bandwidth or the degree of selectivity of the parallel resonant circuit can be defined in a similar manner as for series resonant circuit, *i.e.*

$$\Delta f = f_0/Q.$$

2·8. FURTHER DISCUSSION OF SHARPNESS OF RESONANCE IN PARALLEL CIRCUIT :

The variation of parallel circuit impedance with angular frequency ω is shown in figure 13. The circuit is said to be more selective if it offers high impedance at resonance and a much lower impedance at other frequencies.

PP' at which the impedance is $Z_0/\sqrt{2}$, where Z_0 is the impedance at resonant frequency for medium Q is called width of resonance curve. Then Q, the selectivity, is defined as $\frac{\omega_0}{\Delta\omega}$. In practice $R \to 0$ and thus

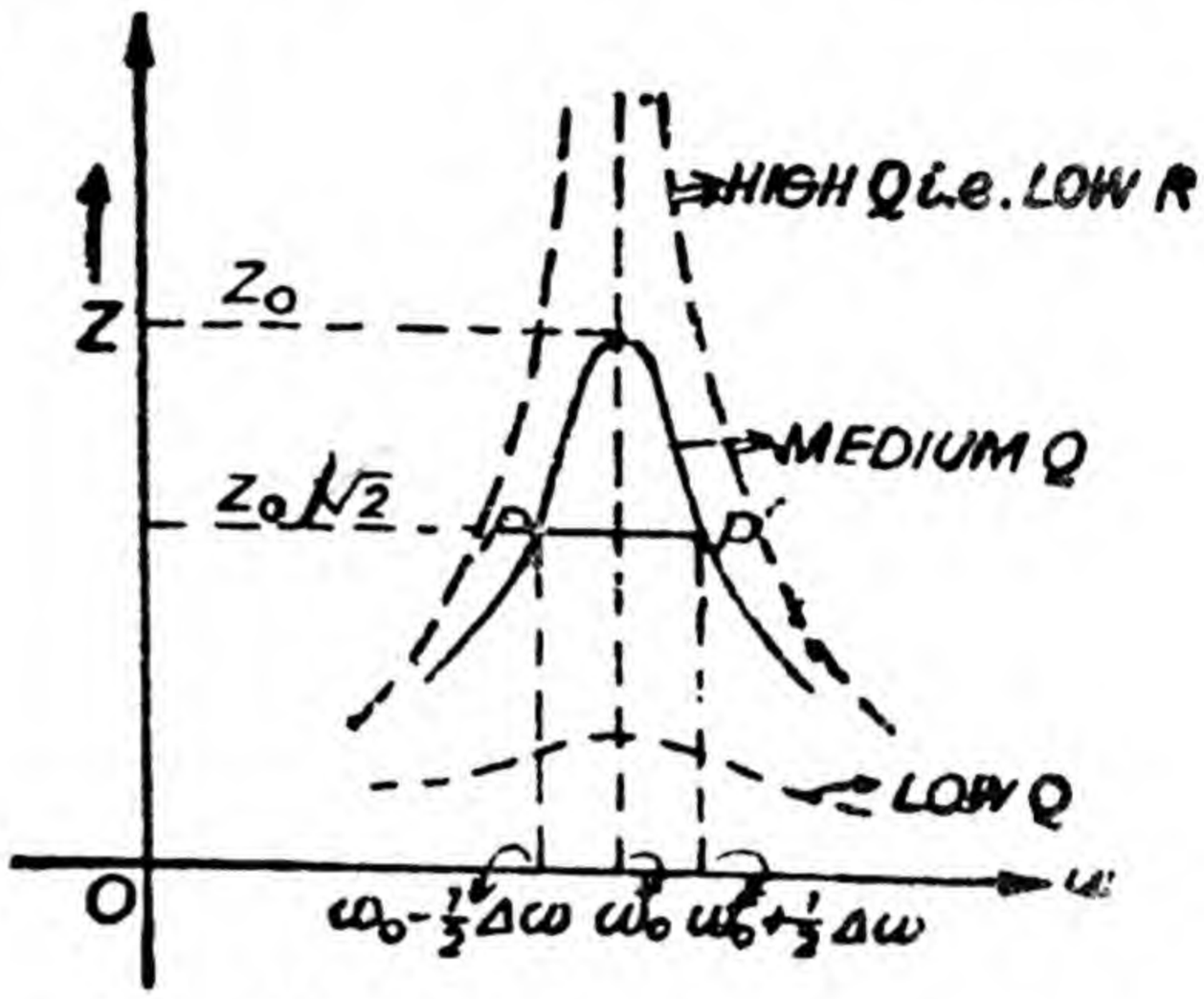


Fig. 13. Impedance curve

$$Z_{R\to 0} = \frac{j\omega L}{(1-\omega^2 LC)+j\omega CR},$$

$$|Z|_{R\to 0} = \frac{\omega L}{\sqrt{\{(1-\omega^2 LC)^2+\omega^2 C^2 R^2\}}},$$

At $\omega=\omega_0$, $Z=Z_0=\frac{L}{RC}$.

If $\omega=\omega_0\pm\frac{1}{2}\Delta\omega$, the terms ωL and $\omega^2C^2R^2$ will be almost same as $\omega_0 L$ and $\omega_0^2C^2R^2$ but the terms $(1-\omega^2LC)$ will no longer be zero $\left(\text{since when } R\to 0,\ \omega_0=\frac{1}{\sqrt{(LC)}}\right)$ so that for this term we assume that $\omega=\omega_0\pm\frac{1}{2}\Delta\omega$, and not only equal to ω.

For $R\to 0$ and $|Z|=\frac{|Z_0|}{\sqrt{2}}$, *i.e.*, $\omega=\omega_0+\frac{1}{2}\Delta\omega$, we have

$$Z=\frac{|Z_0|}{\sqrt{2}}=\frac{\omega_0 L}{\sqrt{[\{1-(\omega_0+\frac{1}{2}\Delta\omega)^2 LC\}^2+\omega_0^2C^2R^2]}},$$

$$\frac{|Z_0|}{\sqrt{2}}=\frac{L/CR}{\frac{1}{\omega_0 CR}\sqrt{[\{1-(\omega_0+\frac{1}{2}\Delta\omega)^2 LC\}^2+\omega_0^2C^2R^2]}}$$

$$=\frac{Z_0}{\frac{1}{\omega_0 CR}\sqrt{[\{1-(\omega_0+\frac{1}{2}\Delta\omega,^2 LC\}^2+\omega_0^2C^2R^2]}}$$

Therefore, $\sqrt{2}\omega_0 CR=\sqrt{[\{1-(\omega_0+\frac{1}{2}\Delta\omega)^2\ LC\}^2+\omega_0^2C^2R^2]}$,

$$\omega_0^2C^2R^2=[1-(\omega_0+\tfrac{1}{2}\Delta\omega)^2\ LC]^2$$

$$=\left[1-(\omega_0^2+\omega_0\Delta\omega+\tfrac{1}{4}\Delta\omega^2)\ \frac{1}{\omega_0^2}\right]^2$$

$$=\left(\frac{\Delta\omega}{\omega_0}\right)^2 \text{ neglecting } \Delta\omega^2 \text{ terms.}$$

$$\therefore \quad \pm\omega_0 CR=\left(\frac{\Delta\omega}{\omega_0}\right) \quad \text{or} \quad \frac{1}{\omega_0 CR}=\left(\frac{\omega_0}{\Delta\omega}\right).$$

But
$$Q=\frac{1}{\omega_0 CR}=\frac{L\omega_0}{R}$$

$$=\frac{L}{R}\cdot\frac{1}{\sqrt{(LC)}}=\frac{1}{R}\sqrt{\left(\frac{L}{C}\right)}.$$

Thus
$$Q=\left(\frac{\omega_0}{\Delta\omega}\right)=\frac{1}{\omega_0 CR}=\frac{L\omega_0}{R}=\frac{1}{R}\sqrt{\left(\frac{L}{C}\right)},$$

which is a measure of the selectivity of the parallel resonant circuit.

2·9. EFFECT OF SOURCE IMPEDANCE ON THE SHARPNESS OF RESONANCE CURVES :

Suppose Z_g is the internal impedance of the generator to which the parallel combination of L and C is connected. Then, by Thevenin's theorem, assuming the generator voltage as open-circuited voltage, the current through load connected will be

$$I_R=\frac{e_{0c}}{Z_g+Z_R}.$$

Therefore, voltage across the combination

$$E_R=\frac{e_{0c}Z_R}{Z_g+Z_R},$$

where Z_R is the impedance of parallel combination.

Now we see that :

(*i*) at the zero internal impedance of the generator ($Z_g=0$), the voltage across the combination will be independent of frequency as represented by a line $Z_g=0$ in the figure 14 :

(*ii*) if the generator has high impedance then at resonant frequency, the parallel circuit impedance being maximum, the current will be minimum. With the result, that the voltage drop across combination will be higher and that across generator impedance will be minimum :

(*iii*) as the frequency is increased or decreased from the resonant value, the parallel circuit impedance decreases and the resistance of the generator is increased. Consequently, voltage across combination falls and that across generator

impedance increases. Greater is the value of generator impedance, *i.e.*, greater the deviation of frequency from resonant frequency, sharper will be the fall *i.e.*, curves are more peaked.

Therefore to secure better sharpness of resonance curve with high impedance generator, a parallel combination should be used.

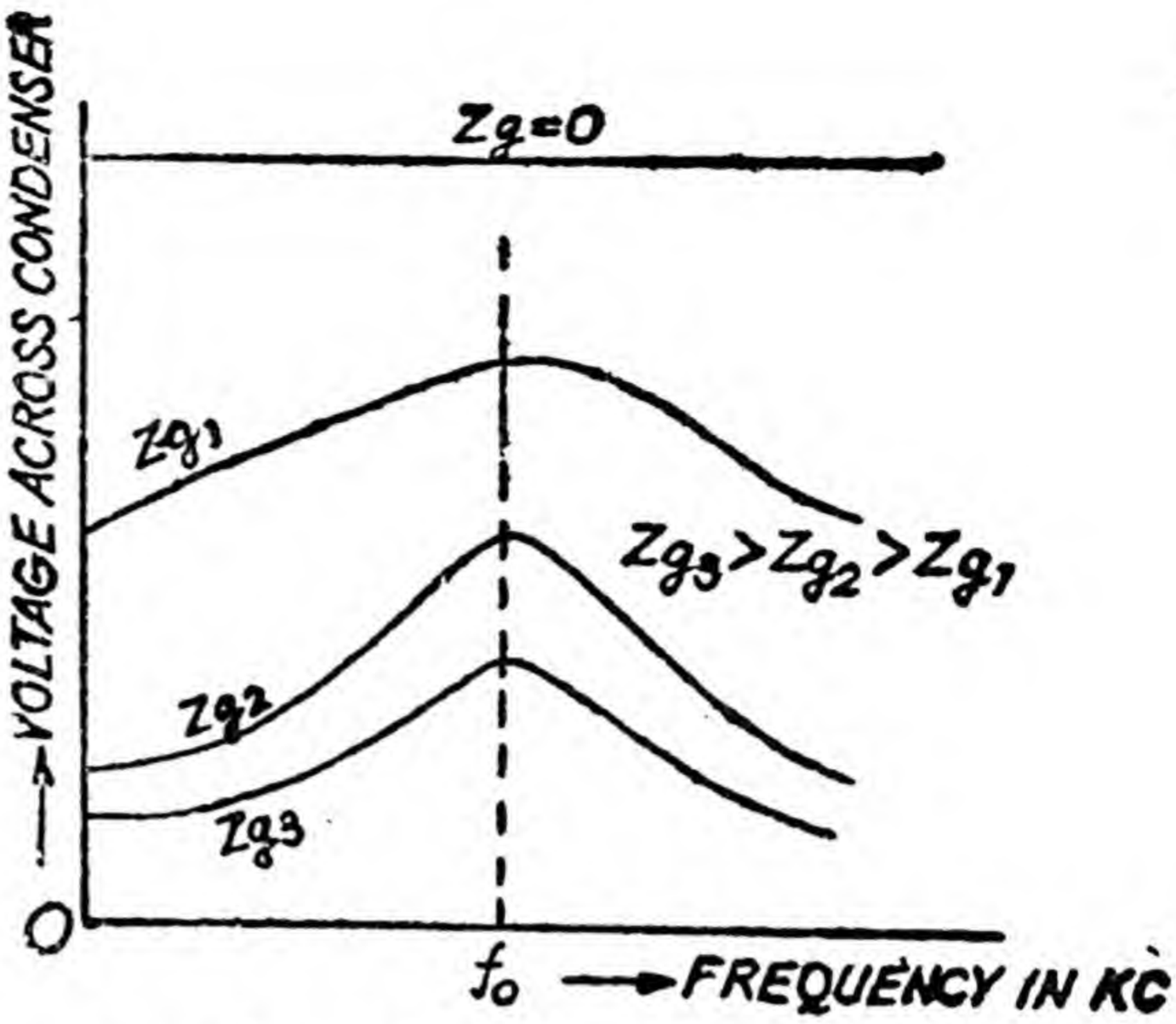


Fig. 14. Effect of generator impedance of parallel resonance curve.

2·10. DESIGN OF AN ANTI-RESONANT CIRCUIT :

In vacuum tube class *C* amplifiers, the anti-resonant circuits are employed as a tuned load to the plate circuit. The reason is that the branch resistance is transformed into the large value appropriate for loading of vacuum tube.

At anti-resonant frequency, the impedance of the circuit is purely resistive and is given by

$$R_{ar}=\frac{R}{\omega^2C^2R^2+(\omega^2LC-1)^2}.$$

Now the condition for unity factor is

$$L=C\,(R^2+\omega^2L^2) \quad \text{or} \quad C=\frac{L}{R^2+\omega^2L^2}.$$

Substituting this value of C in above expression for R_{ar}, we get

$$R_{ar}=\frac{R\,(R^2+\omega^2L)^2}{\omega^2L^2R^2+[\omega^2L^2-(R^2+\omega^2L^2)]^2}$$

$$=\frac{(R^2+\omega^2L^2)}{R}.$$

$$\therefore \quad \omega L=X_L=\sqrt{[R\,(R_{ar}-R)]}. \qquad \ldots(1)$$

$$\therefore \quad Q_L=\frac{\omega L}{R}=\sqrt{\left[\frac{(R_{ar}-R)}{R}\right]}. \qquad \ldots(2)$$

$$\text{Also } \frac{L}{C}=(R^2+\omega^2L^2) \quad \text{or} \quad \frac{\omega L}{\omega C}=(R^2+\omega^2L^2)=RR_{ar}.$$

$$\therefore \quad \frac{1}{\omega C}=X_C=\frac{RR_{ar}}{\omega L}=R_{ar}\sqrt{\left[\frac{R}{(R_{ar}-R)}\right]}=\frac{R_{ar}}{Q_L} \qquad \ldots(3)$$

Relations (1), (2) and (3) can be applied to design an anti-resonant circuit and hence the design formulae.

Example. *It is desired to design an anti-resonant circuit to transform a branch resistance of* 100 *ohms into an effective resistance of* 10,000 *ohms at the anti-resonant frequency* 10^6 *c/s.*

For the design of an anti-resonant circuit, we should find out the value of Q, L and C. From given data, it is obvious that

$$\omega L = 100 \text{ ohms}$$
$$f = 10^6 \text{ c/s.}$$

Effective resistance $= Q$. branch resistance
$= Q.L\omega.$

Therefore $Q.L\omega = 10{,}000$ ohms.

or $Q = 100$ since $L\omega = 100$ ohms.

Resistance of inductive branch

$$R = \frac{L\omega}{Q} = \frac{100}{100} = 1 \text{ ohm and } L\omega = 100$$

$$L = \frac{100}{2\pi f} = \frac{100}{2.3 \cdot 14.10^6} \text{ henry}$$

$$= \frac{10^{-4}}{\cdot 28} = 0 \cdot 159.10^{-4} \text{ henry.}$$

For the value of capacitance,

$$L\omega = \frac{1}{C\omega} = 100.$$

$$C = \frac{1}{100\omega} = \frac{1}{2\pi.10^6.10^2} = \frac{10^{-8}}{6 \cdot 28} = 0 \cdot 1159.10^{-8} \text{ farad.}$$

The required values for the design of anti-resonant circuit are

Resistance = 1 ohms,
Inductance = $\cdot 159.10^{-4}$ henry,
Capacitance = $\cdot 00159\ \mu F$,
$Q = 100.$

2·11. COMPARISON BETWEEN SERIES AND PARALLEL RESONANCE :

	Series Resonance	*Parallel Resonance*
(*a*)	Impedance at minimum value.	Admittance at minimum value.
(*b*)	Inductive and capacitive reactance equal in magnitude.	Inductive and capacitive susceptances equal in magnitude.
(*c*)	Reactive components of voltage combine to equal zero.	Reactive components of current combine to equal zero.
(*d*)	Voltage source constant in maximum magnitude.	Current source constant in maximum magnitude.
(*e*)	Current maximum for constant resistance.	Voltage maximum for constant conductance.

2·12. COUPLED CIRCUITS :

When the two circuits are so arranged that the energy can be transferred from one to another, they are said to be coupled.

In common cases, coupling is produced by two coils (one in each circuit called primary and secondary coils) between which mutual impedance exists. These are called inductively coupled circuits. The effect of mutual inductance is to make possible the transfer of energy from one circuit to the other by transformer action. That is, the alternating current flowing in one circuit (primary circuit) as a result of the voltage applied to it, produces a magnetic flux which induces a voltage in the coupled circuit (secondary circuit). This results in an induced current and a transfer of energy from primary to secondory circuit. There are three types of primary and secondary circuits :

(*a*) primary and secondary circuits both untuned,
(*b*) primary circuit tuned and secondary circuit untuned,
(*a*) primary and secondary both tuned.

2·12-1. TERMS USED IN COUPLED CIRCUITS :

(1) Conductively Coupled Circuits :

Figure 15 shows the two conductively coupled circuits. The circuit 1 is called as primary circuit while circuit 2 as secondary circuit. The impedance (R_{12} or Z_{21}) of the common branch is called as mutual impedance between the circuits. This impedance may be a pure resistance, a pure inductance, a pure capacitance, or some combination of these circuit elements. Let I_1 and I_2 be the currents flowing in circuit 1 and 2 respectively. The directions of I_1 and I_2 are, of course, arbitrary. Now

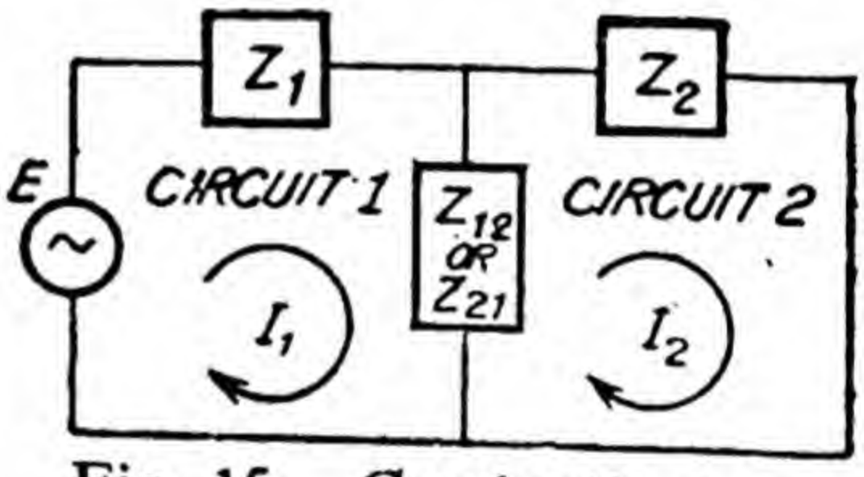


Fig. 15. Conductively coupled circuit.

$$Z_{11}=Z_1+Z_{12} \quad \text{(Impedance of circuit 1 to } I_1) \qquad \ldots(1)$$
$$Z_{22}=Z_2+Z_{21} \quad \text{(Impedance of circuit 2 to } I_2) \qquad \ldots(2)$$

If the circuit parameters are constant,
$Z_{12}=Z_{21}$ (Mutual impedance between circuits 1 and 2).
Applying Kirchhoff's law to circuits 1 and 2, we have

$$Z_{11}I_1-Z_{12}I_2=E \qquad \ldots(3)$$
$$-Z_{21}I_1+Z_{22}I_2=0. \qquad \ldots(4)$$

Solving equation (3) and (4) for I_1 and I_2, we get

$$I=\frac{\begin{vmatrix} E & -Z_{12} \\ 0 & Z_{22} \end{vmatrix}}{\begin{vmatrix} Z_{11} & -Z_{12} \\ -Z_{21} & Z_{22} \end{vmatrix}}=\frac{EZ_{22}}{Z_{11}Z_{22}-R_{12}^2} \qquad \ldots(5)$$

$$I=\frac{\begin{vmatrix} Z_{11} & E \\ -Z_{21} & 0 \end{vmatrix}}{\begin{vmatrix} Z_{11} & -Z_{12} \\ -Z_{21} & Z_{22} \end{vmatrix}}=\frac{EZ_{21}}{Z_{11}Z_{22}-Z_{12}^2}, \qquad \ldots(6)$$

The input impepance of mesh 1 may be determined as

$$Z_{1\,in}=\frac{Z_{11}Z_{22}-Z_{12}^2}{Z_{22}}$$

$$=Z_{11}-\frac{Z_{12}^2}{Z_{22}}. \qquad \ldots(7)$$

This equation shows that the input impedance of the first mesh is modified by the presence of the second mesh because of the term $-Z_{12}^2/Z_{22}$.

The transfer impedance may also be determined as

$$ZT_{12}=\frac{Z_{11}Z_{22}-Z_{12}^2}{-Z_{12}}$$

$$=Z_{12}-\frac{Z_{11}Z_{22}}{Z_{12}} \qquad \ldots(8)$$

(2) Mutual impedance : *The mutual impedance between circuits* 1 *and* 2 *is defined as the ratio of the voltage developed in circuit–*2 *per unit current in circuit-*1 *when the circuit-*2 *is open circuited.*

The definition of mutual impedance may be generalized to apply to two pairs of terminals 11′ and 22′ as shown in figure 16.

$$Z_{21}=\frac{V_1}{I_1}. \qquad \ldots(9)$$

A particular form of capacitive coupling is shown in fig. 17.

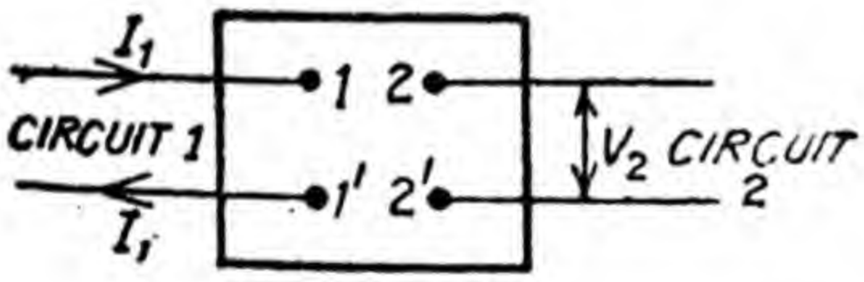


Fig. 16. Circuit-1 coupled to circuit-2 through an arbitrary network.

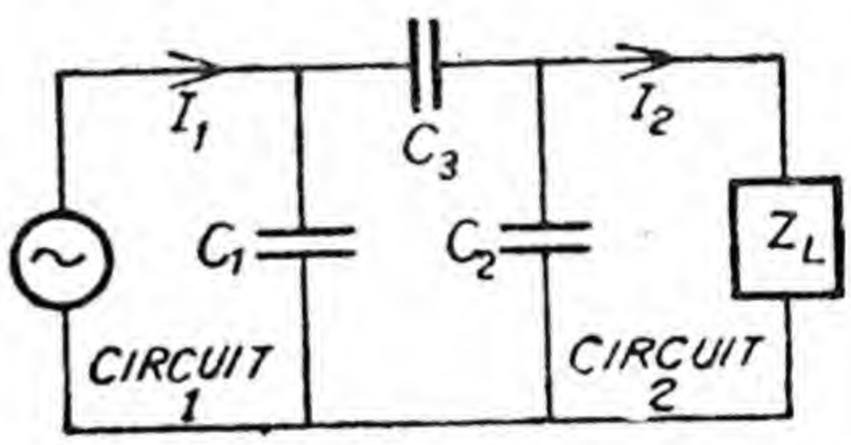


Fig. 17. Circuits coupled through C_1– C_3—C_2 network.

The coupling reactance is

$$X_{coupling}=\frac{\dfrac{V_1X_2}{X_2+X_3}}{\dfrac{V_1(X_1+X_2+X_3)}{X_1(X_2+X_3)}}=\frac{X_1X_2}{X_1+X_2+X_3}, \qquad \ldots(10)$$

where V_1 is the voltage across C_1 and X's are the capacitive reactances of the respective condensers The coupling capacitance between circuit-1 and circuit-2 is

$$C_{coupling}=\frac{1}{\omega X_{coupling}}=\frac{\dfrac{1}{\omega\dfrac{(1/\omega C_1)(1/\omega C_2)}{\left(\frac{1}{\omega C_1}\right)+\left(\frac{1}{\omega C_2}\right)+\left(\frac{1}{\omega C_3}\right)}}}{}$$

$$=C_1+C_2+\frac{C_1C_2}{C_3}. \qquad \ldots(11)$$

(3) Coefficient of Coupling : The coefficient of coupling between circuits 1 and 2 is defined as

$$k=\frac{Z_{12}}{\sqrt{(Z_{11}'Z_{22}')}}=\frac{Z_{21}}{\sqrt{(Z_{11}'Z_{22}')}}, \qquad \ldots(12)$$

where Z_{11}' is the impedance seen looking into terminals $11'$ with terminals $22'$ open circuited.

Z_{22}' is the impedance seen looking into terminals $22'$ with terminals $11'$ open circuited.

(4) Magnetic coupling : The two circuits are called magnetically coupled when a portion of the magnetic flux established by one circuit interlinks with the second circuit or the energy is transferred from one circuit to the other circuit by way of the magnetic field which is common to the two circuits. Figure 18 shows the two coils magnetically coupled. The total flux ϕ_1 due to current i_1 in circuit 1 is divided into two parts.

(*i*) The part ϕ_{11}, which links only with the turns of circuit-1 and is known as leakage flux of circuit-1 with respect to circuit-2.

(*ii*) The part ϕ_{12}, which links with the turns of circuit-2 and is known as mutual flux produced by circuit-1.

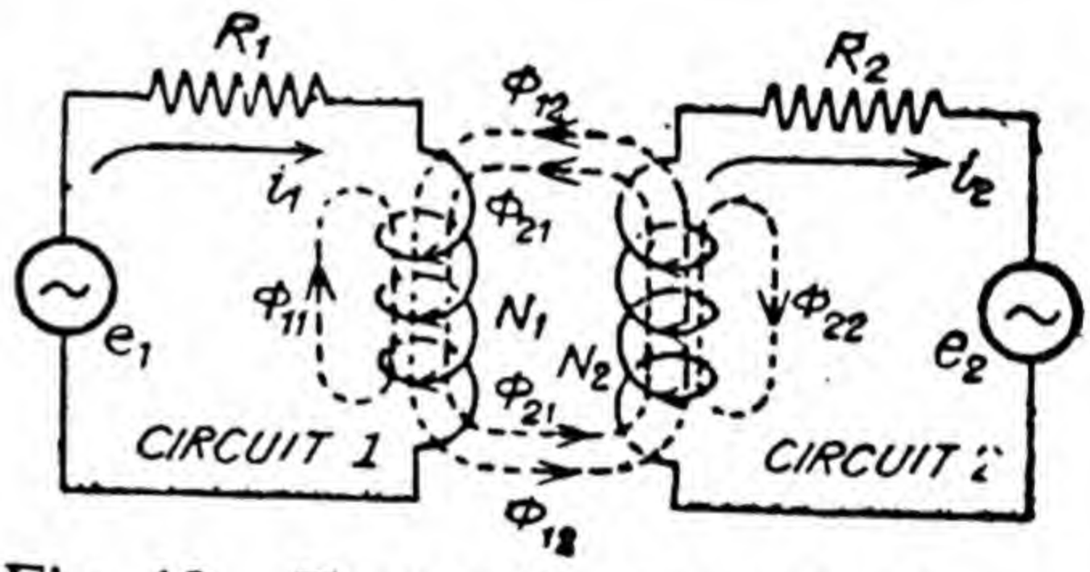


Fig. 18. Two coils magnetically coupled.

$$\therefore \quad \phi_1=\phi_{11}+\phi_{12} \qquad \ldots(13)$$

Similarly, the total flux ϕ_2 due to current i_2 in circuit-2 is divided into two parts :

(*i*) The part ϕ_{22}, which links only with the turns of circuit-2 and is known as leakage flux of circuit-2 with respect to circuit-1

(*ii*) The part ϕ_{21}, which links with the turns of circuit-1 and is known as mutual flux produced by circuit-2.

$$\therefore \quad \phi_2=\phi_{22}+\phi_{21}. \qquad \ldots(14)$$

(5) Mutual inductance : If a current through an inductance varies with time, the flux in it also varies with time. Due to the change in flux an e.m.f. is induced in the coil. It is observed experimentally that the e.m.f. is proportional to the rate of change of the coil current, *i.e.*

$$e=-L\frac{di}{dt} \text{ volts}, \qquad \ldots(15)$$

where L is the constant of proportionality. This constant is known as *self inductance* of the coil.

From Faraday's law, we have

$$e = -N \frac{d\phi}{dt}. \qquad \ldots(16)$$

From equations (15) and (16), we get

$$N \frac{d\phi}{dt} = L \frac{di}{dt}$$

or
$$L = N \frac{d\phi}{dt} \text{ henrys.} \qquad \ldots(17)$$

As shown in figure 18, when a current i_1 varies with time, the flux ϕ_{11} and ϕ_{12} also vary with time. The flux ϕ_{11} produces a voltage in circuit-1 and ϕ_{12} produces a voltage in circuit 2. The e.m.f. e_{12} induced in N_2 is called an e.m.f. of *mutual* induction, since it is an e.m.f. produced in circuit 2 by a change of current in circuit-1. The mutual e.m.f. is proportional to the rate of change of current *i*, *i.e.*,

$$e_{12} = -M_{12} \frac{di_1}{dt} \text{ volts.} \qquad \ldots(18)$$

The constant of proportionality M_{12} is called the *normal inductance* between coils N_1 and N_2.

If the rate of charge of flux ϕ_{12} is considered, then

$$e_{12} = -N_2 \frac{d\phi_{12}}{dt} \text{ volts.} \qquad \ldots(19)$$

Comparing equations (18) and (19), we get

$$M_{12} = N_2 \frac{d\phi_{12}}{di_1} \text{ henrys.} \qquad \ldots(20)$$

Equation (20) shows that the mutual inductance is proportional to the rate of change of flux in N_2 with respect to change in current.

Similarly, the mutual inductance of circuit 1 with respect to circuit 2 is

$$M_{21} = N_1 \frac{d\phi_{21}}{di_2}. \qquad \ldots(21)$$

(6) Mutual Reactance : Considering figure 18, the basic equations of voltage for the two circuits can be written as

$$e_1 = R_1 i_1 + L_1 \frac{di_1}{dt} + M_{21} \frac{di_2}{d}, \qquad \ldots(22)$$

$$e_2 = R_2 i_2 + L_2 \frac{di_2}{dt} + M_{12} \frac{di_1}{dt}. \qquad \ldots(23)$$

If, for example, $i_1 = I_{m1} \sin \omega t$, the voltage drop in current 2 due to mutual inductance is

$$M_{12} \frac{di_1}{dt} = \omega M_{12} I_{m1} \cos \omega t$$

$$= X_{M12} I_{m1} \cos \omega t$$

In general, $\omega M = X_M$. It is called the mutual reactance. This is the impedance function which expresses the ratio of the voltage

of mutual inductance to the exciting current. The voltage of mutual inductance leads the exciting current by 90° and hence the vector expression for the mutual reactance is

$$\overrightarrow{X_M} = j\omega M = \omega M \angle 90°. \quad \text{.. (24)}$$

(7) Coefficient of Magnetic Coupling :

From equation (20) and equation (17), we have

$$\frac{M_{12}}{L_1} = \frac{N_2\, d\phi_{12}/di_1}{N_1\, d\phi_1/di_1} = \frac{N_2\, d\phi_{12}}{N_1\, d\phi_1}. \quad \text{...(25)}$$

The fractional part of ϕ_1 which links with N_2, *i.e.*, (ϕ_{12}/ϕ_1), and the fractional part of ϕ_2 which links with N_2, *i.e.* (ϕ_{21}/ϕ_2) give the degree of coupling that exists between the two windings. If these two fractions are small, the coupling is said to be loose. If all the flux ϕ_1 due to i_1 on coil N_1 links with the turns N_2, *i.e.*, there is no leakage flux, we have

$$\phi_1 = \phi_{12} \text{ and } d\phi_1 = d\phi_{12}$$

$$\therefore \quad \frac{M_{12}}{L_1} = \frac{N_2}{N_1}. \quad \text{...(26)}$$

If a fraction k_1 of the flux ϕ_1 links with the turns N_2, then $\phi_{12} = k_1\phi_1$ and

$$\frac{M_{12}}{L_1} = k_1 \frac{N_2}{N_1} \quad \text{...(27)}$$

Similarly, if the fraction k_2 of the flux ϕ_2 links with the turns of N_1, we have

$$\frac{M_{21}}{L_2} = k_2 \frac{N_1}{N_2}. \quad \text{...(28)}$$

Multiplying equations (27) and (28), we get

$$k_1 k_2 = \frac{M_{12}}{L_1} \times \frac{M_{21}}{L_2}.$$

If the circuit is linear then $M_{12} = M_{21} = M$. We define the coefficients as $k_M = (k_1\, k_2)^{1/2}$,

$$k_M = \frac{M}{\sqrt{(L_1 L_2)}}. \quad \text{...(29)}$$

The coefficient of coupling is expressed as decimal or a percentage, having a maximum value unity or 100 percent.

2·12–2. THE AIR-CORE TRANSFORMER :

The conventional air-core transformer arrangement is shown in fig. 19. The circuit 1 energized by a.c. is called as primary circuit and circuit 2 is called as secondary circuit. Due to the magnetic coupling between the two circuits, a voltage is induced in the secondary circuit. The

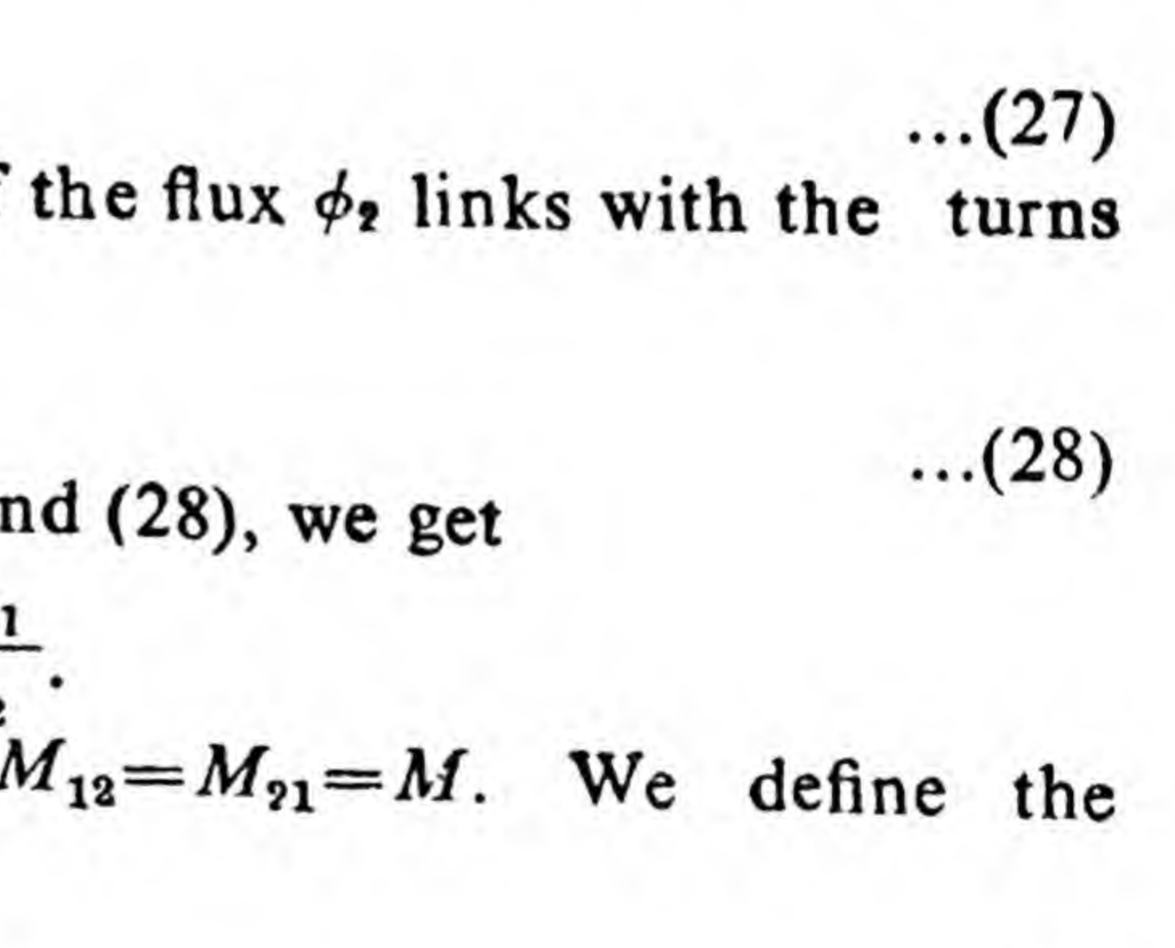


Fig. 19. Conventional air-core transformer arrangement.

induced voltage is proportional to the number of secondary turns N_2 and depends upon the degree of coupling between the two windings. For the positive clockwise direction of i_2, the sign of M is taken as positive. Under the condition $M_{21}\ di_2/dt$ acts in the same direction as $L_1\ di_1/dt$ in the primary winding. For the counter clockwise direction of i_2 the sign of M is taken as negative.

The circuit equations are

$$R_1 i_1 + L_1 \frac{di_1}{dt} + M_{21} \frac{di_2}{dt} = v_1 \qquad \ldots(1)$$

$$(R_2+R)\ i_2 + (L_2+L) \frac{di_2}{dt} + \frac{\int i_2 dt}{C} + M_{12} \frac{di_1}{dt} = 0. \qquad \ldots(2)$$

In terms of effective values, the equations (1) and (2) can be written as

$$(R_1 + j\omega L_1)\ I_1 + j\omega M I_2 = V_1 \qquad \ldots(3)$$

$$(R_2 + j\omega L_2)\ I_2 + \left[R + j\left(\omega L - \frac{1}{\omega C}\right)\right] I_2 + j\omega M I_1 = 0. \qquad \ldots(4)$$

For simplicity, we use the following abbreviations

$Z_1 = (R_1 + j\omega L_1)$, $Z_2 = (R_2 + j\omega L_2)$, $Z_M = j\omega M$

$$Z = \left[R + j\left(\omega L - \frac{1}{\omega C}\right)\right]$$

$$\therefore \quad Z_1 I_1 + Z_M I_2 = V_1 \qquad \ldots(5)$$

$$Z_M I_1 + (Z_2 + Z)\ I_2 = 0 \qquad \ldots(6)$$

Solving these equations, we have

$$I_1 = \frac{V_1 (Z_2 + Z)}{Z_1 (Z_2 + Z) - Z_M^2} \qquad \ldots(7)$$

and

$$I_2 = \frac{-V_1 Z_M}{Z_1 (Z_2 + Z) - Z_M} \qquad \ldots(8)$$

The voltage V_2 which appears across the load impedance is

$$V_2 = \frac{-V_1 Z_M Z}{Z_1 (Z_2 + Z) - Z_M^2}. \qquad \ldots(9)$$

The *equivalent impedance* Z_{e1} referred to the primary side is defined as the ratio of applied voltage to the primary current, *i.e.*

$$Z_{e1} = \frac{V_1}{I_1} \quad \frac{Z_1 (Z_2 + Z) - Z_M z}{(Z_2 + Z)}. \qquad \ldots(10)$$

Transferred impedance : Now we will show that how an impedance placed across the secondary terminals of an air core transformer appears at the primary terminals in the modified form.

The equivalent impedance of an air-core transformer referred to primary side is given by

$$Z_{e1} = \frac{Z_1 (Z_2 + Z) - Z_M^2}{(Z_2 + Z)}$$

$$=Z_1-\frac{Z_M^2}{(Z_2+Z)}$$

$$=(R_1+j\omega L_1)+\frac{\omega^2M^2}{(R_2'+j\omega L_2')} \quad \ldots(11)$$

where $Z_2+Z=(R_2+j\omega L_2)+\left[R+j\left(\omega L-\frac{1}{\omega C}\right)\right]$

$$=\left[(R_2+R)+j\left\{\omega\,(L+L_2)-\frac{1}{\omega C}\right\}\right]$$

$$=R_2'+j\omega L_2'.$$

Rationalizing equation (11), we have

$$Z_{e1}=(R_1+j\omega L_1)+\frac{\omega^2M^2\times(R_2'-j\omega L_2')}{(R_2'+j\omega L_2')\,(R_2'-j\omega L_2')}$$

$$=R_1+j\omega L_1+\frac{\omega^3M^2R_2'}{(R_2'^2+\omega^2L_2'^2)}-\frac{j\omega^3M^2L_2'}{(R_2'^2+\omega^2L_2'^2)}\Bigg]$$

$$=\left[\,R_1+\frac{\omega^2M^2R_2'}{(R_2'^2+\omega^2L_2'^2)}\right]+j\omega\left[\,L_1-\frac{\omega^2M^2L_2'}{(R_2'^2+\omega^2L_2'^2)}\right] \quad \ldots(12)$$

It is observed from equation (12) that R_2' appears at the primary terminals in the modified form as

$$\left[\frac{\omega^2M^2}{(R_2'^2+\omega^2L_2'^2)}\right]R_2'=\frac{X_M^2}{Z_2'^2}\,R_2'. \quad \ldots(13)$$

If R_2' is very small compared with ω_2L_2', then equation (13) takes the form $\left(\frac{N_1}{N_2}\right)^2 R_2'$ approximately.

Thus a resistance R placed across the secondary may appear at primary as $\left(\frac{N_1}{N_2}\right)^2 R$ approximately.

Primary unity power factor resonance: The inductive reactance of equivalent impedance Z_e, caused by the introduction of transformer may be neutralized by using capacitors either in primary or secondary circuits. These capacitors may be arranged either in series or, in parallel with the transformer windings. We know that (equation 12),

$$Z_{e1}=\left[\,R_1+\frac{\omega^2M^2R'_2}{R_2'^2+\omega^2L_2'^2)}\right]+j\omega\left[\,L_1-\frac{\omega^2M^2L'_2}{(R_2'^2+\omega^2L_2'^2)}\right]$$

$$=R_{e1}+jX_{e1}, \quad \ldots(14)$$

where $R_{e1}=\left[\,R_1+\frac{\omega^2M^2R_2'}{(R_2'^2+\omega^2L_2'^2)}\right]$ and

$$X_{e1}=\left[\,\omega L_1-\frac{\omega^2M^2L_2'}{(R_2'^2+\omega^2L_2'^2)}\right]=\left[\,X_1-\frac{X_M^2X_2'}{R_2'^2+X_2'^2}\right].$$

Now by introducing a capacitor (which has a capacitive reactance equal in magnitude to the inductive reactance X_{e1}) in series

KASHMIR UNIVERSITY
Iqbal Library
Acc. No

with the primary, a primary *unity power factor* can be obtained. Thus the value of the capacitive reactance for primary unity power factor should be

$$X_{e1\ (series)}=\left[\ X_1-\frac{X_M^2 X_2'}{R_2'^2+X_2'^2}\ \right]. \qquad ...(15)$$

The capacitor may be placed in parallel with the primary. In this case the susceptance of the parallel capacitor b_C should be equal in magnitude to the susceptance b_L of Y_{e1}, where

$$Y_{e1}=\frac{1}{R_{e1}+jX_{e1}}=\frac{R_{e1}}{R_{e1}^2+X_{e2}^2}-j\frac{X_{e1}}{R_{e1}^2+X_{e1}^2}. \qquad ...(16)$$

The capacitive susceptance of the parallel primary capacitor must be equal to

$$b_{e1\ (parallel)}=\frac{X_{e1}}{R_{e1}^2+X_{e1}^2} \qquad ...(17)$$

Example : *Obtain an expression for the effective impedance of the primary of a circuit shown in figure* 20.

Here primary and secondary both are tuned circuits. Let X_1 and X_2 be the reactances of primary and secondary circuits, then

$$X_1=\omega L_1-\frac{1}{\omega C_1}$$

and $X_2=\omega L_2-\frac{1}{\omega C_2}$.

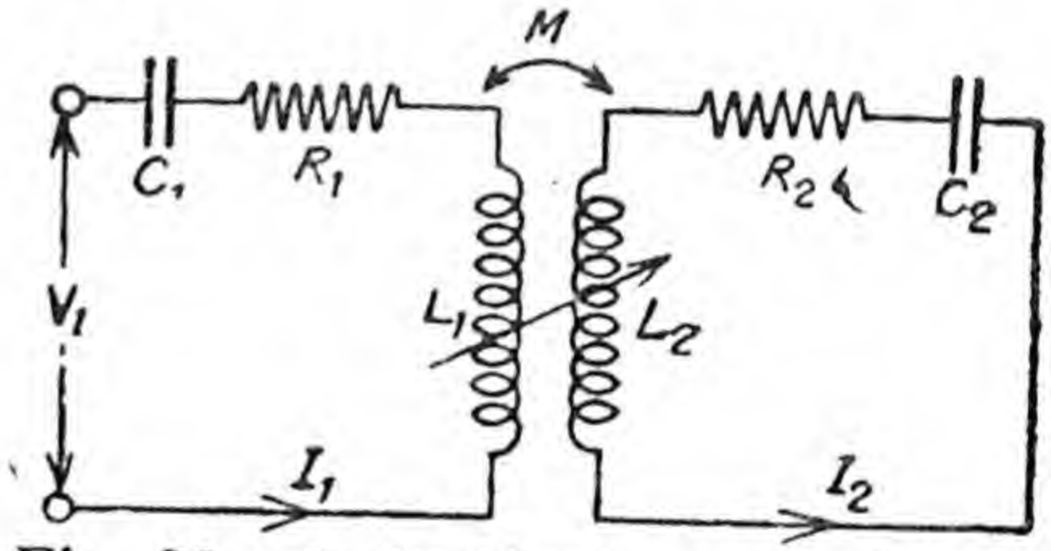


Fig. 20. Inductively coupled circuit.

The current I_1 in the primary circuit will be

$$I_1=\frac{V_1-j\omega MI_2}{(R_1+jX_1)}. \qquad ...(i)$$

Induced currents I_2 in the secondary circuit can be shown to be given by the expression

$$I_2=-\frac{j\omega MI_1}{(R_2+jX_2)}. \qquad ...(ii)$$

Multiplying above and below in equation (*ii*) by (R_2-jX_2)

$$I_2=-\frac{j\omega MI_1}{(R_2+jX_2)}.\frac{(R_2-jX_2)}{(R_2-jX_2)}$$

$$=-\frac{j\omega MI_1\ (R_2-jX_2)}{(R_2^2+X_2^2)}$$

Putting the value of I_2 in equation (*i*), we get

$$I_1\ (R_1+jX_1)=V_1-j\omega M\left\{\frac{-j\omega MI_1\ (R_2-jX_2)}{R_2^2+X_2^2}\right\}$$

$$=V_1-\frac{\omega^2M^2I_1\ (R_2-jX_2)}{R_2^2+X_2^2}$$

$$\therefore \quad V_1 = I_1 \left\{ (R_1 + jX_1) \frac{\omega^2 M^2 (R_2 - jX_2)}{R_2^2 + X_2^2} \right\}$$

$$= I_1 \left\{ \left(R_1 + \frac{\omega^2 M^2 R_2}{R_2^2 + X_2^2} \right) + j \left(X_1 - \frac{\omega^2 M^2 X_2}{R_2^2 + X_2^2} \right) \right\}$$

Let effective impedance of primary be Z_1', then

$$Z_1' = \frac{V_1}{I_1} = \left(R_1 + \frac{\omega^2 M^2 R_2}{R_2^2 + X_2^2} \right) + j \left(X_1 - \frac{\omega^2 M^2 X_2}{R_2^2 + X_2^2} \right)$$

Therefore

$$\text{equivalent primary resistance} = R_1 + \frac{\omega^2 M^2 R_2}{R_2^2 + X_2^2},$$

$$\text{equivalent primary reactance} = X_1 - \frac{\omega^2 M^2 X_2}{R_2^2 + X_2^2}.$$

We see in above expression for resistance and reactance that due to coupling action, the original primary resistance and reactance have been modified by the components of reflected impedance of secondary into primary, *i.e.*

$$\text{resistive component of reflected impedance} = \frac{\omega^2 M^2 R_2}{R_2^2 + X_2^2},$$

$$\text{reactive component of reflected impedance} = \frac{\omega^2 M^2 X_2}{R_2^2 + X_2^2}.$$

The nature of this coupled impedance largely affects the primary and secondary currents. When the coupled impedance is small, then the primary current is very nearly the same as though no secondary were present and the effects produced in the secondary circuit by the primary current will likewise be small. The coupled impedance will be low if the mutual inductance M is very small (*i.e.*, if there is small coupling) or if the secondary impedance is very high. In contrast, if the coupled impedance is large due to strong coupling or small secondary impedance or both, the voltage and current relations that exist in the primary circuit are then largely affected and a very considerable transfer of energy to the secondary occurs.

2·12-3. THE AIR-CORE AUTO-TRANSFORMER :

When the two impedance coils are arranged in such a way as shown in figure 21, the arrangement is called as an auto-transformer. The auto-transformer works as a step up voltage device, when the driving voltage is applied between the terminals *ab* and the load is connected across the terminals *ac*. The auto-transformer works as step down voltage device when the driving voltage is applied between the terminals *ac* and the load is connected between the terminals *ab* or *bc*.

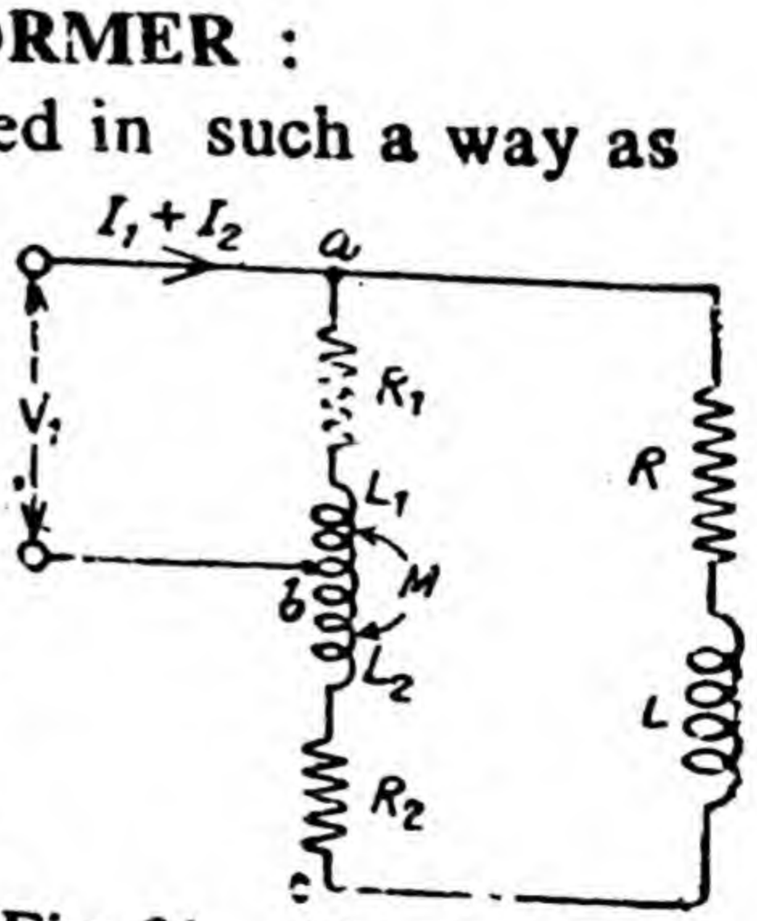


Fig. 21. Air-core auto-transformer.

2·13. BEHAVIOUR OF RESONANT CIRCUITS TUNED TO THE SAME FREQUENCY AND COUPLED TOGETHER :

When the two circuits, resonant at the same frequency are coupled together, the resulting behaviour largely depends upon the degree of coupling. We consider separately the following two cases :

(*a*) *When coefficient of coupling is small.* When the coefficient of coupling (k=·005) is very small, the effect of coupled impedance is negligible. Therefore the curve of primary current as a function of frequency approximately resembles the series resonance curve of primary circuit considered alone as shown in figure 22, (curve for k=·005).

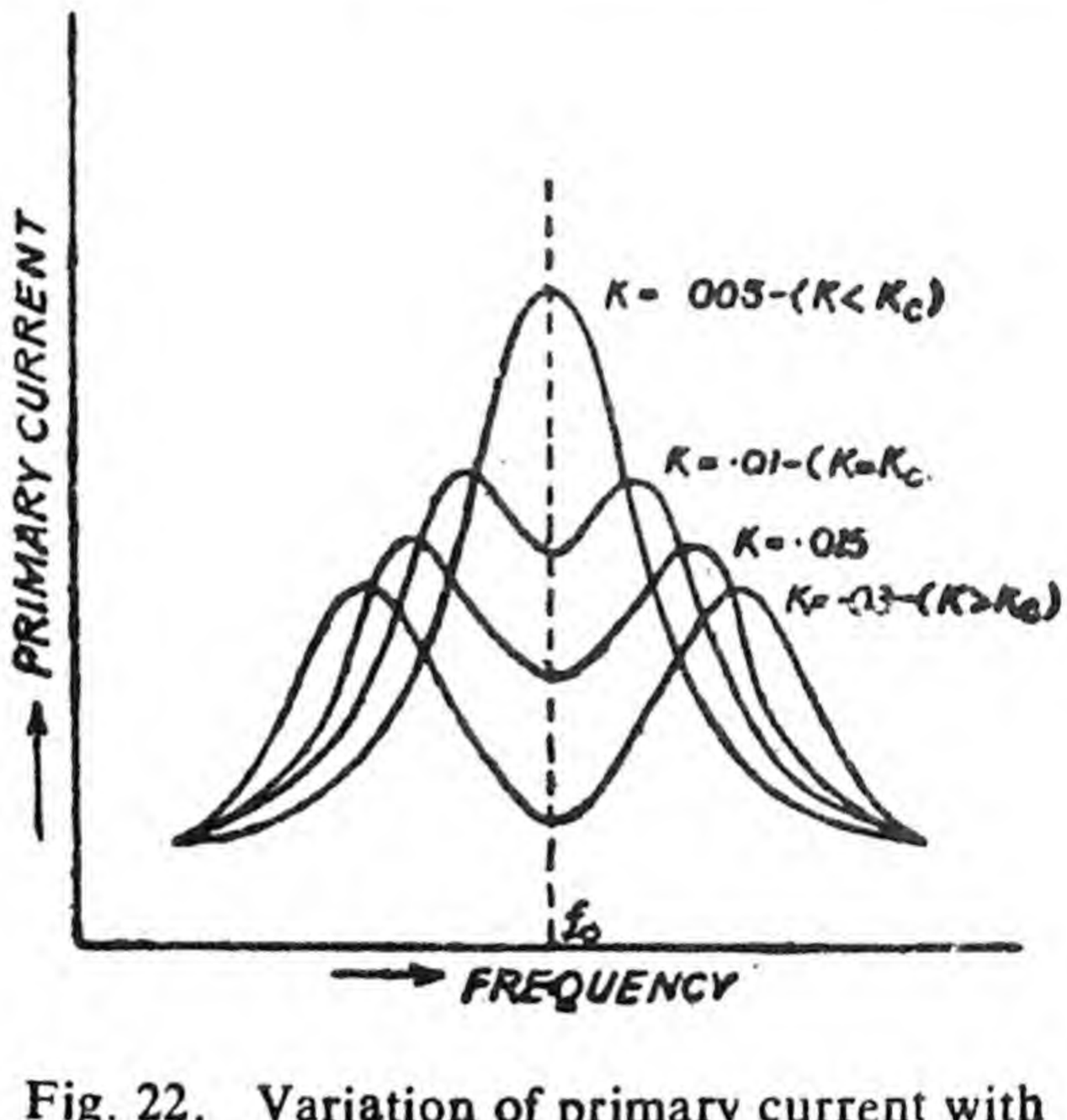


Fig. 22. Variation of primary current with frequency for different coefficient of coupling.

The secondary current at the same time is small and varies with frequency in such a way as to be much more peaked than the resonance curve of the secondary circuit considered as an isolated circuit as shown by curve for k=0·005 in fig. 23.

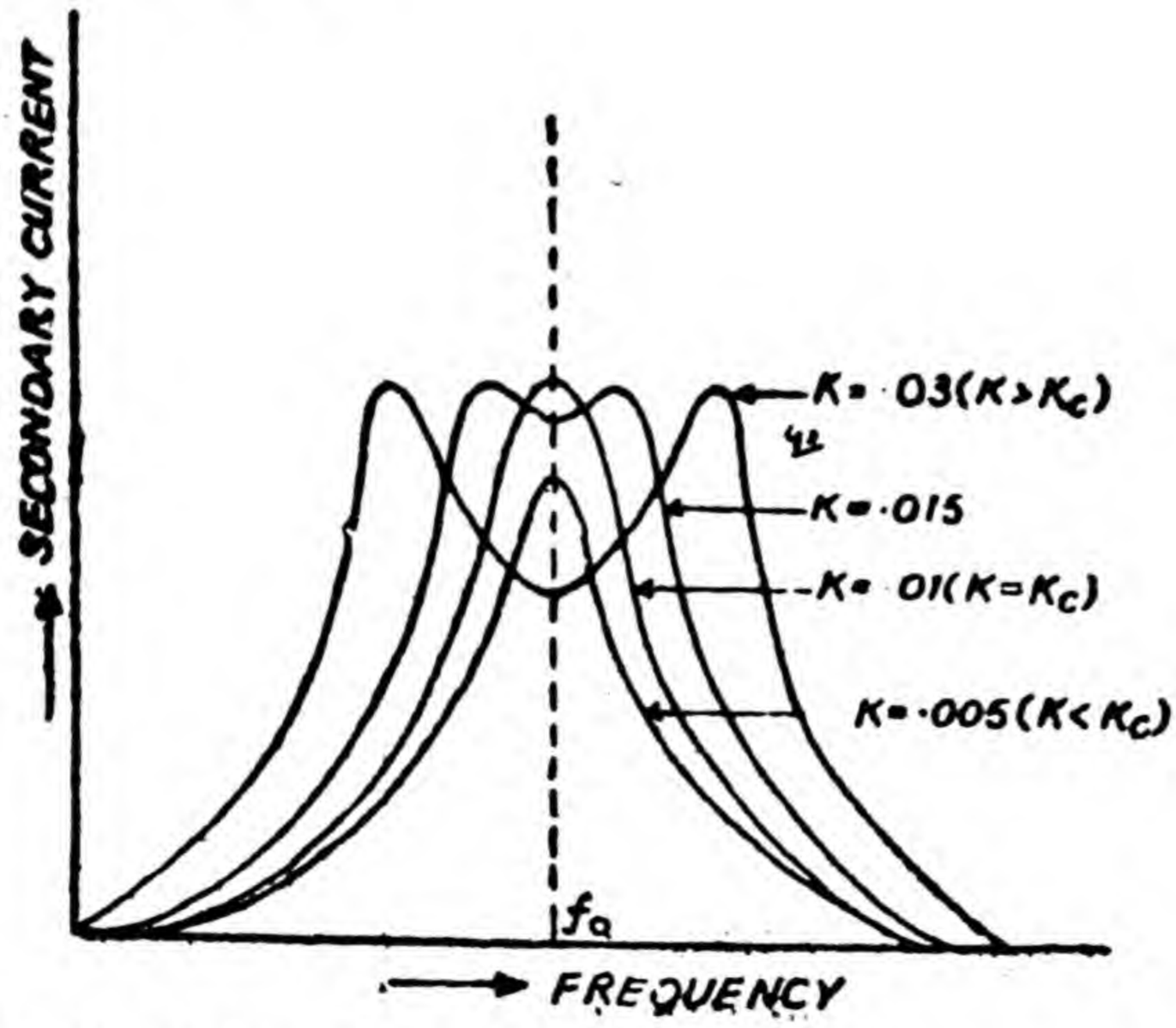


Fig. 23. Variation of secondary current with frequency for different coefficients of coupling.

(*b*) *When coupling increases.* As the coupling is increased, the curve of primary current becomes broader and the peak value of primary current is reduced. At the same time, the secondary current peak becomes higher and the curve of secondary current somewhat broader. This process will continue until a stage is reached at which secondary current has a maximum value. The curve of secondary current is then somewhat broader than the resonance curve of the secondary circuit considered alone, and has a relatively flat top. The primary current now has two peaks because of being greater at frequencies just off resonance than at resonant frequency. All this will happen when the reflected resistance from secondary to the primary is equal to the resistance of the primary. This is known as *critical coupling* and causes the secondary current to have the maximum possible value. Mathematically reflected resistance is

$$R_r=\frac{\omega^2M^2R_2}{R_2^2+X_2^2}$$

At resonance, reactance X_2 is zero,

$$\therefore \quad R_r=\frac{\omega^2M^2}{R_2}.$$

For critical coupling $R_r=R_1$, so that

$$R_1=\frac{\omega^2M^2}{R_2} \text{ and } \omega^2M^2=R_1R_2. \qquad \ldots(1)$$

Since $Q=\frac{L\omega}{R}$, we have

$$R_1=\frac{L_1\omega}{Q_1} \text{ and } R_2=\frac{L_2\omega}{Q_2}.$$

$$\therefore \quad R_1R_2=\frac{\omega^2L_1L_2}{Q_1Q_2}.$$

Putting this value in equation (1), we get

$$\omega^2M^2=\frac{\omega^2L_1L_2}{Q_1Q_2}.$$

$$\therefore \quad \frac{M^2}{L_1L_2}=K^2=\frac{1}{Q_1Q_2}.$$

If $Q_1=Q_2$, then

coefficient of coupling $K=\frac{1}{Q}=K_C$.

If coefficient of coupling is further increased (beyond critical coupling) both primary and secondary currents decrease more at the resonant frequency. The double humps in the primary current become more prominent and the peaks spread further apart. The curve of secondary current also shows double humps. The cause of double humps in the primary current curve is due to the effect of coupled impedance in the primary circuit as it becomes appreciable

when coupling is large. At resonance, the primary current tends to be relatively small because of the very large coupled resistance, while there is a frequency on each side of resonance at which the coupled reactance exactly neutralises the primary reactance, giving zero reactance for the total primary circuit impedance and causing the flow of large current. Thus a double hump is produced. Since the secondary current is a function of the primary current, it tends to have peak values at frequencies where the primary current is at maximum. Hence the secondary current curve also shows the double humps as the coefficient of coupling is increased beyond the critical value, the peaks being spaced further apart as the coupling is increased. It is obvious from fig. 23 ($K=0{\cdot}01$) that when the coefficient of coupling equals the critical value and if $Q_1=Q_2$, then the curve of secondary current as a function of frequeney has the maximum flatness that is possible in the vicinity of resonant frequency.

Width of band pass : As we have deduced previously, the

$$\text{reflected reactance in primary}=-\frac{\omega^2M^2X_2}{X_2^2+R_2^2}.$$

Since resistances of both primary and secondary are small compared with reactances X_1 and X_2 (as we are considering the circuits with high Q values), we can take net reactance of primary equal to

$$X_1-\frac{\omega^2M^2}{X_2}.$$

At resonant frequency f_0, reactance becomes zero, that is

$$X_1-\frac{\omega^2M^2}{X_2}=0.$$

$$\omega^2M^2=X_1X_2$$

$$=\left(\omega L_1-\frac{1}{\omega C_1}\right)\left(\omega L_2-\frac{1}{\omega C_2}\right)$$

$$=\omega^2L_1L_2\left(1-\frac{1}{\omega^2L_1C_1}\right)\left(1-\frac{1}{\omega^2L_2C_2}\right),$$

$$\therefore\quad M^2=L_1L_2\left(1-\frac{1}{\omega^2L_1C_1}\right)\left(1-\frac{1}{\omega^2L_2C_2}\right)$$

$$\therefore\quad M^2=L_1L_2\left(1-\frac{\omega_0^2}{\omega^2}\right)\left(1-\frac{\omega_0^2}{\omega^2}\right).$$

Since $K^2=\frac{M^2}{L_1L_2}$, above equation can be written as

$$K^2=\left(1-\frac{\omega_0^2}{\omega^2}\right)^2$$

$$\therefore \quad \omega_0^2=\omega^2(1\pm K)$$

or $$f_0^2=f^2(1\pm K)$$

or $$f=\frac{f_0}{\sqrt{(1\pm K)}}=\frac{f}{1\pm\frac{K}{2}} \text{ approximately.}$$

Above relation shows two frequencies one less and other greater than the resonant frequency f_0, *i e.*

$$f_1=\frac{f_0}{1+\frac{K}{2}}, \qquad f_2=\frac{f_0}{1-\frac{K}{2}}.$$

We have seen from curves for secondary and primary currents that the double hump occurs at frequencies slightly less and greater than resonant frequency. The width of passband can be shown to be equal to

$$f_2-f_1=f_0\left(1-\frac{K}{2}\right)^{-1}-f_0\left(1+\frac{K}{2}\right)^{-1}=f_0K,$$

$$\therefore \quad \Delta f=f_0K=\frac{f_0}{Q}, \text{ since } K=\frac{1}{Q}$$

From above equation it is obvious that width of passband depends upon coefficient of coupling and varies inversely as Q value of the circuits.

2·14. FIDELITY OF RECEIVERS :

In order to obtain high fidelity and high selectivity the ideal response curve should have a flat top and straight sides. An approach to this ideal condition is made by using two resonant circuits tuned to the same frequency and coupled to each other. The circuit is called a band-pass amplifier and used as an intermediate frequency amplifier in super-heterodyne receiver.

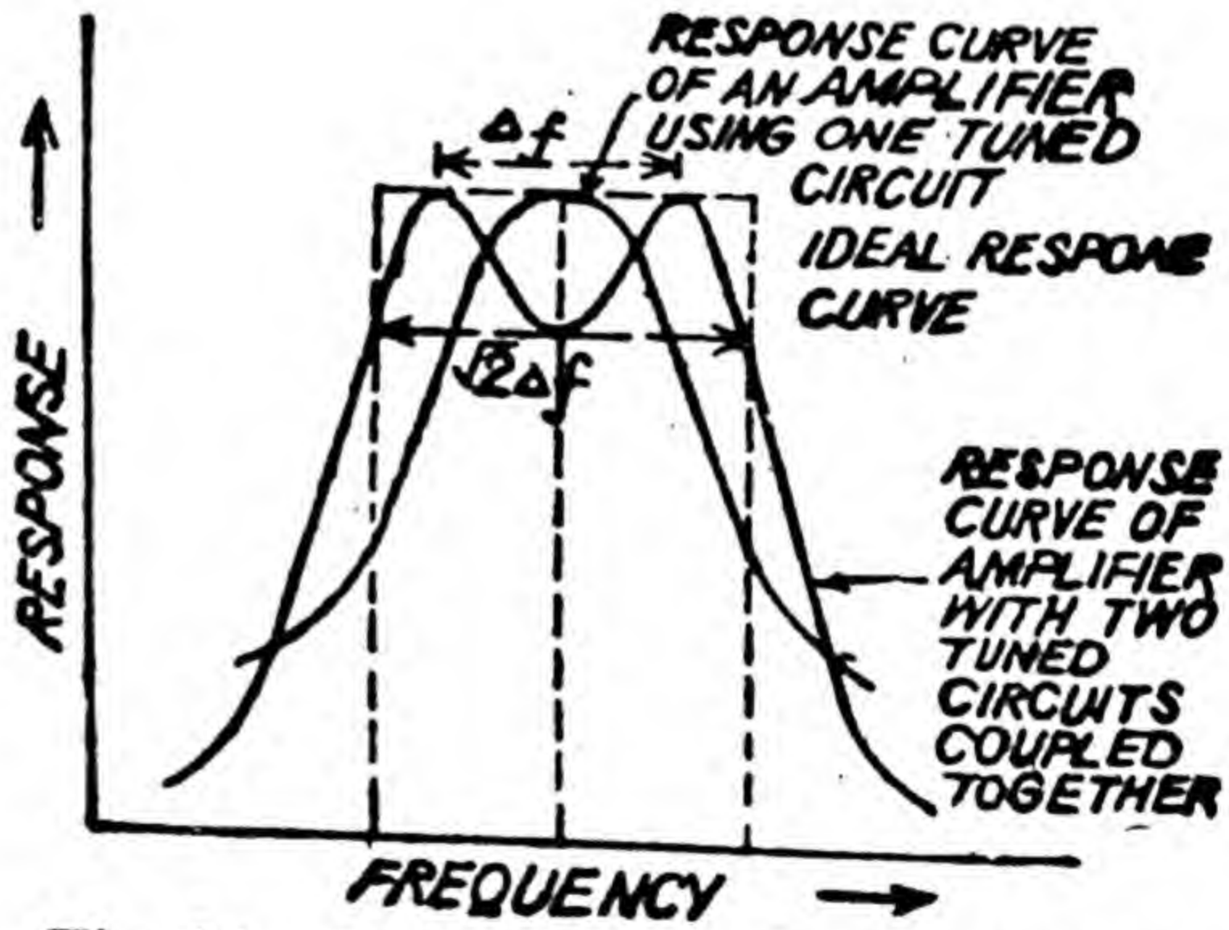


Fig. 24. Response curve of amplifier with tunned circuit.

The slope of the sides of the response curve will vary with the Q of the circuit. Increasing the value of Q will make the curve steeper and provide greater attenuation of the signals of frequencies outside desired band. Thus high Q circuit has better selectivity, *i.e.* sufficient ability to discriminate against undesired signals. In intermediate frequency circuits, where high fidelity to response is of major importance, overcoupling is frequently used to broaden the nose of the curve by double peaking. This passes efficiently the

upper side bands due to higher (audio frequency) modulating frequencies. For most purposes, the best bandpass characteristic is that which corresponds to a coefficient of coupling equal to the critical value ($K=K_C$).

Example. To calculate the Q of the tuned circuits when the bandwidth between point of the curve (fig. 24), where gain equals the gain at resonant frequency one megacycles is 20 *kc/s*.

It can be shown that the relation for the bandwidth between points on the curve where gain equals the gain at centre frequency is

$$(\Delta f)_{resonance}=\sqrt{2}.(\Delta f)_{secondary\ current\ peaks}$$
$$=\sqrt{2}f_0.K$$
$$=\sqrt{2}\,\frac{f_0}{Q}=\frac{f_0}{Q_{effective}}$$

where $$Q_{effective}=\frac{Q}{\sqrt{2}}.$$

Therefore $$Q_{effective}=\frac{10^6}{20.10^3}=50$$

and $$Q_{circuit}=\sqrt{2}\ Q_{effective}=50\sqrt{2}=70{\cdot}5.$$

NUMERICAL EXAMPLES

1. *A series circuit contains an inductance of one henry, a capacitance of 16·5 microfarads and a resistance of 120 ohms. An alternating e.m.f. of 100 volts, 50 cycles is applied to the circuit. Find the value of current and the phase difference between the impressed voltage and current.*

The imdedance of the series circuit is given by

$$Z=\left[R^2+\left(\omega L-\frac{1}{\omega C}\right)^2\right]^{1/2}$$
$$=\left[(120)^2+\left\{2\pi\times50\times1-\frac{10^6}{2\pi\times50\times16{\cdot}5}\right\}^2\right]^{1/2}$$
$$=[(120)^2+(314-194)^2]^{1/2}=168\text{ ohms.}$$

Current $$I=\frac{E}{Z}=\frac{100}{168}=0{\cdot}6\text{ amperes.}$$

The phase difference θ is given by

$$=\tan^{-1}\frac{\omega L-1/\omega C}{R}=\tan^{-1}\frac{120}{120}=45^\circ$$

The applied voltage is ahead of current by 45°.

2. *Find the resonant frequency of a circuit containing an inductance of 2 henrys and a capacitance of 8μF.*

We know that $$f=\frac{1}{2\pi\sqrt{(LC)}}$$
$$=\frac{1}{6{\cdot}28\sqrt{(2\times8\times10^{-6})}}$$

$$= \frac{1}{6\,28 \times 4 \times 10^{-3}}$$

$$= \frac{1}{0{\cdot}0251} = \mathbf{39{\cdot}8\ c/s.}$$

3. ***Find the resonant frequency of a circuit containing an inductance of*** $300\mu H$ ***and a capacitance of*** $500\mu\mu F$.

$$300\mu H = 300 \times 10^{-6} H = 3 \times 10^{-4} H$$

$$500\mu\mu F = 500 \times 10^{-12} \text{ farad} = 5 \times 10^{-10} \text{ farad}$$

$$f = \frac{1}{2\pi\sqrt{(LC)}}$$

$$= \frac{1}{6{\cdot}28\sqrt{(3 \times 5 \times 10^{-4} \times 10^{-10})}}$$

$$= \frac{1}{6{\cdot}28 \times 10^{-7} \times 3{\cdot}8}$$

$$= \frac{1}{24{\cdot}35 \times 10^{-7}}$$

$$= \mathbf{412000\ c/s.}$$

4. ***A series LCR circuit has*** $L = 100$ ***micro-henrys,*** $C = 100$ ***microfarad and*** $R = 0{\cdot}1$ ***ohms. Find the resonant frequency and the temperature of the circuit at this frequency. A voltage of 5 millivolts at the resonant frequency is applied to the circuit. Determine the resulting voltage across L. What is the impedance of the series circuit at a frequency less than the resonant frequency by*** 5000 ***cycles*** ?

Given $L = 100.10^{-6}$ henry, $C = 100.10^{-12}$ farad, $R = 0{\cdot}1$ ohms,

$V = 5.10^{-3}$ volts.

The resonant frequency of a series circuit is

$$f_0 = \frac{1}{2\pi\sqrt{(LC)}} = \frac{1}{2\pi\sqrt{(100.10^{-6}.100.10^{-12})}} = \frac{10^7}{2\pi}$$

$$= \mathbf{1{\cdot}59.10^6\ c/s.}$$

Impedance at resonant frequency in series LCR circuit is purely resistive *i.e.*

Impedance at resonance $= R = \mathbf{0{\cdot}1\ omhs.}$

Now the current at resonant frequency $= \frac{V}{R}$ amp.

$$= \frac{5{\cdot}10^{-3}}{0{\cdot}01}$$

$$= {\cdot}05 \text{ amp}$$

Therefore. voltage across $L = I.\omega L$

$$= {\cdot}05.2\pi f_0 L$$

$$= {\cdot}05.2\pi.\frac{10^7}{2\pi}.\ 100.10^{-6}$$

$$= \mathbf{50\ volts.}$$

Now impedance of series LRC circuit at any frequency is

$$Z=\sqrt{\left\{R^2+\left(\omega L-\frac{1}{\omega C}\right)^2\right\}}$$

$$=\sqrt{\left\{R^2+\left(2\pi f\,L-\frac{1}{2\pi f.C}\right)^2\right\}}$$

where $f=f_0-5000$

$=159.10^6-5.10^3$

$=159.10^4-0{\cdot}5.10^4=158{\cdot}5.10^4$ c/s.

$$\therefore\quad Z=\sqrt{\left\{(\cdot1)^2+\left(2\pi.158{\cdot}5.10.100.10^{-6}-\frac{1}{2\pi.158{\cdot}5.10^4\ 100.10^{-12}}\right)^2\right\}}$$

$$=\sqrt{\left\{\cdot01+\left(317\pi-\frac{10^6}{317\pi}\right)^2\right\}}$$

$$=\sqrt{\left\{\cdot01+\left(995{\cdot}4-\frac{10^6}{995{\cdot}4}\right)^2\right\}},$$

$=\sqrt{\{\cdot01+92{\cdot}16\}}$

$=9{\cdot}601$ ohms.

EXERCISES AND PROBLEMS

1. When has a series circuit its lowest impedance ?
2. What happens to the impedance of a series resonant circuit when the frequency is changed from the resonant frequency ?
3. What is the effect on the resonant frequency of the circuit by : (*a*) doubling the inductance, (*b*) halving the capacitance, (*c*) doubling the capacitance and halving the inductance.
4. Show by drawing several resonance curves on the same axis, the effect of increasing the resistance of a series resonant circuit. What is the effect of increasing resistance on the maximum current through the circuit ?
5. What is effect on impedance of tuning a parallel circuit to resonance ?
6. How can a parallel resonance circuit be used to reject a radio-signal ?
7. State similarities and differences between series and parallel resonance.
8. What is the resonant frequency of a circuit containing an inductance of 6 henrys and a capacitance of 9 μF ?
9. What value of inductance when placed in series with a 20μF condenser will produce resonance at a frequency of 50 cycles per second ?
10. What value of capacitance when placed in series with an inductance of 40μH, will resonate to a frequency of 2000 kc ?
11. A high Q parallel resonant circuit is being used as a receiver tuning circuit. A variable capacitor with a minimum capacitance of 20μF, and a maximum capacitance of 350 $\mu\mu F$, is used to adjust the resonance frequency. (*a*) What value of fixed inductance should be used to make 530 kc the lowest obtainable resonance frequency ? (*b*) Using the value of inductance obtained in part (*a*), what will be the frequency range of the circuit ?
12. In a series RLC circuit, the impressed e. m. f.$=440$V, $L=5$ mH. As the operating frequency is varied, the maximum current is found to be 5 amp.

and with this current, the voltage drop across L exceeds that across R by a factor of 10. What is the resonant frequency of the circuit ?

13. Two identical inductors, each having an effective resistance of 10 ohms and a Q of 18·72 at an operating frequency of 1000 cycles, have a coefficient of coupling of 0·2 between them. What is the value of the mutual impedance ?

14. A coil of inductance 5 mH and resistance 50 ohms is joined in parallel with a capacitor of $500\mu\mu F$. These two together are joined in series with a non-inductive resistor of 50,000 ohms to a source of alternating p. d. of magnitude 1 V. and frequency 100 kc/s. Calculate the p. d. between the terminals of the capacitor.

15. An alternating voltage of 100 V (r. m. s) at 50 cycles/sec. is applied to a circuit consisting of a capacitor of $2\mu F$ in parallel with a coil having an inductance of 5H and a resistance of 100 ohms. What will the amplitude and phase of the current through the inductor and the amplitude and phase of the current taken from the mains ?

16. A circuit consist of a resistance R in parallel with a coil having an inductance L and resistance r; in series with the combination is a capacitance C. An alternating sinusoidal voltage is applied to the circuit in which $C=1F$, $L=0{\cdot}08H$, $r={\cdot}001$ ohms and $R=250$ ohms.
If the current in the inductive resistance is 10 mA when the angular frequency of the supply $\omega=5000$ rad/sec., determine by means of a vector diagram or otherwise,
(a) the current in R,
(b) the total current supplied,
(c) the total power dissipated in the circuit and,
(d) the voltage applied across the circuit.

3 FOUR TERMINAL NETWORKS AND NETWORK THEOREMS

3·0. FOUR TERMINAL NETWORKS :

A four terminal network is provided with two pairs of terminals one of which, for reference, is termed as input while other as output. The associated electrical quantities at the two pairs of terminals are the voltages E_1 and E_2, and the currents I_1 and I_2. Their reference direction is shown in fig. (1). With the change of these pairs, signs of some of the associated terms are changed.

At any frequency, the properties of a four terminal network can be expressed in terms of any three independent properties of the network. These independent constants can be defined as :

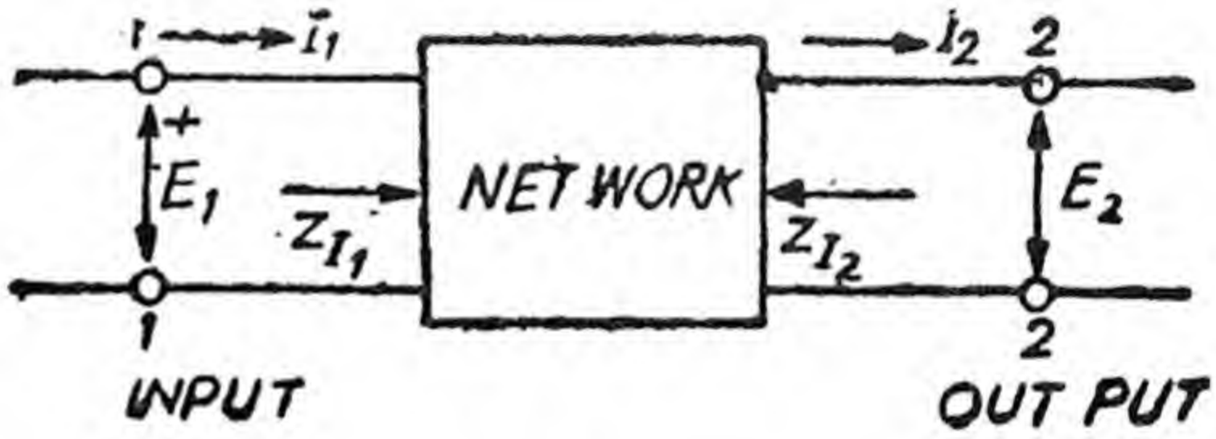


Fig. .

(*i*) image impedances and image transfer constant.
(*ii*) open and short circuited impedances, and
(*iii*) iterative impedances and iterative transfer constant.

We shall deal with all these constants in further articles.

Let Z_{11} be the impedance between the input terminals without any consideration to output terminals effect then :

$Z_{11}=\frac{E_1}{I_1}$, when $I_2=0$ *i.e.*, terminals 2–2 are open circuited.

Similarly, the impedance Z_{22} between output terminals will be

$Z_{22}=\frac{E_2}{I_2}$, when $I_1=0$ *i.e.*, terminals 1–1 are open circuited.

But if we consider a sort of coupling between input and output, which of course is the case, then one should think about mutual impedances or transfer impedances too *i.e.*,

$R_{12}=\frac{E_1}{I_2}$, when $I_1=0$ called the transfer impedance with 1–1 open circuited.

and $Z_{21}=\frac{E_2}{I_1}$, when $I_2=0$ called the transfer impedance with 2–2 open-circuited.

Thus the voltage and current relations can be put in the form

$$Z_{11}I_1+Z_{12}I_2=E_1$$

and

$$Z_{21}I_1+Z_{22}I_2=E_2.$$

Expressing in matrix form,

$$\begin{Bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{Bmatrix}\times\begin{Bmatrix} I_1 \\ I_2 \end{Bmatrix}=\begin{Bmatrix} E_1 \\ E_2 \end{Bmatrix}, \text{ called } Z \text{ matrix}$$

$$\begin{Bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{Bmatrix}\times\begin{Bmatrix} E_1 \\ E_2 \end{Bmatrix}=\begin{Bmatrix} I_1 \\ I_2 \end{Bmatrix}, \text{ called } Y\text{-matrix}$$

or

$$\begin{Bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Bmatrix}\times\begin{Bmatrix} E_2 \\ I_2 \end{Bmatrix}=\begin{Bmatrix} E_1 \\ I_1 \end{Bmatrix}, \text{ called } A\text{-matrix.}$$

In the latter matr it is quite easy to see that

$$A_{11}=\frac{Z_{11}}{Z_{21}},$$

$$A_{11}=Z_{12}-\frac{Z_{22}\times Z_{11}}{Z_{21}},$$

$$A_{21}=\frac{1}{Z_{21}} \text{ and } A_{22}=\frac{Z_{22}}{Z_{21}}.$$

Image Parameters—If internal impedance Z_g of generators between input terminals is related to the load impedance Z_L between output terminals such that the impedance looking into the network from terminals 1-1 with the load connected equal to the generator impedance, and similarly the impedance looking into the network from terminal 2–2 with the generator connected equals the load impedance, then the network is said to be operated under image-impedance conditions. The generator and load impedance required to produce this condition are properties of the network, termed as image impedances, which in this case can be designated as Z_{I1} $(=Z_S)$ and Z_{I2} $(=Z_L)$ for the input terminals respectively fig. (2).

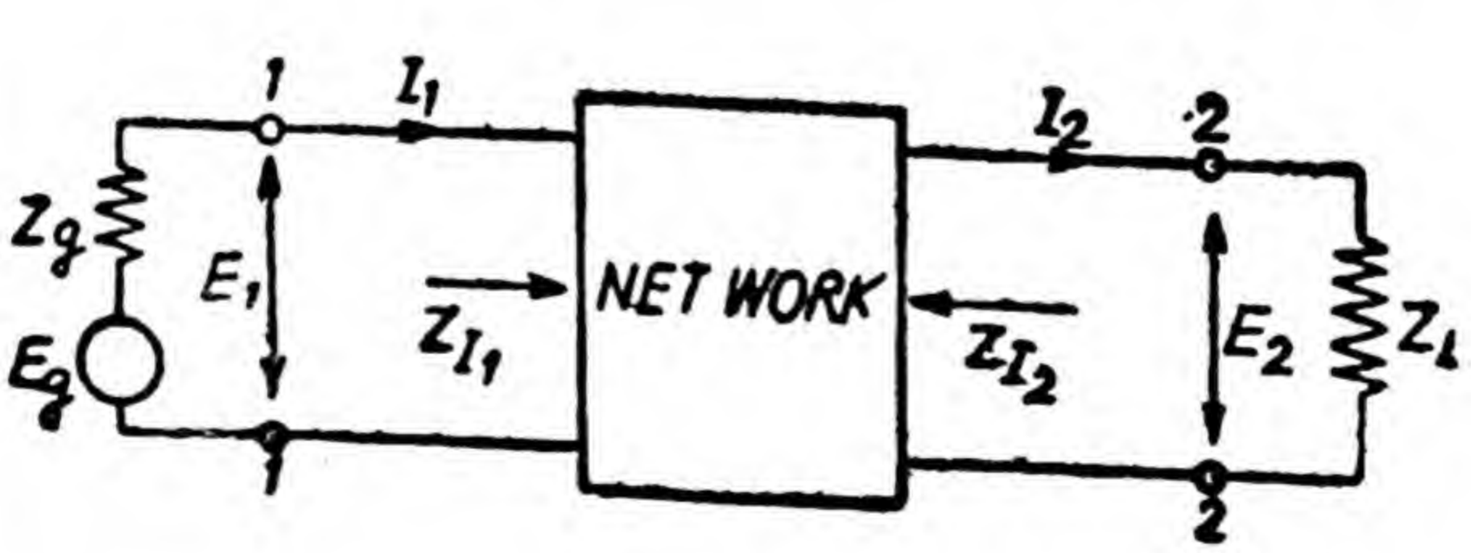


Fig. 2.

Thus, the two image impedances of a four terminal network Z_{i1} and Z_{I2} are two impedances such that, if end 1 of the network is terminated in Z_{I1}, then the impedance measured at end 2 is Z_{I2} and *vice versa.*

To find a relation in image and linear parameters, we proceed as follows :

Let the output terminals be terminated in Z_{I_2} fig. (3). Then

$$E_2 = Z_{I_2} I_2,$$
$$E_1 = Z_{I_1} I_1,$$

We know that

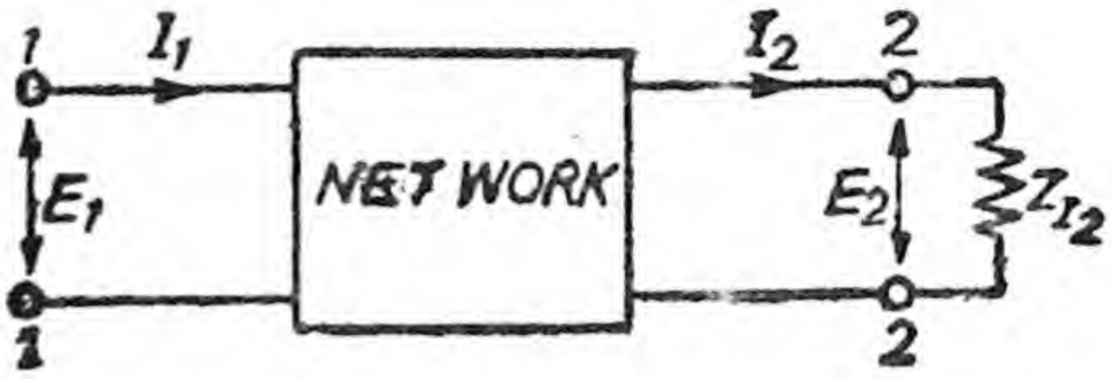


Fig. 3.

$$\begin{Bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Bmatrix} \times \begin{Bmatrix} E_2 \\ I_2 \end{Bmatrix} = \begin{Bmatrix} E_1 \\ I_1 \end{Bmatrix}$$

i.e.,
$$A_{11}E_2 + A_{12}I_2 = E_1, \quad \ldots(1)$$
$$A_{21}E_2 + A_{22}I_2 = I_1. \quad \ldots(2)$$

Putting value of E_2 and E_1 in (1) and (2), we get

$$Z_{I_1} I_1 = (A_{11}Z_{I_2} + A_{12}) I_2$$

and
$$I_1 = (A_{21}Z_{I_2} + A_{22}) I_2.$$

Dividing both,

$$Z_{I_1} = \frac{A_{11}Z_{I_2} + A_{12}}{A_{21}Z_{I_2} + A_{22}}. \quad \ldots(3)$$

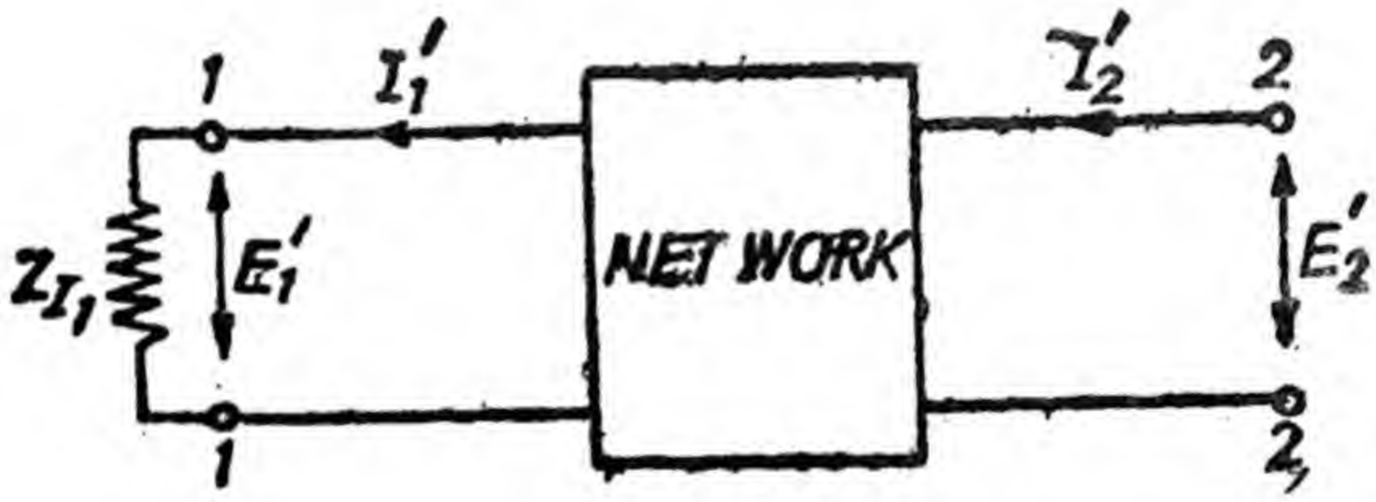


Fig. 4,

Let the terminals 1-1 be terminated in Z_{I_1}, fig. (4), then

$$E_1' = Z_{I_1} I_1',$$
$$E_2' = Z_{I_2} I_2'.$$

We know that

$$\begin{Bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Bmatrix} \times \begin{Bmatrix} E_2' \\ -I_2' \end{Bmatrix} = \begin{Bmatrix} E_1' \\ -I_1' \end{Bmatrix},$$

i.e.,
$$A_{11}E_2' - A_{12}I_2' = E_1' \quad \ldots(4)$$
and
$$A_{21}E_2' - A_{22}I'_2 = -I_1' \quad \ldots(5)$$

Putting value of E_1' and E_2' in above equation, we get

$$Z_{I_1}I_1' = A_{11}Z_{I_2}I_2' - A_{12}I_2'$$

and
$$-I_1' = A_{21}Z_{I_2}I_2' - A_{22}I_2'.$$

Dividing both,

$$Z_{I_1} = \frac{A_{11}Z_{I_2} - A_{12}}{-A_{21}Z_{I_2} + A_{22}} \quad \ldots(6)$$

Therefore from equations (3) and (6), we get

$$\frac{A_{11}Z_{I_2}+A_{12}}{A_{21}Z_{I_2}+A_{22}}=-\frac{A_{11}Z_{I_2}-A_{12}}{A_{12}Z_{I_2}+A_{22}}$$

or $-A_{11}A_{21}Z^2{}_{I_2}+A_{11}A_{22}Z_{I_2}+A_{12}A_{22}-A_{12}A_{21}Z_{I_2}$

$$=A_{11}A_{21}Z_{I_2}{}^2-A_{12}A_{21}Z_{I_2}+A_{11}A_{21}Z_{I_2}-A_{12}A_{22}$$

or $$A_{11}A_{21}Z_{I_2}{}^2=A_{12}A_{22}$$

$$\therefore \quad Z_{I_2}{}^2=\frac{A_{12}}{A_{21}}.\frac{A_{22}}{A_{11}} \qquad \text{...(7)}$$

Similarly it can be found from equations (3) or (4) that

$$Z_{I_2}{}^2=\frac{A_{12}A_{11}}{A_{21}A_{22}}. \qquad \text{...(8)}$$

Image transfer constant γ : This constant, along with two image impedance ZI_1 and ZI_2, completely specifies the performance of a network. This is defined as one half the natural logarithm of the complex ratio of the steady state volt amperes entering and leaving the network then the network is terminated in its impedance, *i.e.*,

$$\gamma=\tfrac{1}{2}\log_e\frac{E_1I_1}{E_2I_2} \quad \text{or} \quad e^{2\gamma}=\frac{E_1I_1}{E_2I_2}$$

Putting values of E_1 and I_1 from equations (1) and (2), we get

$$e^{2\gamma}=\frac{(A_{11}E_2+A_{12}I_2)(A_{21}E_2+A_{22}I_2)}{E_2I_2}.$$

Putting $E_2=Z_{I_2}.I_2$ (since it is terminated in its image impedance) in above expression, we get

$$e^{2\gamma}=\frac{(A_{11}Z_{I_2}I_2+A_{12}I_2)(A_{21}Z_{I_2}I_2+A_{22}I_2)}{Z_{I_2}I_2{}^2}$$

$$=(A_{11}+A_{12}/Z_{I_2})\ (A_{21}Z_{I_2}+A_{22}).$$

But we have shown that

$$Z_{I_2}=\frac{A_{12}}{A_{21}}\times\frac{A_{22}}{A_{11}},$$

Hence $e^{2\gamma}=\{A_{11}+A_{12}\sqrt{}(A_{21}A_{11}/A_{12}A_{22})\}\times\{A_{21}\sqrt{}(A_{12}A_{22}/A_{21}A_{11})+A_{22}\}$

$$=\sqrt{}(A_{11}/A_{22})\ \{\sqrt{}(A_{11}A_{22})+\sqrt{}(A_{12}A_{21})\}$$

$$\times\sqrt{}(A_{22}/A_{11})\ \{\sqrt{}(A_{21}A_{12})+\sqrt{}(A_{11}A_{22})\}$$

$$\therefore\ e^{\gamma}=\sqrt{}(A_{11}A_{22})+\sqrt{}(A_{12}A_{21}).$$

Also $$e^{-\gamma}=\frac{1}{\sqrt{}(A_{11}A_{22})+\sqrt{}(A_{12}A_{21})}$$

or $$e^{-\gamma}=\frac{\sqrt{}(A_{11}A_{22})-\sqrt{}(A_{12}A_{21})}{A_{11}A_{22}-A_{12}A_{21}}.$$

But $$A_{11}A_{22}-A_{12}A_{21}=1.$$

$$\therefore\ e^{-\gamma}=\sqrt{}(A_{11}A_{22})-\sqrt{}(A_{12}A_{21})$$

$$\left.\begin{aligned} \text{Therefore } \sinh\gamma &= \frac{e^{\gamma}-e^{-\gamma}}{2} = \sqrt{(A_{12}A_{21})}, \\ \cosh\gamma &= \frac{e^{\gamma}+e^{-\gamma}}{2} = \sqrt{(A_{11}A_{22})}, \\ \tanh\gamma &= \sqrt{\left(\frac{A_{12}A_{21}}{A_{11}A_{22}}\right)} \end{aligned}\right\} \quad \ldots(9)$$

The expression for $\tanh\gamma$ *and image impedances can also be obtained in terms of open or short circuit driving-point impedances at each pair of terminals.*

Let Z_{01} = Input impedance at end 1 with terminals at end 2 open circuited.

Z_{02} = Input impedance at end 2 with the terminals at end 1 open-circuited.

Z_{s1} = Input impedance at end 1 with terminals at end 2 short-circuited.

Z_{s2} = Input impedance at end 2 with the terminals at end 1 short-circuited.

We have seen that

$$E_1 = A_{11}E_2 + A_{12}I_2,$$
$$I_1 = A_{21}E_2 + A_{22}I_2.$$

Now when end 2 is open, $I_2 = 0$.

$$\therefore \quad Z_{01} = \frac{E_1}{I_1} = \frac{A_{11}}{A_{21}}.$$

and when end 1 is open, $I_1 = 0$,

$$\therefore \quad Z_{02} = \frac{E_2}{(-I_2)} = \frac{A_{22}}{A_{21}}$$

Also when end 2 is shorted, $E_2 = 0$,

$$Z_{s1} = \frac{E_1}{I_1} = \frac{A_{12}}{A_{22}},$$

and when end 1 is shorted, $E_1 = 0$,

$$\therefore \quad Z_{s1} = \frac{E_2}{(-I_2)} = \frac{A_{12}}{A_{11}}.$$

The expressions for image impedances are then as

$$\left.\begin{aligned} Z_{I_1}^2 &= \frac{A_{11}}{A_{21}} \times \frac{A_{12}}{A_{22}} \\ &= Z_{01} \times Z_{s1} \\ \text{and} \qquad Z_{I_2}^2 &= \frac{A_{22}}{A_{11}} \times \frac{A_{12}}{A_{11}} \\ &= Z_{02} \times Z_{s2} \\ \text{Also} \qquad \tanh^2\gamma &= \frac{A_{12}}{A_{22}} \times \frac{A_{21}}{A_{11}} = \frac{Z_{s1}}{Z_{01}} \\ &= \frac{A_{12}}{A_{11}} \times \frac{A_{21}}{A_{22}} = \frac{Z_{s2}}{Z_{02}} \end{aligned}\right\} \quad \ldots(10)$$

It is now quite simple to express every '*A*' parameter in terms of two image impedances Z_{I1} and Z_{I2}, and the image transfer constant γ as discussed below :

We have shown previously,

$$\sinh^2 \gamma = A_{12}A_{21},\ \cosh^2 \gamma = A_{11}A_{22}$$

and $$Z_{I1}.Z_{I2} = \sqrt{\left(\frac{A_{11}}{A_{21}}.\frac{A_{12}}{A_{22}}\right)}.\sqrt{\left(\frac{A_{22}}{A_{21}}.\frac{A_{12}}{A_{11}}\right)} = \frac{A_{12}}{A_{21}}$$

$$\frac{Z_{I1}}{Z_{I2}} = \frac{A_{11}}{A_{22}}.$$

Hence $$A_{11}^2 = [A_{11}A_{22}] \times \frac{A_{11}}{A_{22}} = \cosh^2 \gamma \frac{Z_{I2}}{Z_{I1}}$$

or $$A_{11} = \cosh \gamma . \sqrt{\left(\frac{Z_{I1}}{Z_{I2}}\right)}$$

and $$A_{22}^2 = [A_{11}A_{22}].\frac{A_{22}}{A_{11}} = \cosh^2 \gamma \frac{Z_{I2}}{Z_{I1}}$$

or $$A_{22} = \cosh \gamma \sqrt{\left(\frac{Z_{I2}}{Z_{I1}}\right)}$$

Similarly, $$A_{12}^2 = [A_{12}A_{21}] \times \frac{A_{12}}{A_{21}} = \sinh^2 \gamma \times Z_{I1}\, Z_{I2}$$

$$A_{12} = \sinh \gamma \sqrt{(Z_{I1}\, Z_{I2})}$$

and $$A_{21} = \frac{\sinh \gamma}{\sqrt{(Z_{I1}\, Z_{I2})}}.$$

Thus '*A*' matrix of the network can be written as

$$[A] = \begin{Bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Bmatrix}$$

$$\begin{pmatrix} \cosh \gamma \left(\frac{Z_{I1}}{Z_{I2}}\right) & \sinh \gamma \sqrt{(Z_{I1}\, Z_{I2})} \\ \frac{\sinh \gamma}{\sqrt{(Z_{I1}\, Z_{I2})}} & \cosh \gamma \Big/ \left(\frac{Z_{I2}}{Z_{I1}}\right) \end{pmatrix} \qquad \ldots(11)$$

Now for an example, we shall calculate the image impedance and image transfer constant γ of *L*-section as shown in figure (5).

It is obvious that

$Z_{01} = \frac{1}{2}Z_1 + 2Z_2$,

$Z_{s1} = \frac{1}{2}Z_1$,

$Z_{I1} = \sqrt{(Z_1Z_2 + \frac{1}{4}Z_1^2)}$

$Z_{02} = 2Z_2$,

$Z_{s2} = \dfrac{\frac{1}{2}Z_1 \times 2Z_2}{\frac{1}{2}Z_1 + 2Z_2}$

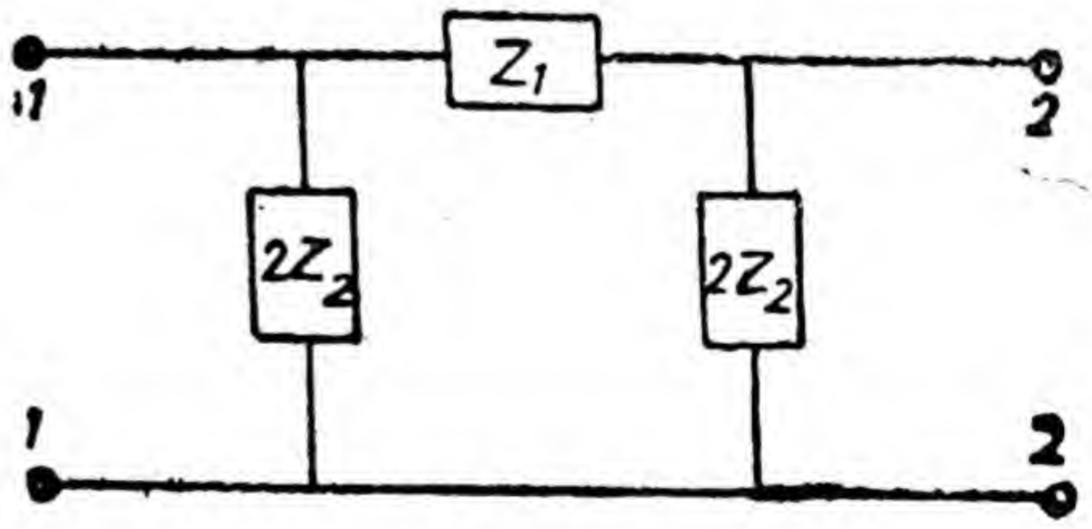


Fig. 5.

$$\therefore\ Z_{I_2}=\sqrt{(Z_{02}\times Z_{s2})}=\sqrt{\left(\frac{2Z_1\cdot Z_2^2}{\frac{1}{2}Z_1+2Z_2}\right)}=\frac{Z_1Z_2}{\sqrt{(Z_1Z_2+\frac{1}{4}Z_1^2)}}$$

and $$\tanh^2\gamma=\frac{Z_{s1}}{Z_{01}}=\frac{\frac{1}{2}Z_1}{\frac{1}{2}Z_1+2Z_2}$$

or $$\coth^2\gamma=1+\frac{4Z_2}{Z_1}$$

or $$\frac{\cosh^2\gamma}{\sinh^2\gamma}-1=\frac{4Z_2}{Z_1}$$

$$\frac{1}{\sinh^2\gamma}=\frac{4Z_2}{Z_1}$$

$$\therefore\quad \sinh^2\gamma=\frac{Z_1}{4Z_2}$$

$$\cosh 2\gamma=1+2\sinh^2\gamma=1+\frac{Z_1}{2Z_2}$$

or $$2\gamma=\cosh^{-1}\left(1+\frac{Z_1}{2Z_2}\right),$$

$$\gamma=\tfrac{1}{2}\cosh^{-1}\left(1+\frac{Z_1}{2Z_2}\right).$$

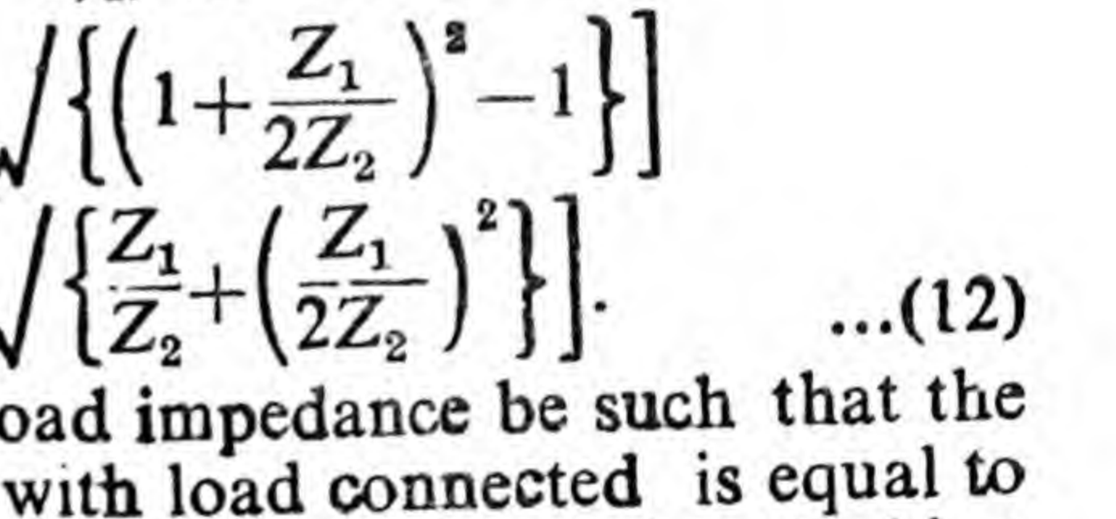
Since $\cosh^{-1}x=\log[x+\sqrt{(x^2-1)}]$, we can have

$$\gamma=\tfrac{1}{2}\log\left[\left(1+\frac{Z_1}{2Z_2}\right)+\sqrt{\left\{\left(1+\frac{Z_1}{2Z_2}\right)^2-1\right\}}\right]$$

$$=\tfrac{1}{2}\log\left[\left(1+\frac{Z_1}{2Z_2}\right)+\sqrt{\left\{\frac{Z_1}{Z_2}+\left(\frac{Z_1}{2Z_2}\right)^2\right\}}\right]. \qquad ...(12)$$

Iterative Parameters : If the load impedance be such that the input impedance of the network with load connected is equal to the load impedance and also the impedance observed by looking into the network from the output terminals towards the generator with the internal impedance of the generator connected across the input terminals of the network equals the generator impedance, then the network is said to be operating under iterative impedance basis. These two impedance are designated as Z_{1K} and Z_{2K} respectively and are the properties of network.

Thus the iterative impedance of a network is that impedance which, when connected to one pair of terminals produces a like impedance at the other pair of terminals.



Fig, 6.

Let the second pair be terminated in Z_{1K}. The impedance at 1 is Z_{1K} provided it is the iterative impedance there. Now

$$Z_{1K}=Z_1+\frac{Z_3\times(Z_2+Z_{1K})}{Z_2+Z_3+Z_{1K}}$$ (from fig. 6.)

or $Z_{1K}Z_2+Z_{1K}Z_3+Z_{1K}^2=Z_1Z_2+Z_1Z_3+Z_1Z_{1K}+Z_2Z_3+Z_3Z_{1K}$

or $Z_{1K}^2+Z_{1K}(Z_2-Z_1)-(Z_1Z_2+Z_2Z_3+Z_3Z_1)=0,$

which is a quadratic equation in Z_{1K} and therefore the roots of the equation are

$$Z_{1K}=\frac{(Z_1Z_2)}{2}\pm\tfrac{1}{2}\sqrt{[(Z_2-Z_1)^2+4(Z_1Z_2+Z_2Z_3+Z_3Z_1)]}. \quad ...(13)$$

Now to calculate Z_{2K}, let the terminal 1 be terminated in Z_{2K} and the impedance at 2 is Z_{2K}, provided it is iterative impedance there. Therefore

$$Z_{2K}=Z_2+\frac{Z_3(Z_2+Z_{2K})}{Z_1+Z_3+Z_{2K}}$$

Solving and arranging in quadratic form Z_{2K}, we get

$$Z_{2K}=\frac{(Z_2-Z_.)}{2}\pm\tfrac{1}{2}\sqrt{[(Z_2-Z_1)^2+4(Z_1Z_2+Z_2Z_3+Z_3Z_1)]}. \quad ...(14)$$

Equations (13) and (14) for iterative impedances show that there are two values for each iterative impedance but practically there may be only one value.

The third property necessary to specify completely the network characteristic is iterative transfer constant p and is defined by the equation

$$\frac{I_2}{I_1}=e^{-p},$$

where I_2 and I_1 are the input and output currents of the network respectively, when operated under iterative conditions. Iterative transfer constant is the same for either direction of transmission.

For L section (fig. 7), the iterative transfer constant is given by the relation

$$e^p=\frac{I_1}{I_2}=\frac{Z_b+Z_{1K}}{Z_b}$$

$$=1+\frac{Z_{1K}}{Z_b}. \quad ...(15)$$

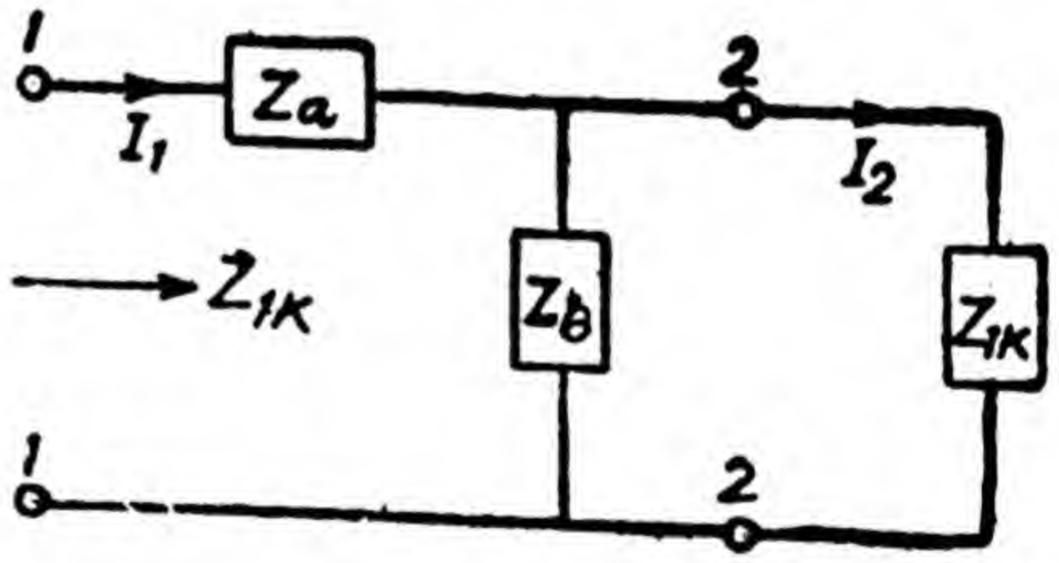


Fig. 7.

3·1. IMPEDANCE MATCHING :

Two impedances are said to be matched when they have the same magnitude and the same phase angle. For example, in fig. 8 load impedance Z_L is said to be matched to the generator impedance Z_g if both possess the same magnitude and same phase angle. However, if Z_L and Z_g are not matched then to secure matching, a network having $Z_{I_1}=Z_g$ and $Z_{I_2}=Z_L$ but with zero attenuation, is inserted between generator and

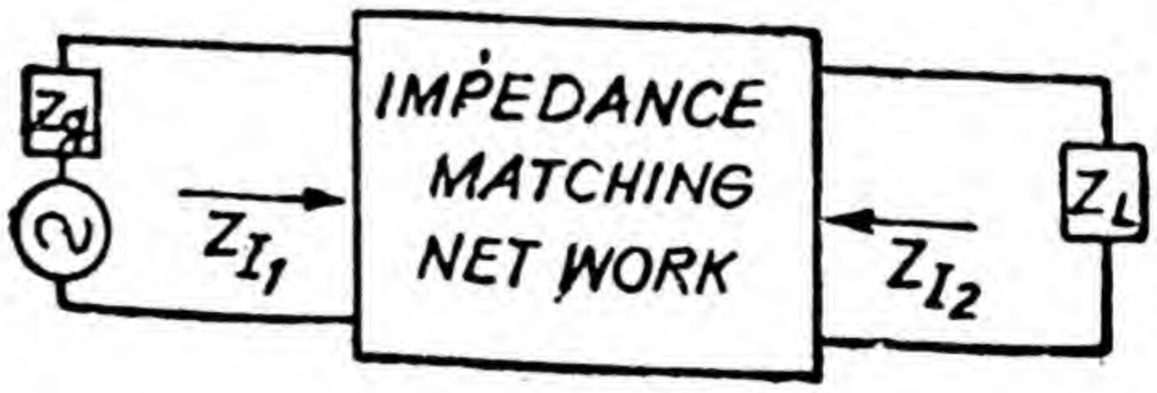


Fig. 8.

load. Due to this insertion, generator sees a load impedance $Z_{I_1} = Z_g$ and load takes power from a source of internal impedance $Z_{I_2} = Z_L$, *i.e.* matching is on image impedance basis. The ratio

$$\frac{\text{load current without matching}}{\text{load current with matching}}$$

is termed as mismatching factor, reflection or transition factor and is denoted by k. Mismatching results in the loss of load current. In decibels, it can be expressed as :

$$\text{mismatching in } db = 20 \log_{10} (1/k).$$

Matching, when the generator impedance is resistive, leads to the condition at which the power delivered by the generator to the load is maximum while mismatching to the loss of load power. Under certain conditions, when the generator impedance has reactive component, mismatching results in an increase in the load current.

3·2. INSERTION LOSS :

When a network is inserted between generator impedance and load impedance, loss in output current is resulted. Thus this loss is expressed as the ratio of current in the load impedance when the network is present to the load current in the absence of network. This is called insertion loss. We shall obtain an expression for this loss.

Let us consider a four terminal network connected between a

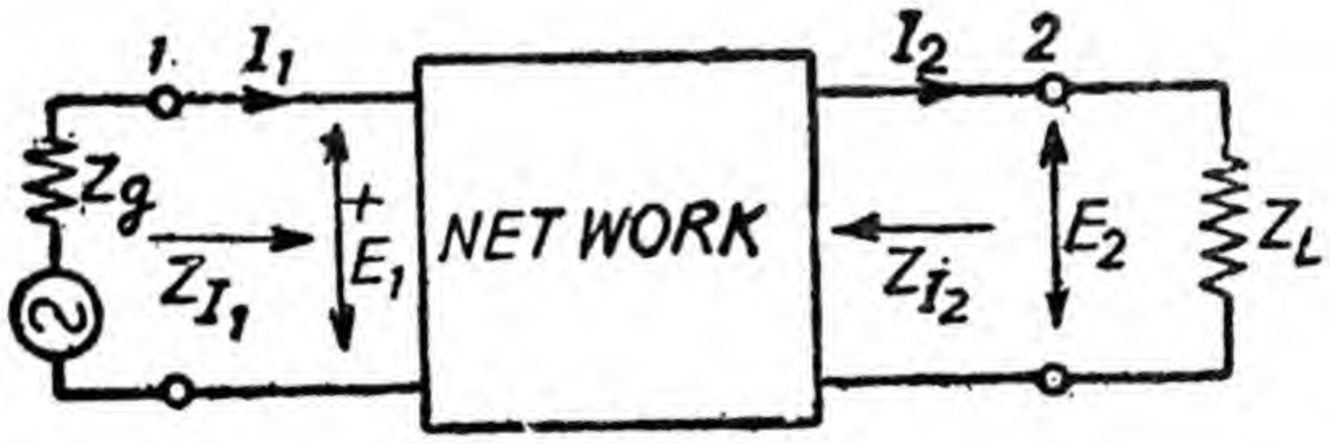


Fig. 9.

load impedance Z_L and a generator of voltage e and internal impedance Z_g connected at the input terminals.

Matrix expressing relation in output and input electrical quantities

$$\begin{Bmatrix} E_1 \\ I_1 \end{Bmatrix} = \begin{Bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Bmatrix} \times \begin{Bmatrix} E_2 \\ I_2 \end{Bmatrix}$$

Here $E_1 = e - I_1 Z_g$,

$E_2 = I_2 Z_L$.

Our aim is to calculate the insertion loss, *i.e.*, the ratio I_2/I_2', where I_2 is the current when network is present while I_2' is the current when network is absent. It is obvious that when network is absent, the total impedance of the circuit is $(Z_g + Z_L)$ and voltage is e, hence

$$I_2' = \frac{e}{Z_g + Z_L}.$$

Thus our problem is now to calculate I_2 in terms of image impedance Z_{I_1}, Z_{I_2}, image transfer constant γ, Z_g and Z_L.

From matrix,

$$I_1 = A_{21}E_2 + A_{22}I_2$$
$$= (A_{21}Z_L + A_{22})\, I_2,$$

after subtraction the value of E_2.

And
$$F_1 = A_{11}E_2 + A_{12}I_2 = e - I_1 Z_g$$

or
$$A_{11}I_2 Z_L + A_{12}I_2 = e - I_1 Z_g,$$
$$(A_{11}Z_L + A_{12})\, I_2 = e - I_1 Z_g.$$

Now putting the value of I_1,

$$(A_{11}Z_L + A_{12})\, I_2 = e - (A_{21}Z_L + A_{22})\, I_2 Z_g.$$

$$\therefore \quad \frac{e}{I_2} = A_{12} + A_{11}Z_L + A_{21}Z_L Z_g + A_{22}Z_g.$$

We shall substitute the values of 'A' parameters in terms of image parameters and image transfer constant γ.

The 'A' matrix is

$$\begin{Bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Bmatrix} = \begin{bmatrix} \cosh\gamma \sqrt{\left(\frac{Z_{I_1}}{Z_{I_2}}\right)}, & \sinh\gamma \sqrt{(Z_{I_1}.\, Z_{I_2})} \\ \frac{\sinh\gamma}{\sqrt{(Z_{I_1}\, Z_{I_2})}}, & \cosh\gamma \sqrt{\left(\frac{Z_{I_2}}{Z_{I_1}}\right)} \end{bmatrix}$$

Substituting the values of A_{11}, A_{12}, A_{21}, A_{22}, we get

$$\frac{e}{I_1} = \sinh\gamma \sqrt{(Z_{I_1}\, Z_{I_2})} + Z_L \cosh\gamma \sqrt{\left(\frac{Z_{I_1}}{Z_{I_2}}\right)} + Z_g \cosh\gamma \sqrt{\left(\frac{Z_{I_2}}{Z_{I_1}}\right)} + Z_L Z_g \frac{\sinh\gamma}{\sqrt{(Z_{I_1}\, Z_{I_2})}}$$

or
$$\frac{e}{I_g}\sqrt{(Z_{I_1}\, Z_{I_2})} = \sinh\gamma [Z_{I_1}Z_{I_2} + Z_L Z_g] + \cosh\gamma\, [Z_{I_1}Z_L + Z_g Z_{I_2}].$$

Putting $\sinh\gamma = \frac{e^{\gamma} - e^{-\gamma}}{2}$, $\cosh\gamma = \frac{e^{\gamma} + e^{-\gamma}}{2}$,

$$\frac{2e\sqrt{(Z_{I_1}\, Z_{I_2})}}{I_2} = (e^{\gamma} - e^{-\gamma})\, [Z_{I_1}\, Z_{I_2} + Z_L Z_g] + (e^{\gamma} + e^{-\gamma})\, [Z_{I_1}\, Z_L + Z_g Z_{I_2}]$$
$$= e^{\gamma}\, [Z_{I_1}\, Z_{I_2} + Z_L Z_g + Z_{I_1}\, Z_L + Z_g Z_{I_2}]$$
$$- e^{-\gamma}\, [Z_{I_1}\, Z_{I_2} + Z_L Z_g - Z_{I_1}\, Z_L - Z_g Z_{I_1}]$$
$$= e^{\gamma}[Z_{I_1} + Z_g]\, [Z_{I_2} + Z_L] - e^{-\gamma}\, (Z_{I_1} - Z_g]\, [Z_{I_2} - Z_L]$$
$$= e^{\gamma}.[Z_{I_1} + Z_g]\, [Z_{I_1} + Z_L] \left\{ 1 - e^{2\gamma} \frac{Z_{I_1} - Z_g\, Z_{I_2} - Z_L}{Z_{I_1} + Z_g\, Z_{I_2} + Z_L} \right\}$$

Putting $\rho_1 = \frac{Z_{I_1} - Z_g}{Z_{I_1} + Z_g}$ and $\rho_2 = \frac{Z_{I_2} - Z_L}{Z_{I_2} + Z_L}$

$$\frac{2e\sqrt{(Z_{I_1}\, Z_{I_2})}}{I_2} = e^{\gamma}\, [Z_{I_1} + Z_g]\, [Z_{I_2} + Z_L]\, [1 - e^{-2\gamma}\rho_1\rho_2].$$

$$\therefore \quad I_2=\frac{e\sqrt{(4Z_{I_1}Z_{I_2})}}{e^{\gamma}[Z_{I_1}+Z_g][Z_{I_2}+Z_L][1-e^{-2\gamma}\rho_1\rho_2]}.$$

Multiplying numerator and denominator by $\sqrt{(4Z_gZ_L)}$ and (Z_0+Z_L) and putting in suitable form, we get

$$I_2=\frac{e}{Z_0+Z_L}\times\frac{Z_0+Z_L}{\sqrt{(4Z_0Z_L)}}\times\frac{\sqrt{(4Z_0Z_{I_1})}}{Z_{I_1}+Z_0}\times\frac{\sqrt{(4Z_LZ_{I_2})}}{(Z_g+Z_L)}\times e^{-\gamma}\frac{1}{(1-e^{-2\gamma}\rho_1\rho_2)}$$

Putting $k=\dfrac{\sqrt{(4Z_0Z_L)}}{Z_L+Z_g}$, $\quad k_1=\dfrac{\sqrt{(4Z_0Z_{I_1})}}{Z_{I_1}+Z_L}$.

$$k_2=\frac{\sqrt{(4Z_{I_2}Z_L)}}{Z_{I_2}+Z_L} \text{ and } \sigma=\frac{1}{(1-e^{2\gamma}\rho_1\rho_2)}$$

$$\therefore \quad I_2=\frac{e}{Z_g+Z_L}\times\frac{k_1k_2}{k}\times e^{-\gamma}\sigma$$

or
$$I_2=I_2'\,\frac{k_1k_2}{k}.e^{-\gamma}.\sigma$$

$$\therefore \quad \frac{I_2}{I_2'}=\frac{k_1k_2}{k}.e^{-\gamma}.\sigma.$$

Let $\gamma=\alpha+j\beta$*,

therefore $\left|\dfrac{I_2}{I_2'}\right|=\left|\dfrac{k_1}{k}\right|\times\left|k_2\right|.e^{-\gamma}\,|\,\sigma\,|.$ ...(1)

Insertion loss in decibels $=20\log_{10}\left|\dfrac{I_2'}{I_2}\right|$

$$=-20\log_{10}\left|\frac{1}{k}\right|+20\log_{10}\left|\frac{1}{k_1}\right|+20\log_{10}\left|\frac{1}{k_2}\right|+8{\cdot}686\alpha+20\log_{10}\left|\frac{1}{\sigma}\right|$$

$$=-20\log_{10}\left|\frac{Z_L+Z_0}{\sqrt{(4Z_0Z_L)}}\right|+20\log_{10}\left|\frac{Z_{I_1}+Z_0}{\sqrt{(4Z_{I_1}Z_0)}}\right|$$

$$+20\log_{10}\left|\frac{Z_{I_2}+Z_L}{\sqrt{(4Z_{I_2}Z_L)}}\right|+\text{attenuation loss in decibels}$$

$$+20\log_{10}\left|1-e^{-\gamma}.\frac{Z_0-Z_{I_1}}{Z_0+Z_{I_1}}.\frac{Z_L-Z_{I_0}}{Z_L+Z_{I_2}}\right| \qquad ...(2)$$

In equation (1),
k is called external reflection factor between Z_g and Z_L,
k_1 is input reflection factor between Z_{I_1} and Z_L,
k_2 is output reflection factor between Z_{I_2} and Z_L,
ρ_1 is the reflection coefficient at input,
ρ_2 is the reflection coefficient at output,
σ is called the interaction factor.
Both these equations (1) and (2) represent the insertion loss.

*$\gamma=\alpha+j\beta$ where α is called attenuation constant and β phase constant.

Example. *Design a network to match a* 500 *ohms generator to a* 200 *ohms load with the minimum loss. Calculate the loss.*

Since R_1 must be greater than R_2, take

$$R_1 = 500 \text{ ohms},$$
$$R_2 = 200 \text{ ohms}.$$

Therefore,

$$Z_a = \sqrt{\{R_1 (R_1 - R_2)\}} = \sqrt{(500 \times 300)} = 387 \text{ ohms},$$

$$Z_b = \frac{R_1 R_2}{Z_a} = \frac{500 \times 300}{387} = 258 \text{ ohms}.$$

$$\frac{R_1}{R_2} = \frac{500}{200} = 2{\cdot}5$$

Now $$[\text{Loss}]_{min} = \frac{2R_1}{R_2} - 1 + 2\sqrt{\left\{\frac{R_1}{R_2}\left(\frac{R_1}{R_2} - 1\right)\right\}}$$

$$= 2 \times 2{\cdot}5 - 1 + \sqrt{\{2{\cdot}5\ (2{\cdot}5 - 1)\}}$$
$$= 5 - 1 + 2\sqrt{(3{\cdot}75)}$$
$$= 4 + 2.1{\cdot}93 = 4 + 3{\cdot}66$$
$$= \mathbf{7{\cdot}86}.$$

Loss in decibels

$$= 10 \log_{10} [\text{Loss}]_{min}$$
$$= 10 \log_{10} 7{\cdot}86$$
$$= 10 \times {\cdot}8954 = \mathbf{8{\cdot}954}.$$

Network is shown in figure (10).

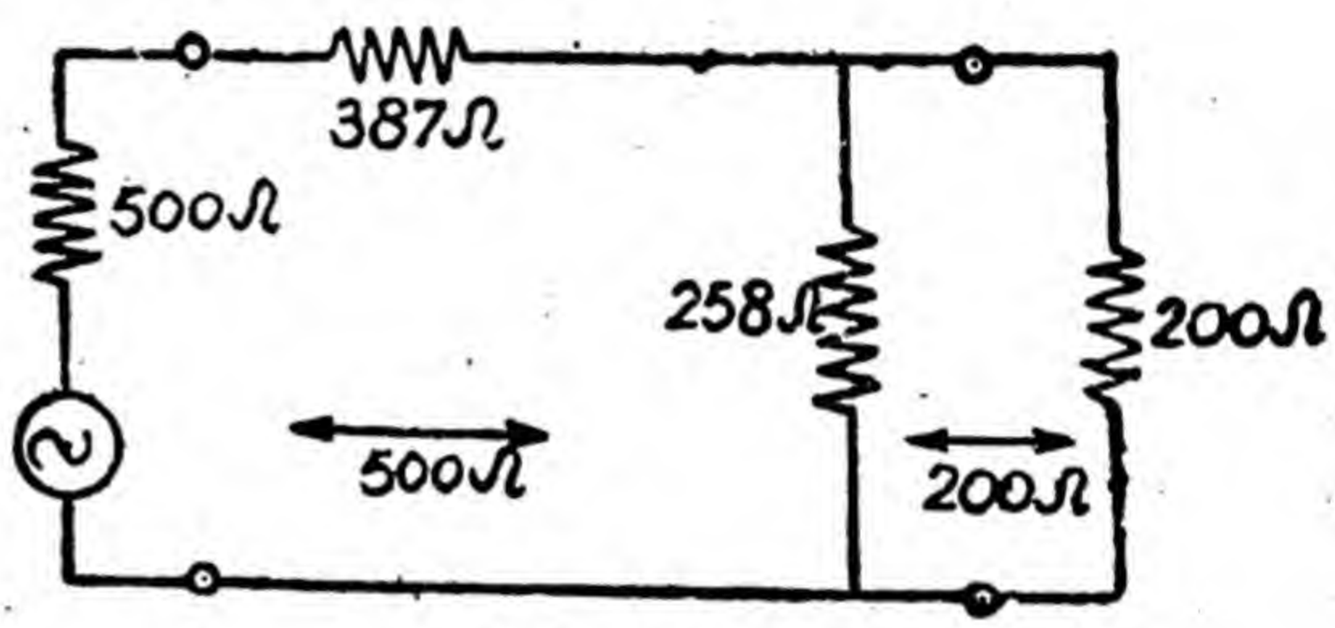


Fig. 10.

3·3. SUPERPOSITION THEOREM :

In any linear network containing impedances and energy sources, the current flowing in any element is the vec tor sum of the currents that are separately caused to flow in that element by each energy source.

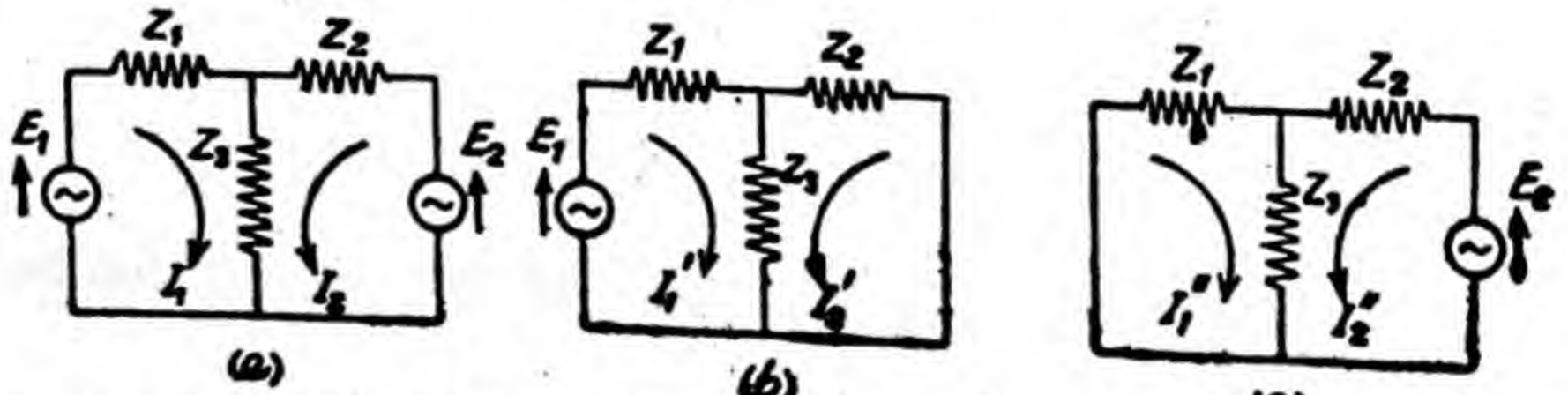


Fig. 11 (a) Two mesh network 11 (b) Two mesh network when E_2 is removed 11 (c) Two mesh network when E_1 is removed

Consider the two voltage sources E_1 and E_2 and three impedances Z_1, Z_2 and Z_3 as shown in fig. 11. Let I_1 and I_2 be the two mesh currents then the mesh equations can be written as

$$E_1 = (Z_1 + Z_3)\ I_1 + Z_3\ I_2 \qquad \ldots(1)$$
$$E_2 = Z_3\ I_1 + (Z_2 + Z_3)\ I_2 \qquad \ldots(2)$$

From equation (2), we have

$$E_2 - (Z_2 + Z_3)\ I_2 = Z_3\ I_1 \quad \text{or} \quad I_1 = \frac{E_2}{Z_3} - \frac{(Z_2 + Z_3)\ I_2}{Z_3}$$

Substituting the value of I_1 in equation (1), we have

$$E_1=(Z_1+Z_3)\left[\frac{E_2}{Z_3}-\frac{(Z_2+Z_3)}{Z_3}I_2\right]+Z_3I_2$$

$$=\frac{(Z_1+Z_3)\,E_2}{Z_3}-\frac{(Z_1+Z_3)(Z_2+Z_3)\,I_2}{Z_3}+Z_3\,I_2$$

$$=\frac{(Z_1+Z_3)\,E_2}{Z_3}-\frac{(Z_1Z_2+Z_2Z_3+Z_1Z_3)\,I_2}{Z_3}$$

$$\therefore\quad I_2=-\frac{E_1Z_3}{Z_1Z_2+Z_2Z_3+Z_1Z_3}+\frac{(Z_1+Z_3)\,E_2}{Z_1Z_2+Z_2Z_3+Z_1Z_3} \qquad \ldots(a)$$

Substituting the value of I_2 in equation (1), and solving it we get the value of I_1 which is given as

$$I_1=\frac{(Z_2+Z_3)\,E_1}{Z_1Z_2+Z_2Z_3+Z_1Z_3}-\frac{Z_3E_2}{Z_1Z_2+Z_2Z_3+Z_1Z_3} \qquad \ldots(b)$$

Now considering the circuit of fig. 11 (b), we have the mesh equotions

$$E_1=(Z_1+Z_3)\,I_1'+Z_3I_2' \qquad \ldots(3)$$

$$0=Z_3\,I_1'+(Z_2+Z_3)\,I_2'. \qquad \ldots(4)$$

Solving equations (3) and (4) we have

$$I_2'=\frac{-Z_3\,E_1}{(Z_1Z_2+Z_2Z_3+Z_1Z_3)} \qquad \ldots(c)$$

$$I_1'=\frac{(Z_2+Z_3)\,E_1}{(Z_1Z_2+Z_2Z_3+Z_1Z_3)}. \qquad \ldots(d)$$

Referring to circuit of fig. 11 (c), the mesh equations are

$$0=(Z_1+Z_3)\,I_1''+Z_3\,I_2'' \qquad \ldots(5)$$

$$E_2=Z_3I_1''+(Z_2+Z_3)\,I_2'' \qquad \ldots(6)$$

Solving these equations, we have

$$I_2''=\frac{(Z_1+Z_3)\,E_2}{Z_1Z_2+Z_2Z_3+Z_1Z_3}. \qquad \ldots(e)$$

$$I_1''=\frac{-Z_3E_2}{Z_1Z_2+Z_2Z_3+Z_1Z_3} \qquad \ldots(f)$$

From equations (a), (b), (c), (d), (e), and (f), we have

$$I_1=I_1'+I_1''$$

and

$$I_2=I_2'+I_2''$$

This proves the theorem. Now we shall prove this theorem in case of a number of meshes.

General Case :

In any linear network containing linear impedances and several sources, the voltage across or the current through any impedance may be calculated by adding algebraically all the individual voltages or currents caused by each source acting alone. with all other voltage sources replaced by short circuits and all other current sources replaced by open circuits.

By the term, '*linear network*' we mean that current in all branches is directly proportional to the driving voltage or e. m f. impressed.

Consider a linear network containing p-meshes with voltage sources E_{11}, E_{22}, E_{pp} and impedances Z_{11}, $Z_{22}...Z_{pp}$ in Ist, II nd,... pth mesh respectively. The voltage law equations can be written as :

for first mesh :

$$Z_{11}I_1+Z_{12}I_2+Z_{13}I_3+\ldots+R_{1m}I_m+\ldots+Z_{1p}I_p=E_{11},$$

where Z_{11} is the self impedance for first mesh and Z_{12} is the mutual impedance of meshes first and second. Similarly, the voltage law equations for rest meshes can be written as

$$Z_{21}I_1+Z_{22}I_2+Z_{23}I_3+\ldots+Z_{2m}I_m+\ldots+Z_{2p}I_p=E_{22}$$
$$Z_{31}I_1+Z_{32}I_2+Z_{33}I_3+\ldots+Z_{3m}I_m+\ldots+Z_{3p}I_p=E_{33}$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\ldots \quad \ldots\ldots \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad \ldots\ldots$$
$$Z_{p1}I_1+Z_{p2}I_2+Z_{p3}I_3+\ldots+Z_{pm}I_m+\ldots+Z_{pp}I_p=E_{pp}$$

The above voltage law equations can be easily solved with the help of famous Cramer's rule. Let I_m be the loop current flowing in mesh m, then I_m can be calculated by Cramer's rule as follows :

$$I_m=\frac{\begin{vmatrix} Z_{11} & Z_{12}\ldots & E_{11} & Z_{1p} \\ Z_{21} & Z_{22}\ldots & E_{22} & Z_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ Z_{p1} & Z_{p2}\ldots & E_{pp}\ldots & Z_{pp} \end{vmatrix}}{\begin{vmatrix} Z_{11} & Z_{12}\ldots & Z_{1m}\ldots & Z_{1p} \\ Z_{21} & Z_{22}\ldots & Z_{2m}\ldots & Z_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ Z_{p1} & Z_{p2}\ldots & Z_{pm} & Z_{pp} \end{vmatrix}}=\frac{C_{1m}E_{11}+C_{2m}E_{22}+\ldots+C_{pm}E_{pp}}{\Delta}$$

where, $C_{mn}=(-1)^{m+n}M^{mn}$

M_{mn} is the determinant which is left by removing m^{th} and n^{th} column from Δ and Δ is the pure impedance determinant

$$\Delta=\begin{vmatrix} Z_{11} & Z_{22}\ldots & Z_{1m}+Z_{1p} \\ Z_{21} & Z_{22}\ldots & Z_{2m}\ldots Z_{2p} \\ \ldots & \ldots & \ldots \\ Z_{p1} & Z_{p2}\ldots & Z_{pm}\ldots Z_{pp} \end{vmatrix}$$

So $$I_m=\frac{1}{\Delta}\left[C_{1m}E_{11}+C_{2m}E_{22}+\ldots+C_{pm}E_{pp}\right].$$

or $$I_m=aE_{11}+bE_{22}+cE_{33}+\ldots\ldots,$$

where $a=\frac{C_{1m}}{\Delta}$, $b=\frac{C_{2m}}{\Delta}$ and so on.

Thus I_m is the total loop current flowing in m^{th} mesh due to individual currents and hence the superposition theorem has been proved.

Note : In short circuiting the voltage sources, their internal impedances should be taken into account.

Advantages :

(1) This theorem permits the solution of networks without setting up a large number of simultaneous equations because at a time one generator is used.

(2) If the new generators are introduced into the system, it is not necessary to solve the network from begining provided the internal impedances of the generators are zero.

(3) If voltage of different frequencies are introduced, this theorem permits a solution to be obtained for each individual frequency. As these solutions are independent of each other, the currents of each frequency flow as if the other frequencies were absent.

Problem 1 : Use the superposition theorem to find the current i_x in branch cd in the following circuit.

According to the superposition theorem first we shall calculate the current i_x due to voltage source $10V$ by short circuiting the $5V$ source, then branches *cd* and *ce* will be parallel to each other and in series with *ac* branch; so total resistance is given as

$$R_{T_1}=2+\frac{1\times 3}{1+3}$$

$$=2+\frac{3}{4}=\frac{11}{4} \text{ ohms.}$$

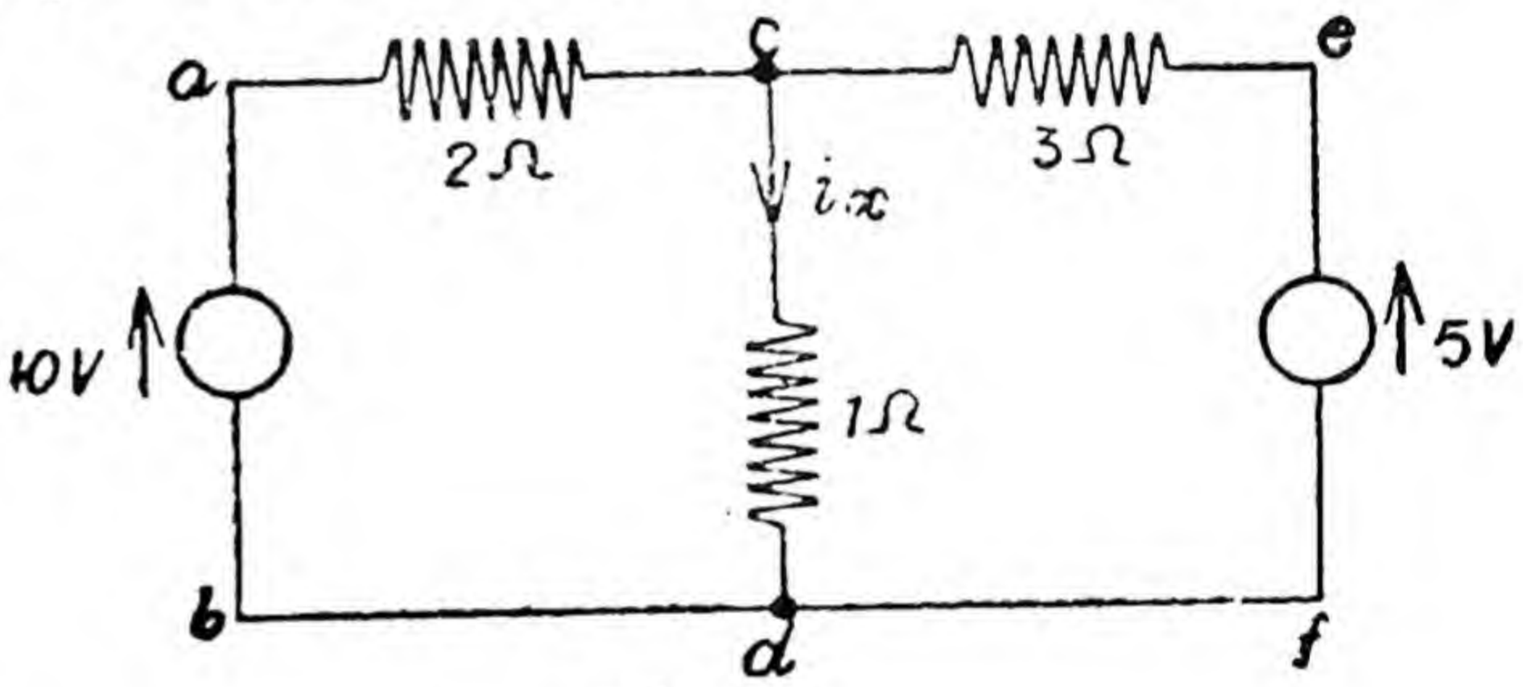


Fig. 12.

So the current generated by $10V$ source is

$$I_1=\frac{10}{11/4}=\frac{40}{11} \text{ amp.}$$

Henee $$ix_1=\frac{40}{11}\times\frac{3}{4}=\frac{30}{11} \text{ amp.}$$

Now we shall calculate the current due to $5V$ source by short circuiting $10V$ source.

We see that branches *cd* and *ac* are parallel with each other and are in series with *ce* branch.

So, $$R_{T_2}=3+\frac{2\times 1}{2+1}=3+\frac{2}{3}=\frac{11}{3}$$

So the current generated by $5V$ source is

$$I_2=\frac{5}{11/3}=\frac{15}{11}$$

Hence $$ix_2=\frac{15}{11}\times\frac{2}{3}=\frac{10}{11}$$

So $$i_x=ix_1+ix_2=\frac{30}{11}+\frac{10}{11}=\frac{40}{11}=3{\cdot}64 \text{ amp.}$$

Problem 2. Use the superposition theorem to find the current i_x in branch ce in the following circuit.

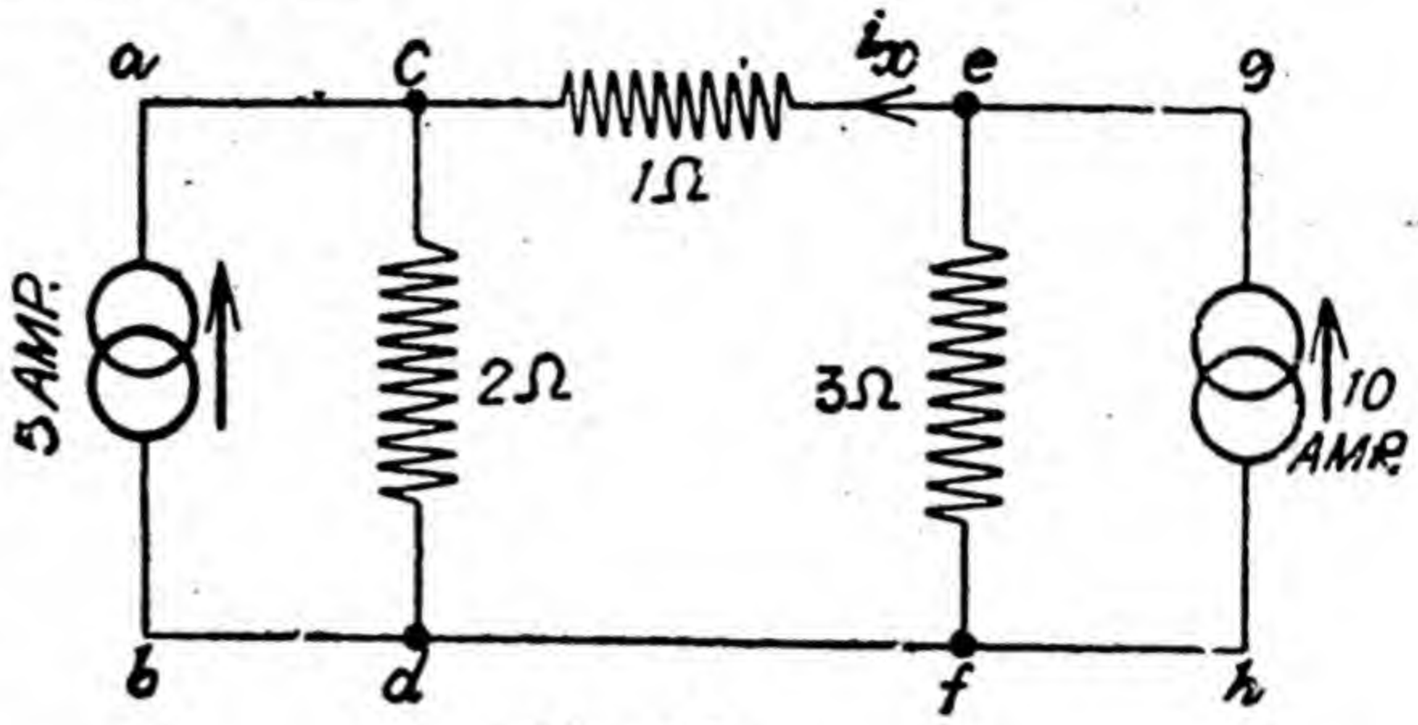


Fig. 13.

The same procedure will be adopted here for calculating i_x as we have applied in first problem with a difference that here we open circuit the current sources.

First open circuit 10 amp. source, then branches *ce* and *ef* will be in series with each other and in parallel with branch *cd*. So current 5 amp. will be divided at node *c* in *cd* and at *e* in *ef* branches. Applying the current division law at *c*, we have

$$i_{oe} = -i_{x_1} = \frac{5 \times 2}{2+1+3} = \frac{10}{6}$$

$$\therefore \quad i_{x_1} = -\frac{10}{6}$$

Now open circuit 5 amp. source and calculate current i_{x_2} due to 10 amp. source. We apply the current division law at *e*, we have

$$i_{x_2} = \frac{10 \times 3}{2+1+3} = \frac{30}{6}.$$

So $$i_x = i_{x_1} + i_{x_2} = -\frac{10}{6} + \frac{30}{6} = 3{\cdot}33 \text{ amp.}$$

Problem 3. Calculate the current in branch bc for the network shown in figure 14.

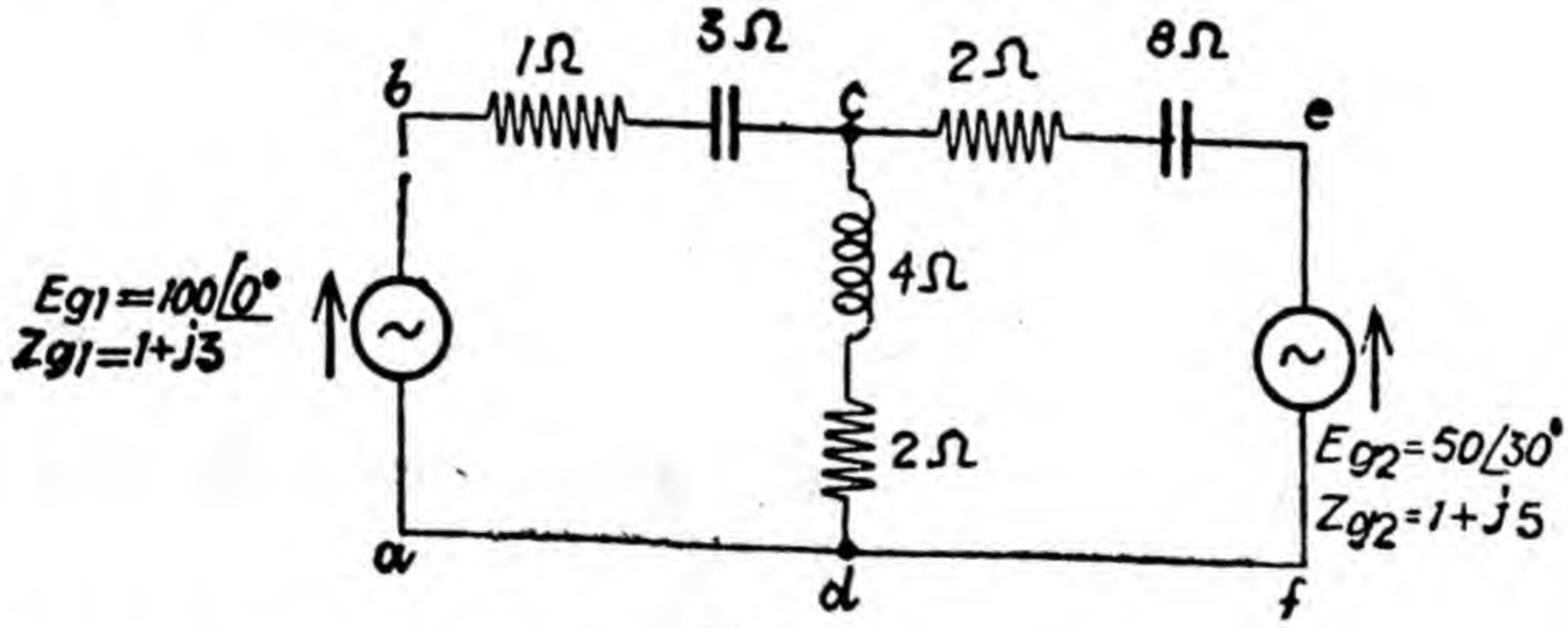


Fig. 14

This is more typical example in comparison with the above two examples but, however, the method of calculating the current is same. Assume $E_{g1}=0$, *i.e.* short circuit the branch *ab* but the internal impedance $Z_{g1}=1+j$ should be taken into account. So branches *cb* and *ba* will be in series with each other and parallel with *cd* branch, hence,

$$Z_{ca} = \frac{Z_{cba}\, Z_{cd}}{Z_{cba} + Z_{cb}} = \frac{(1-j3+1+j3)\,(2+j4)}{2+2+j4}$$

$$= 1{\cdot}5 + j0{\cdot}5 \text{ ohms.}$$

$$Z_{ea} = Z_{foo} + Z_{oa} = 1 + j5 + 2 - j8 + 1{\cdot}5 + j0{\cdot}5$$

$$= 4{\cdot}5 - j2{\cdot}5 \text{ ohms.}$$

So $$I_{eo} = \frac{50 \angle 0^\circ}{4{\cdot}5 - j2{\cdot}5} = \frac{50\ (\cos 30 + j \sin 30)}{4{\cdot}5 - j2{\cdot}5}$$

$$= 5 + j8{\cdot}34 \text{ amperes.}$$

So current in branch *cb* due to E_{g2} is given as

$$I_{cb_1} = \frac{I_{eo} Z_{ca}}{Z_{cba}} = \frac{(5 + j8{\cdot}34)\ (1{\cdot}5 + j0{\cdot}5)}{1 + j3 + 1 - j3}$$

$$= \frac{(5 + j8{\cdot}34)\ (1{\cdot}5 + j0{\cdot}5)}{2}$$

$$I_{cb_1} = 1{\cdot}66 + j7{\cdot}50 \text{ amperes.}$$

Now assume $E_{g2} = 0$

$$Z_{cf} = \frac{Z_{cef}\ Z_{cd}}{Z_{cef} + Z_{cd}} = \frac{(3 - j3)\ (2 + j4)}{(3 - j3 + 2 + j4)} = 3{\cdot}69 + j0{\cdot}462 \text{ ohms.}$$

$$Z_{bf} = Z_{abc} + Z_{ef} = 1 + j3 + 1 - j3 + 3{\cdot}69 + j0{\cdot}462$$

$$= 5{\cdot}69 + j0{\cdot}462 \text{ ohms.}$$

$$I_{bc_2} = \frac{100 \angle 0^\circ}{5{\cdot}69 + j0{\cdot}462} = \frac{100\ (\cos 0 + \sin 0)}{5{\cdot}69 + j0\ 462}$$

$$= \frac{100 + j0}{5{\cdot}69 + j0{\cdot}462} = 17{\cdot}43 - j1{\cdot}417 \text{ amps.}$$

So $$I_{bc} = I_{bc_1} + I_{bc_2} = -I_{bc_1} + I_{bc_2}$$

$$= -1{\cdot}76 - j7{\cdot}50 + 17{\cdot}43 - j1{\cdot}417$$

$$I_{bc}^* = 15{\cdot}77 - j8{\cdot}917 \text{ amps.}$$

3·4. RECIPROCITY THEOREM :

In any passive linear bilateral network containing bilateral linear impedances and sources of e.m.f. the ratio of a voltage E introduced in one mesh to the current I in any other mesh is the same as the ratio obtained if the positions of E and I are interchanged. other e.m. fs. being removed.

In other words, the interchange of an ideal current source and an ideal voltage source in passive linear bilateral circuit will not change the voltmeter and ammeter readings.

Or

In any linear network, if a source of e.m. f. E located in one mesh produces a current I_2 in second mesh, then if the same source E is placed in the second mesh and it produces a current I_1' in first mesh, we have $I_2 = I_1'$.

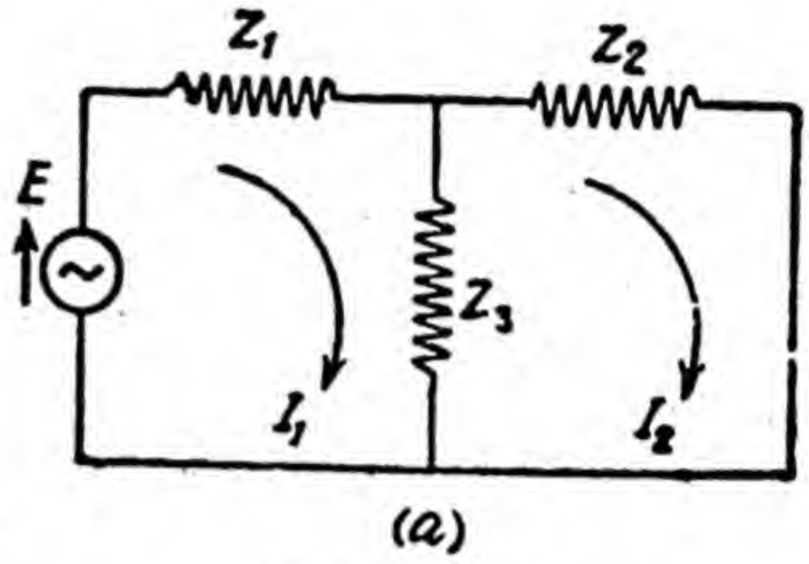
---

*If it is asked to write I_{bc} in terms of angle we can write as

$$I_{bc} = \sqrt{[(15{\cdot}77)^2 + 8\ 917)^2]}\ \tan \frac{-8{\cdot}917}{15{\cdot}77}$$

By the term *passive network* we mean, the networks which do not contain any internal e. m. fs. whereas network containing sources of e.m.fs. are called *active networks. Bilateral elements* are those circuit elements like inductance, resistance, and capacitance which transmit current equally well in either direction whereas *unilateral elements* are those circuit elements like rectifiers and vacuum tubes which transmit effectively only in one direction.

Consider the circuit shown in fig. 15. Let us calculate the value of $\frac{E}{I_2}$ with the help of Kirchoffs law.



(a)

Fig. 15.

Applying this law we have

$$E=(Z_1+Z_3)\,I_1-Z_3I_2 \qquad \ldots(1)$$

$$0=-Z_3I_1+(Z_2+Z_3)\,I_2. \qquad \ldots(2)$$

Solving the equations (1) and (2), we get

$$I_2=\frac{\begin{vmatrix} Z_1+Z_3 & E \\ -Z_3 & 0 \end{vmatrix}}{\begin{vmatrix} Z_1+Z_3 & -Z_3 \\ -Z_3 & Z_2+Z_3 \end{vmatrix}}$$

$$=\frac{Z_3E}{(Z_1+Z_3)(Z_2+Z_3)-Z_3^2}=\frac{Z_3E}{Z_1Z_2+Z_2Z_3+Z_1Z_3}$$

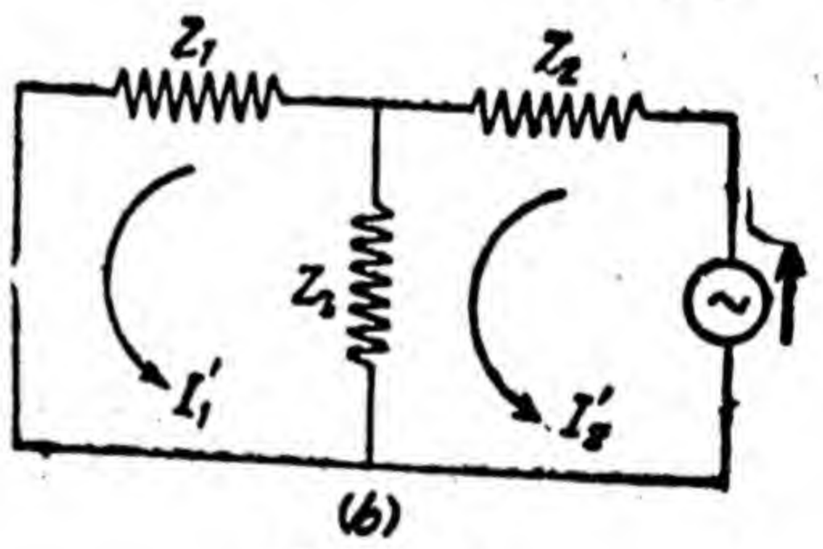
$$\therefore \quad \frac{E}{I_2}=\frac{Z_1Z_2+Z_2Z_3+Z_1Z_3}{Z_3}. \qquad \ldots(a)$$

Now let the position of the battery be changed and the circuit be as shown in figure 16. Applying Kirchoff's law, we have

$$E=(Z_2+Z_3)\,I_2'-Z_3I_1' \qquad \ldots(3)$$

$$0=-Z_3I_2'+(Z_1+Z_3)\,I_1' \qquad \ldots(4)$$

Solving equations (3) and (4), we get



(b)

Fig. 16.

$$I_1'=\frac{\begin{vmatrix} (Z_2+Z_3) & E \\ -Z_3 & 0 \end{vmatrix}}{\begin{vmatrix} (Z_2+Z_3) & -Z_3 \\ -Z_3 & (Z_1+Z_3) \end{vmatrix}}$$

$$=\frac{Z_3E}{(Z_2+Z_3)\,(Z_1+Z_3)-Z_3^2}=\frac{Z_3E}{Z_1Z_2+Z_2Z_3+Z_1Z_3}$$

$$\therefore \quad \frac{E}{I_1'}=\frac{Z_1Z_2+Z_2Z_3+Z_1Z_3}{Z_3}. \qquad \ldots(b)$$

As the right hand side of equations (*a*) and (*b*) are equal, hence

$$\frac{E}{I_2}=\frac{E}{I_1'}.$$

This proves the reciprocity theorem. Now we shall consider the general case.

General Case :

Consider a linear network containing p meshes with voltage sources $E_{11}, E_{22} \ldots E_{pp}$ and impedances $Z_{11}, Z_{22} \ldots Z_{pp}$ in 1st, IInd, pth mesh respectively. The voltage law can be written as :

For Ist mesh $\quad Z_{11}I_1+Z_{12}I_2+\ldots+Z_{1x}I_x+\ldots+Z_{1p}I_p=E_{11}$

For IInd mesh $\quad Z_{21}I_1+Z_{22}I_2+\ldots+Z_{2x}I_x+\ldots+Z_{2p}I_p=E_{22}$

$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$

For x mesh $\quad Z_{x1}I_1+Z_{x2}I_2+\ldots+Z_{xx}I_x+\ldots+Z_{xp}I_p=E_{xx}$

$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$

For p mesh $\quad Z_{p1}I_1+Z_{p2}I_2+\ldots+Z_{px}I_x+\ldots+Z_{pp}I_p=E_{pp}.$

The above set of equations can be solved with the help of Cramer's rule and the current I_x flowing in x mesh can be given as

$$I_x=\frac{C_{1x}}{\Delta}E_{11}+\frac{C_{2x}}{\Delta}E_{22}+\frac{C_{3x}}{\Delta}E_{33}+\ldots+\frac{C_{xx}E_{xx}}{\Delta}+\ldots+\frac{C_{px}}{\Delta}E_{pp}.$$

where C_{mn} and Δ are the same as stated in superposition theorem.

Now let all *emfs* except the *emf* E_{mm} in the nth mesh, are zero and due to this E_{mm} the current I_n flowing in n^{th} mesh is given by

$$I_n=\frac{C_{1n}}{\Delta}0+\frac{C_{2n}}{\Delta}0+\ldots+\frac{C_{mn}}{\Delta}E_{mm}+\ldots+\frac{C_{pn}}{\Delta}0.$$

$$\therefore \quad I_n=\frac{C_{mn}}{\Delta}E_{mm}. \qquad \ldots(5)$$

Now let all *emfs* except the *emf* E_{mm} in the nth mesh, are zero and due to this E_{nn} the current I_m flowing in m^{th} mesh is given by

$$I_m=\frac{C_{1m}}{\Delta}0+\frac{C_{2m}}{\Delta}0+\ldots+\frac{C_{nm}}{\Delta}E_{nn}+\ldots+\frac{C_{pm}}{\Delta}0$$

$$\therefore \quad I_m=\frac{C_{nm}}{\Delta}E_{nn}. \qquad \ldots(6)$$

It is obvious that sources are same

i.e. $\quad E_{mm}=E_{nn}=E, \qquad \ldots(7)$

$$C_{mn}=(-1)^{m+n}M_{mn}$$

and $\quad C_{nm}=(-1)^{n+m}M_{nm}.$

Clearly $\quad C_{mn}=C_{nm}$ since $M_{mn}=M_{nm}. \qquad \ldots(8)$

Now applying equations (7) and (8) to equations (5) and (6), we have $\quad I_m=I_n=I.$

Hence $\quad \frac{E_{mm}}{I_n}=\frac{E_{nn}}{I_m}=\frac{E}{I}.$

Problem 1. Given the network shown in fig. (17) prove that if 100 voits are inserted in bc and branch ef is left closed, the current flowing in ef will then be exactly the same as the current that flowed in bc when this same voltage was applied at ef.

Let us first calculate the current in branch *bc* due to $E_g = 100 \angle 0$ inserted in *ef* branch. We find

$$Z_{ao} = \frac{Z_{abc}\, Z_{ad}}{Z_{abc} + Z_{ad}} = \frac{(3+j4)\,(-j10)}{3+i4-j10} = 6{\cdot}67 + j3{\cdot}33 \text{ ohms.}$$

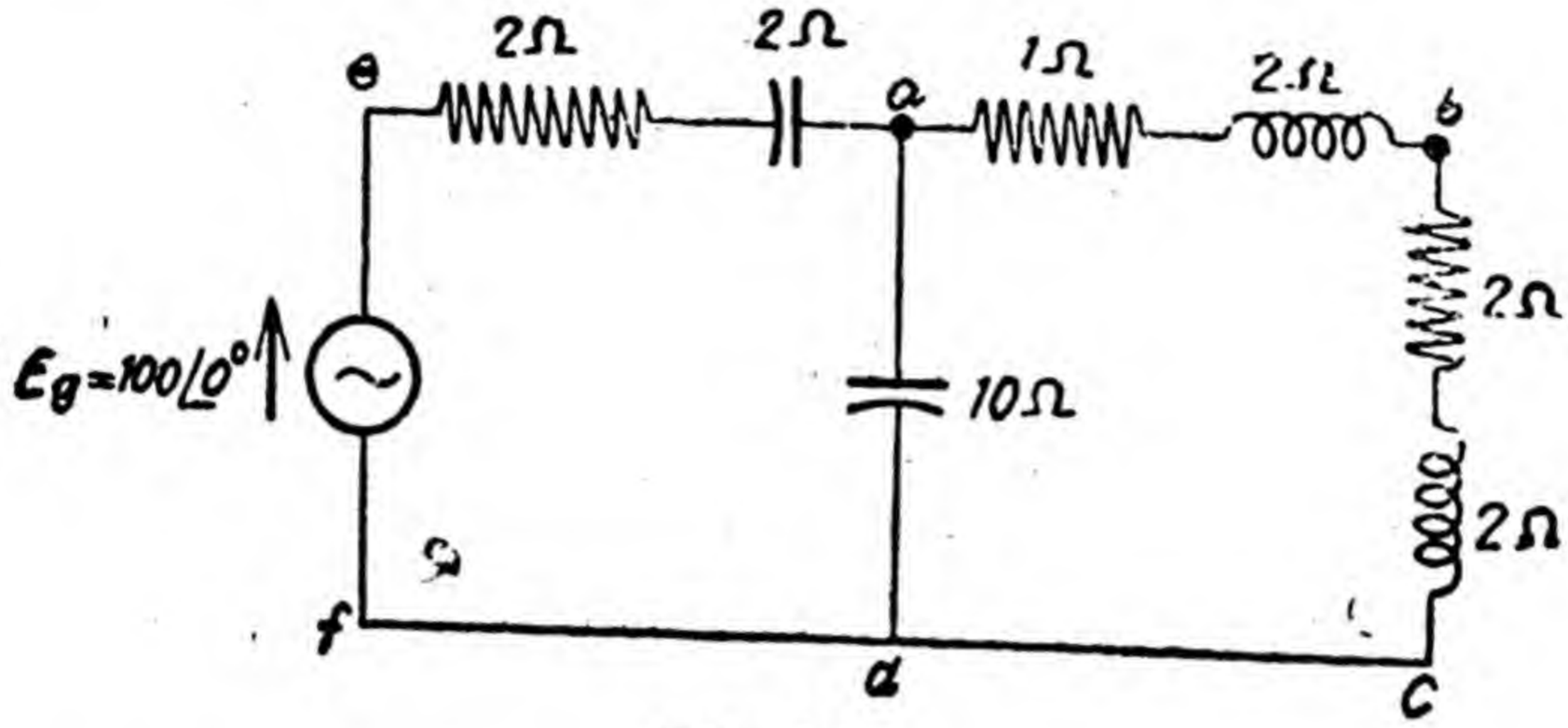


Fig. 17

$$Z_{ec} = Z_{ea} + Z_{ac} = 2 - j2 + 6{\cdot}67 + j3{\cdot}33$$
$$= 8{\cdot}67 + j\ 1{\cdot}33 \text{ ohms.}$$
$$I_{ec} = \frac{V_{fe}}{Z_{eo}} = \frac{100 + j0}{8{\cdot}67 + j1{\cdot}33} = 11{\cdot}27 - j1{\cdot}732 \text{ amps.,}$$
$$V_{ae} = I_{ec} Z_{ac} = (11{\cdot}27 - j1{\cdot}732)\,(6{\cdot}67 + j3{\cdot}33)$$
$$= 81 + j26 \text{ volts.}$$
$$\therefore \quad I_{cc} = \frac{V_{ae}}{Z_{abc}} = \frac{81 + j26}{3 + j4} = 13{\cdot}88 - j9{\cdot}84 \text{ amps.}$$

Now if 100 volts are inserted in branch *bc* and *ef* remains as short circuited. The current in *ef* will be calculated as follows :

$$Z_{af} = \frac{(2-j2)\,(-j10)}{(2-j2-j10)} = 1{\cdot}352 - j18{\cdot}92 \text{ ohms.}$$
$$Z_{caf} = 3 + j4 + 1{\cdot}352 + j1{\cdot}8692 = 4{\cdot}352 + j2{\cdot}108 \text{ ohms.}$$
$$I_{caf} = \frac{100 + j0}{4{\cdot}352 + j2{\cdot}108} = 18{\cdot}6 - j9{\cdot}02 \text{ amps.}$$
$$V_{ef} = (18{\cdot}6 - j9{\cdot}02)\,(1{\cdot}352 - j1{\cdot}892)$$
$$= 8{\cdot}07 - j47{\cdot}4 \text{ volts.}$$
$$\therefore \quad I_{ef} = \frac{8{\cdot}07 - j47{\cdot}4}{2 - j2} = 13{\cdot}88 - j9{\cdot}84 \text{ amps.}$$

Thus we see that I_{ef} is exactly same as I_{bc}.

Problem 2. In the same network as shown in problem 1, if 100 $\angle 0°$ volts are inserted in ef branch, find the current in branch ad. Verify the result by calculating the current in fe when 100 $\angle 0°$ volts are inserted in branch ad.

Let us first calculate the current in branch *ad* due to $E_g = 100 \angle 0°$ inserted in branch *ef*. We have

$$Z_{ae}=\frac{Z_{abe}\,Z_{ad}}{Z_{abc}+Z_{ad}}=\frac{(3+j4)\,(-j10)}{3+j4-j10}=6{\cdot}67+j3{\cdot}33 \text{ ohms.}$$

$$Z_{eo}=Z_{ea}+Z_{ao}=2-j2+6{\cdot}67+j3{\cdot}33$$
$$=8{\cdot}67+j1{\cdot}33 \text{ ohms.}$$

$$I_{eo}=\frac{V_{fe}}{Z_{ae}}=\frac{120+j0}{8{\cdot}67+j1{\cdot}33}=11{\cdot}27-j1{\cdot}732 \text{ amps.}$$

$$V_{ao}=I_{eo}Z_{ao}=(11{\cdot}27-j\ 1{\cdot}732)\ (6{\cdot}67+j\ 3{\cdot}33)$$
$$=81+j\ 26 \text{ volts.}$$

$$\therefore \qquad I_{ad}=\frac{V_{ao}}{Z_{ad}}=\frac{81+j\ 26}{-j\ 10}=-{\cdot}6+j8{\cdot}1 \text{ amps.}$$

Now insert 100 $\angle 0$ volts in branch *ad* and short circuit the branch *ef*.

In this case branches *abc* and *aef* are parallel with each other in series with *ad* branch. So that impedance can be given as

$$Z_T=Z_{ad}+\frac{Z_{abc}\times Z_{aef}}{Z_{abo}+Z_{aef}}$$
$$=\left[-j\ 10+\frac{(3+j\ 4)\ (2-j\ 2)}{3+j\ 4+2-j\ 2}\right]$$
$$=\left[-j\ 10+\frac{4+j\ 2}{5+j\ 2}\right]$$
$$=\frac{34-j\ 48}{5+j\ 2}.$$

So current

$$I_{ad}=\frac{V_{ac}}{Z_T}=\frac{100+j\ 0}{34-j\ 48}\times(5+j\ 2) \text{ amps.}$$

Applying the current division law at *a*, we get

$$I_{ef}=\frac{(100+j\ 0)}{34-j\ 48}\times\frac{5+j\ 2}{1}\times\frac{3+j\ 4}{5+j\ 2}$$
$$=\frac{(100+j\ 0)\ (3+j\ 4)}{34-j\ 48}$$
$$=\frac{50\ (3+j\ 4)\ (17+j\ 24)}{(17-j\ 24)\ (17+j\ 24)}$$
$$=\frac{50\ (3+j\ 4)\ (17+j\ 24)}{17^2+24^2}$$
$$=\frac{50\ (-45+j140)}{865}.$$

$$I_{ef}=-2{\cdot}6+j\ 8{\cdot}1 \text{ amps.}$$

Hence $I_{ef}=I_{ad}$.

3·5. THEVENIN'S THEOREM :

Any two terminal linear network containing linear impedances and generators can be replaced with an equivalent circuit consisting of a voltage source E' in series with an impedance Z'. The value of

E' is the open circuit voltage between the terminals of the network, and Z' is the impedance measured between the terminals with all other generators being removed, (but not their internal impedances).

Consider a two terminal network, containing one active mesh and other passive mesh. Let us also assume that load impedance Z_R is appearing between two terminals as shown in fig. (18*a*). Fig. (18*b*) represents the Thevenin's equivalent network.

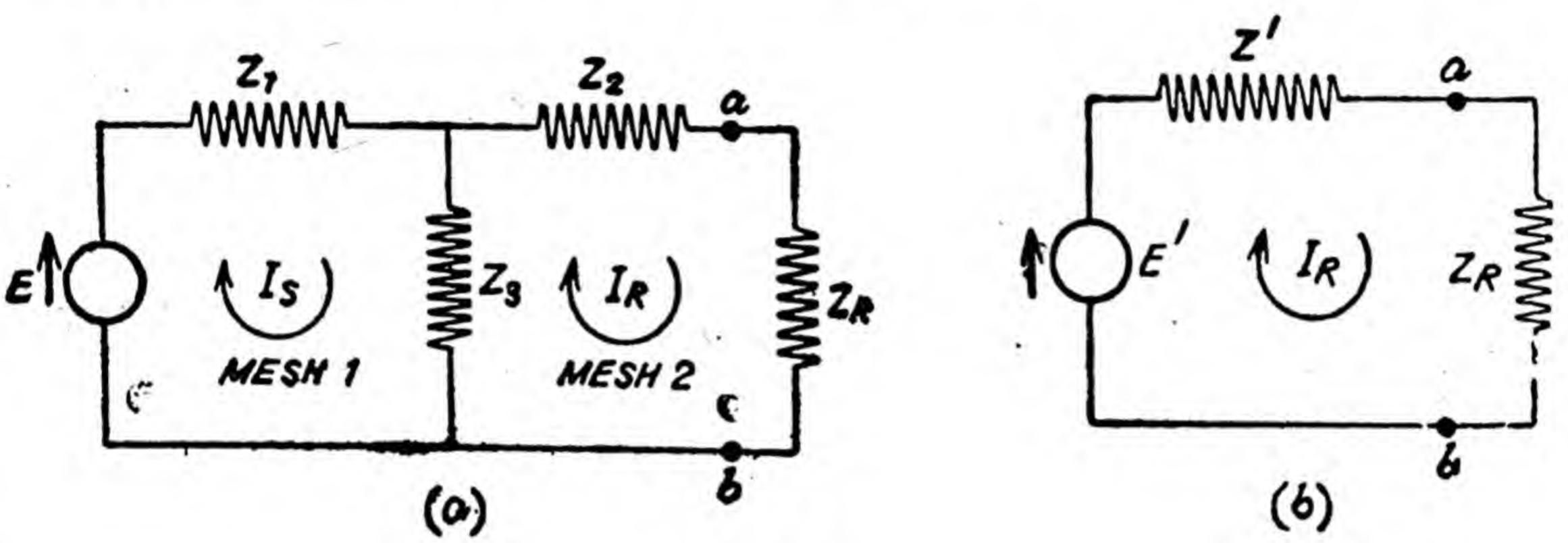


Fig. 18.

In fig. (18*a*) I_S and I_R represent the loop currents flowing in active and passive network respectively. Our aim is to calculate E' and Z' from fig. (18*a*) and then to show that it is equivalent to a circuit shown in fig. (18*b*).

Applying the voltage law equations in mesh 1 and mesh 2 respectively, we have

$$Z_1 I_S + Z_3 I_S - Z_3 I_R = E \qquad \ldots(1)$$

and $$(Z_2 + Z_3 + Z_R) I_R - Z_3 I_S = 0. \qquad \ldots(2)$$

The above two equations can be easily solved with the help of Cramer's rule. The current I_R flowing through load impedance Z_R is given as

$$I_R = \frac{\begin{vmatrix} Z_1+Z_3 & E \\ -Z_3 & 0 \end{vmatrix}}{\begin{vmatrix} Z_1+Z_3 & -Z_3 \\ -Z_3, & Z_2+Z_3+Z_R \end{vmatrix}} = \frac{EZ_3}{(Z_1+Z_3)(Z_2+Z_3+Z_R)-Z_3^2}$$

or $$I_R = \frac{\frac{EZ_3}{Z_1+Z_3}}{Z_2+Z_3+Z_R-\frac{Z_3^2}{Z_1+Z_3}} = \frac{\frac{EZ_3}{Z_1+Z_3}}{Z_2+Z_R+\frac{Z_1Z_3}{Z_1+Z_3}} \qquad \ldots(3)$$

This is the value of load current. By inspection of fig. (18*a*) the open circuit voltage at *a*, *b*, terminals is

$$E' = E\left(\frac{Z_3}{Z_1+Z_3}\right) \qquad \ldots(4)$$

and the impedance of the network measured between *a*, *b* terminals with all emf's generators short circuited is

$$Z' = Z_2 + \frac{Z_1 Z_3}{Z_1 + Z_3}. \qquad \ldots(5)$$

From equations (3), (4) and (5), it can be easily seen that

$$I_R = \frac{E'}{Z' + Z_R} \qquad \ldots(6)$$

Now from fig. (18*b*), we see that

$$I_R = \frac{E'}{Z' + Z_R}. \qquad \ldots(7)$$

Equations (6) and (7) are the same, hence theorem has been proved for networks containing one generator. It may be generalized to any number of generators by the application of superposition theorem permitting each generator and associated circuit to be considered separately.

E' and Z' as given by equations (4) and (5) are known as Thevenin's components.

Problem 1. Find the open circuit voltage and Thevenin resistance of the two terminal network shown in fig. (19).

Comparing fig. (19) with (18*a*) we have

$E = 100\ V$
$Z_1 = 20\Omega,\ Z_2 = 0,$
$Z_3 = 30\Omega.$

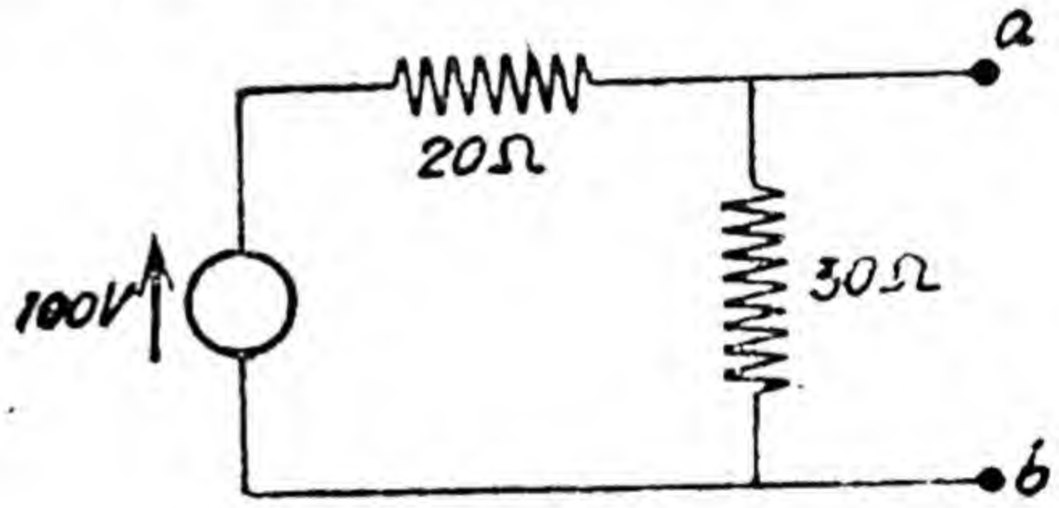


Fig. 19.

Then by Thevenin components we have open circuited voltage as

$$E' = \frac{EZ_3}{Z_1 + Z_3} \text{ or } E' = \frac{100 \times 30}{20 + 30} = 60V$$

and
$$Z' = Z_2 + \frac{Z_1 Z_3}{Z_1 + Z_3}$$

or
$$Z' = 0 + \frac{30 \times 20}{30 + 20} = \frac{30 \times 20}{50} = 12\Omega.$$

Hence $E' = 60V$ and $Z' = 12\Omega$.

Problem 2. Find the short circuit current and Thevenin resistance of the two terminal net work shown in fig. (20).

As given in question, short circuit *ab* by connecting terminals *ab* with resistanceless wire, then applying current division law, we have

$$I_{ab} = \frac{8}{6+2} \times 6 = \frac{8}{8} \times 6 = 6 \text{ amps.}$$

This is the short circuit current.

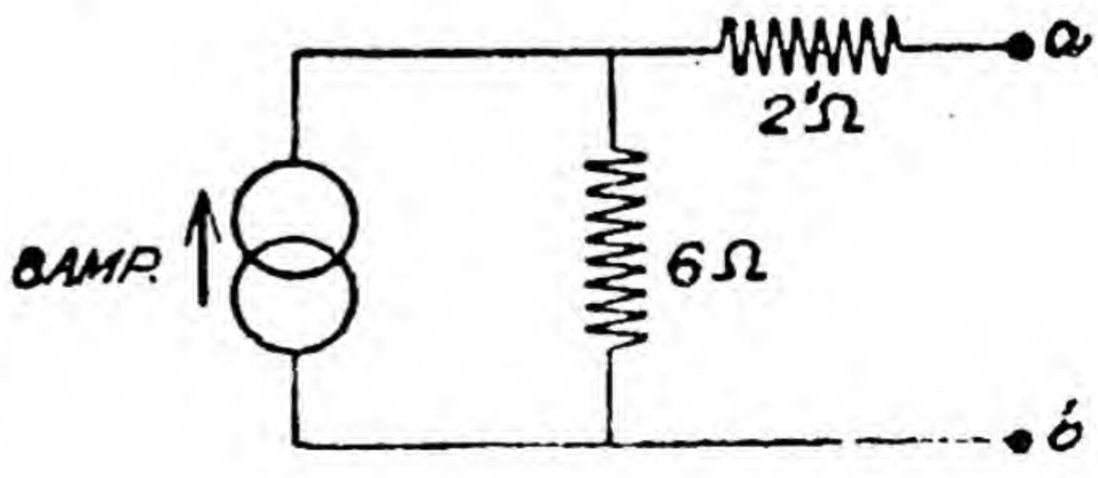


Fig. 20.

Now for finding out Z', make current source as open circuit then
$$Z'=2+6=8 \text{ ohms.}$$
Hence short circuit current is 6 amps.
and
$$Z'=8 \text{ ohms.}$$

Problem 3. Obtain Thevenin's equivalent circuit for the network shown in fig. 21a.

Sol. We shall first calculate the open-circuited voltage between A and B as follows.

The loop eq. for big closed mesh may be written as
$$10I+6I+2I$$
$$=3 \text{ volt}-3 \text{ volt}=0$$
or $I=0$

Hence (i) current across 1 Ω resistor will also be zero and (ii) voltage drop $(=r\times I)$ across 6 Ω resistor and along 1 Ω resistor will evidently be zero. Therefore the voltage across A and B will be same as across 12Ω resistor *i.e.*=3 volts.

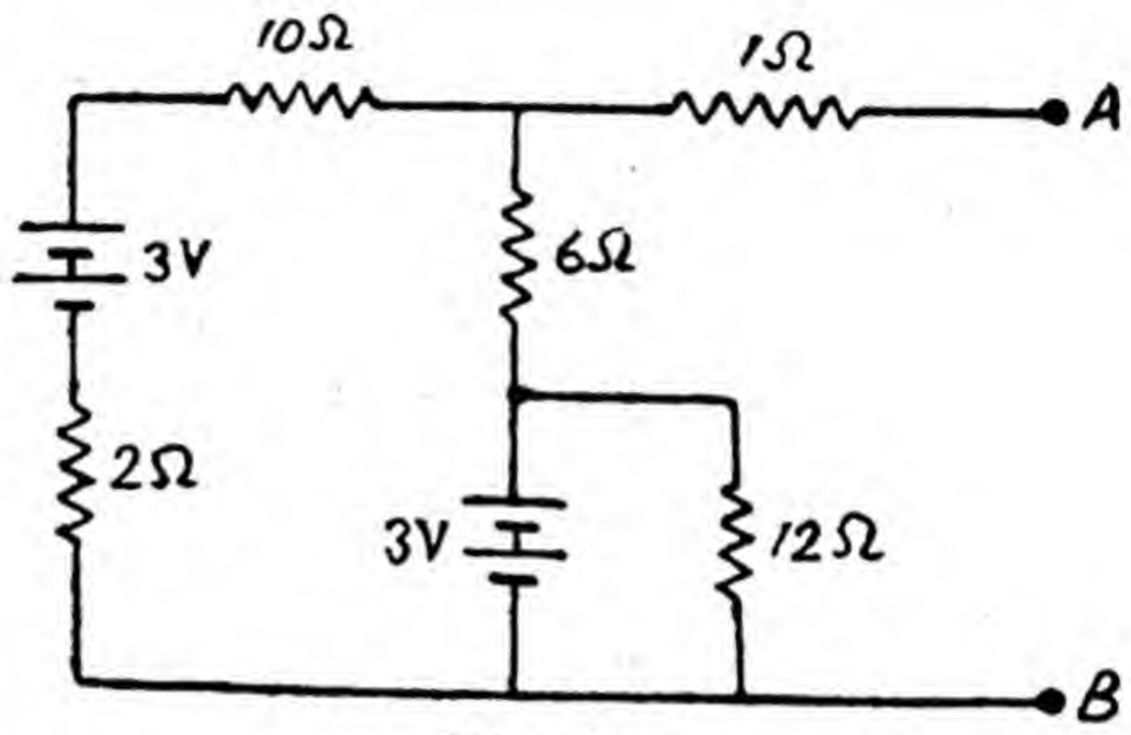


Fig. 21*a*

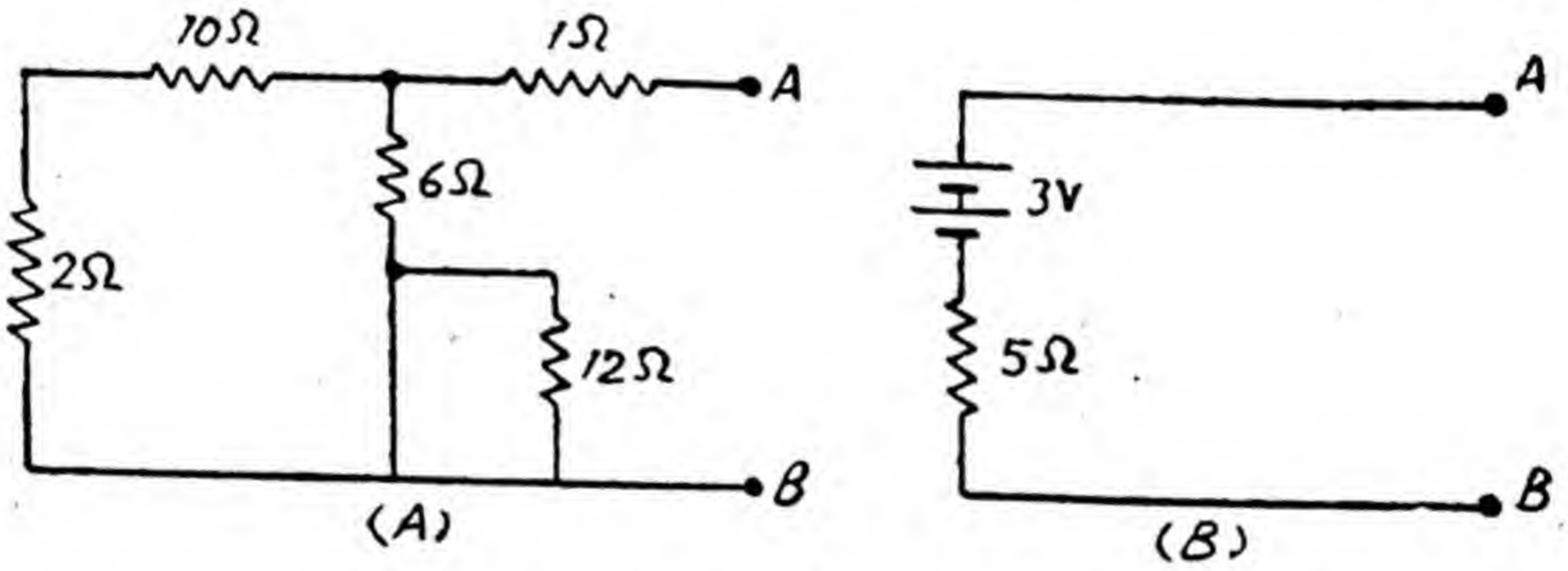


Fig. 21*b*

Now if each battery is replaced by zero resistance, as required for Thevenins equivalent circuit, then given circuit reduces to that shown in fig. 21*b* (A). In this, it is evident that 12Ω resistor becomes short circuited and so becomes ineffective. Therefore the equivalent resistance across A and B may be computed as follows : 10 and 2Ω resistor may be supposed as connected in series and 6Ω as in parallel with them. Hence effective resistance R' of these will be as given by

$$\frac{1}{R'}=\frac{1}{10+2}+\frac{1}{6} \quad \text{or} \quad R'=\frac{12}{3}=4$$

and further since 1 Ω resistor may be taken as connected in series with this, hence total resistance across A and $B=1+4=5$ **ohms.**

Therefore Thevenin's equivalent circuit for given network may be shown as in fig. 21*b* (B).

Problem 4. Using Thevenin's theorem, calculate the current in load R_L in the given circuit (fig. 22A).

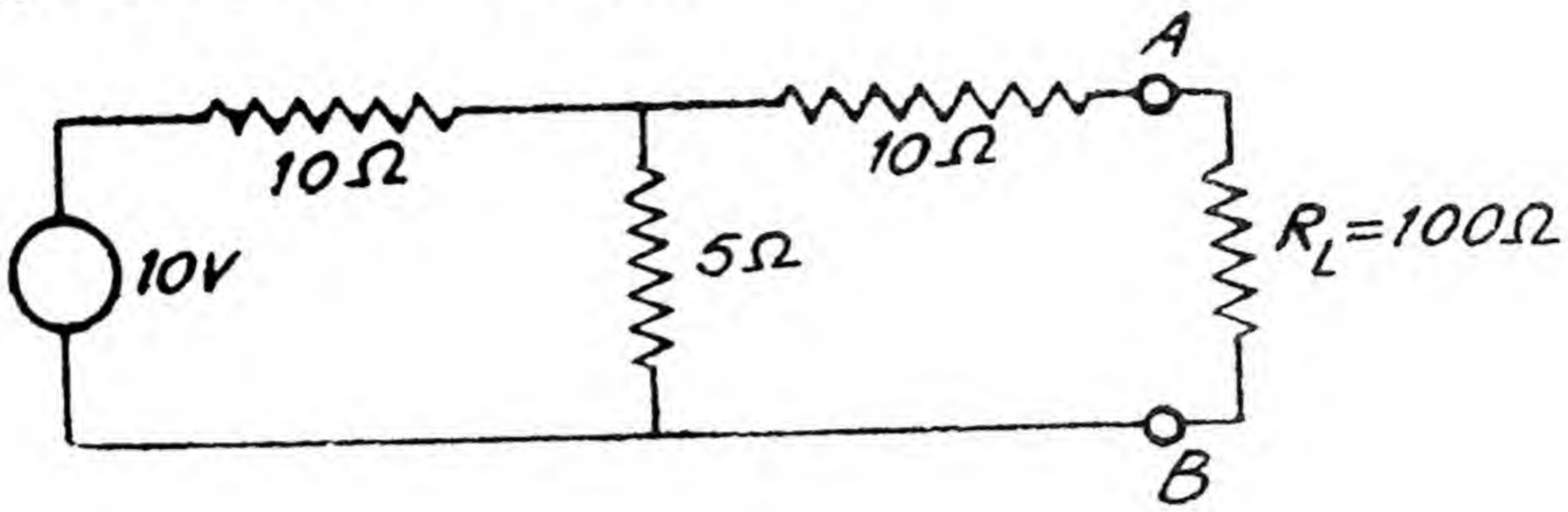


Fig. 22A

Sol. First of all, we shall calculate the open circuited voltage E' across A and B in the absence of R_L as shown in fig. 22B (a).

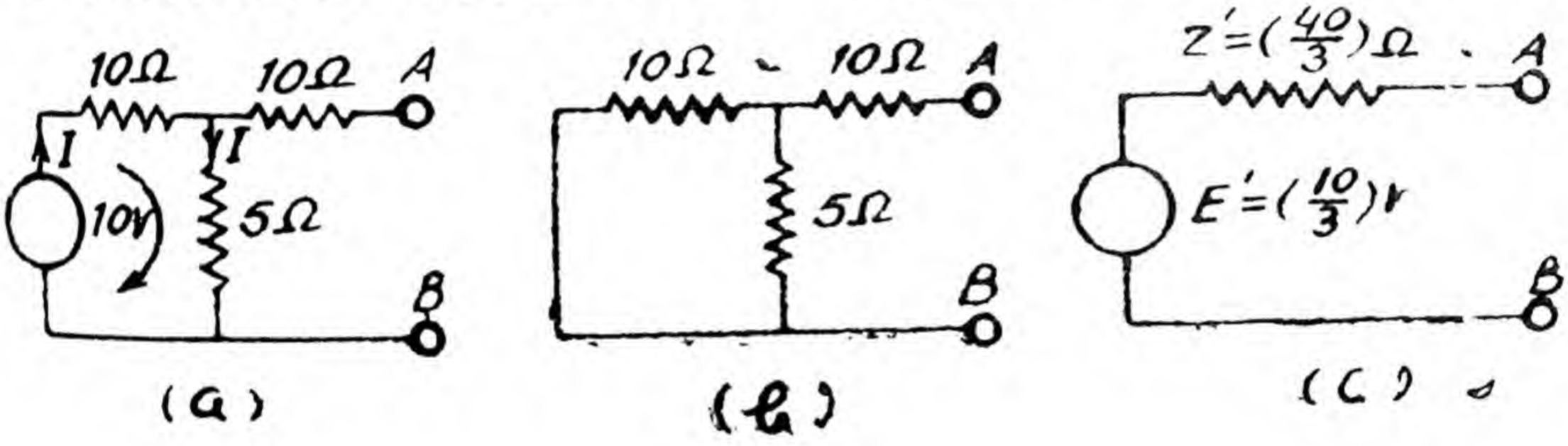


Fig. 22B

The loop equation is

$$10I+5I=10 \quad \text{or} \quad I=\frac{10}{15}=\frac{2}{3} \text{ amp.}$$

As no current flows through the resistor (10Ω) near point A, hence voltage across it is zero. The voltage across AB is the same as voltage across 5Ω resistor *i.e.*

Voltage across AB = Voltage across 5Ω resistor

$$= \text{current} \times \text{resistance} = (2/3) \times 5 = (10/3)\ V$$

Now we shall calculate the impedance across AB when the voltage source is replaced by zero resistance as shown in fig. 10b. As viewed across A, B, 10 Ω and 5 Ω resistances are parallel and then 10 Ω resistor is in series. Hence the impedance Z' between A and B is

$$Z'=10+\frac{10\times 5}{10+5}=\frac{40}{3}\Omega$$

The Thevenin's circuit is shown in fig. 22B (c).

The current in the load, when connected across R_L is given by

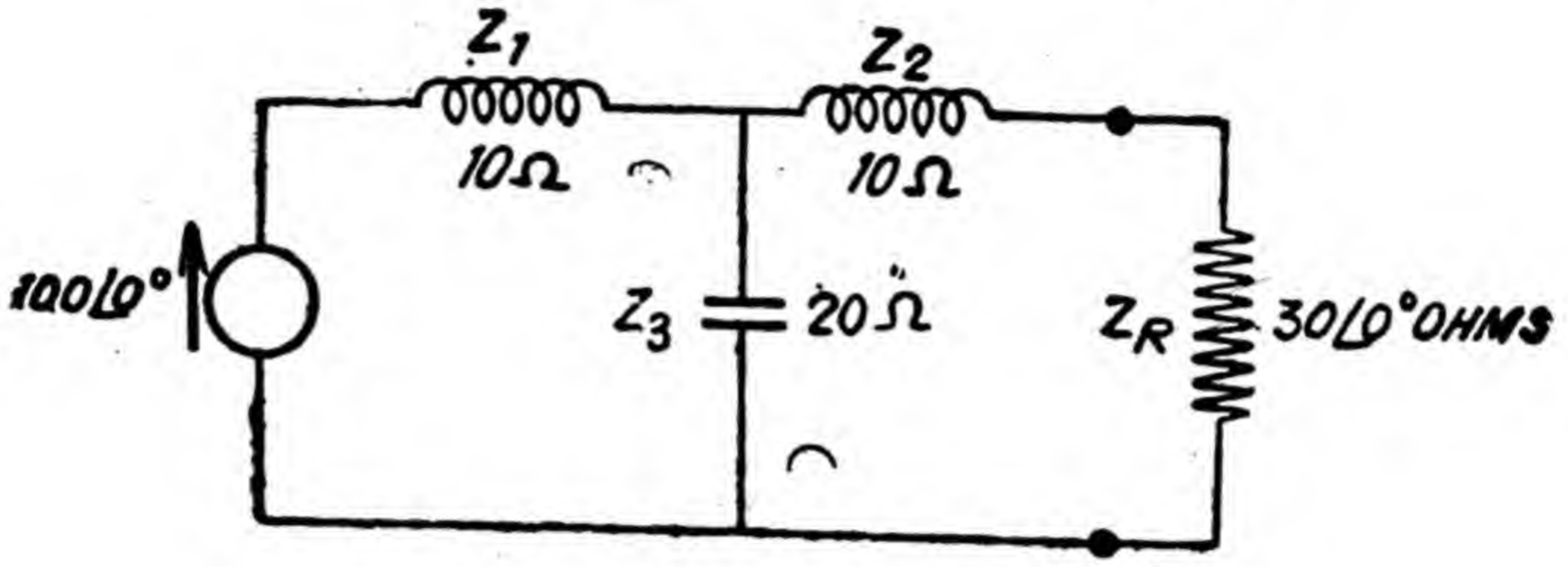
$$I_L=\frac{\text{Voltage}}{\text{Resistance}}=\frac{(10/3) \text{ Volt}}{\{(40/3)+100\} \text{ ohm}}=\frac{1}{34} \text{ amp.}$$

Problem 5. For the network shown in fig. (23), calculate the current through load impedance Z_R, using Thevenin's theorem.

Let I_R be the current flowing through Z_R, given by

$$I_R=\frac{E'}{Z'+Z_R}$$

where $$E'=\frac{EZ_3}{Z_1+Z_3} \text{ and } Z'=Z_2+\frac{Z_1Z_3}{Z_1+Z_3}$$



Fig, 23

So $$E'=\frac{(100+j\,0)(-20\,j)}{j\,10-j\,20}=\frac{(100+j\,0)(-j\,20)}{-j\,10}$$

$$=(200+j\,0) \text{ volts.}$$

and $$Z'=j\,10+\frac{(j\,10)(-j\,20)}{j\,10-j\,20}$$

$$=j\,10+j\,20=j\,30 \text{ ohms.}$$

So $$I_R=\frac{E'}{Z'+Z_R}=\frac{200+j\,0}{j\,30+30+j\,0}=\frac{200}{30+j\,30}$$

$$I_R=(3{\cdot}33-j\,3{\cdot}33) \text{ amps.}$$

ar $$I_R=\sqrt{\{(3{\cdot}33)^2+(3{\cdot}33)^2\}}\tan^{-1}\left(-\frac{3{\cdot}33}{3{\cdot}33}\right)$$

$$I_R=4{\cdot}72\angle -5^\circ \text{ amps.}$$

Problem 6. Given the following network, find the impedance Z_0 looking into the terminals ab, with the help of the Thevenin's theorem and calculate also that the drop across ab is $150\angle 0°$ volts. Calculate the current in a load impedance $Z_L=10-j\,7{\cdot}5$ ohms connected across ab.

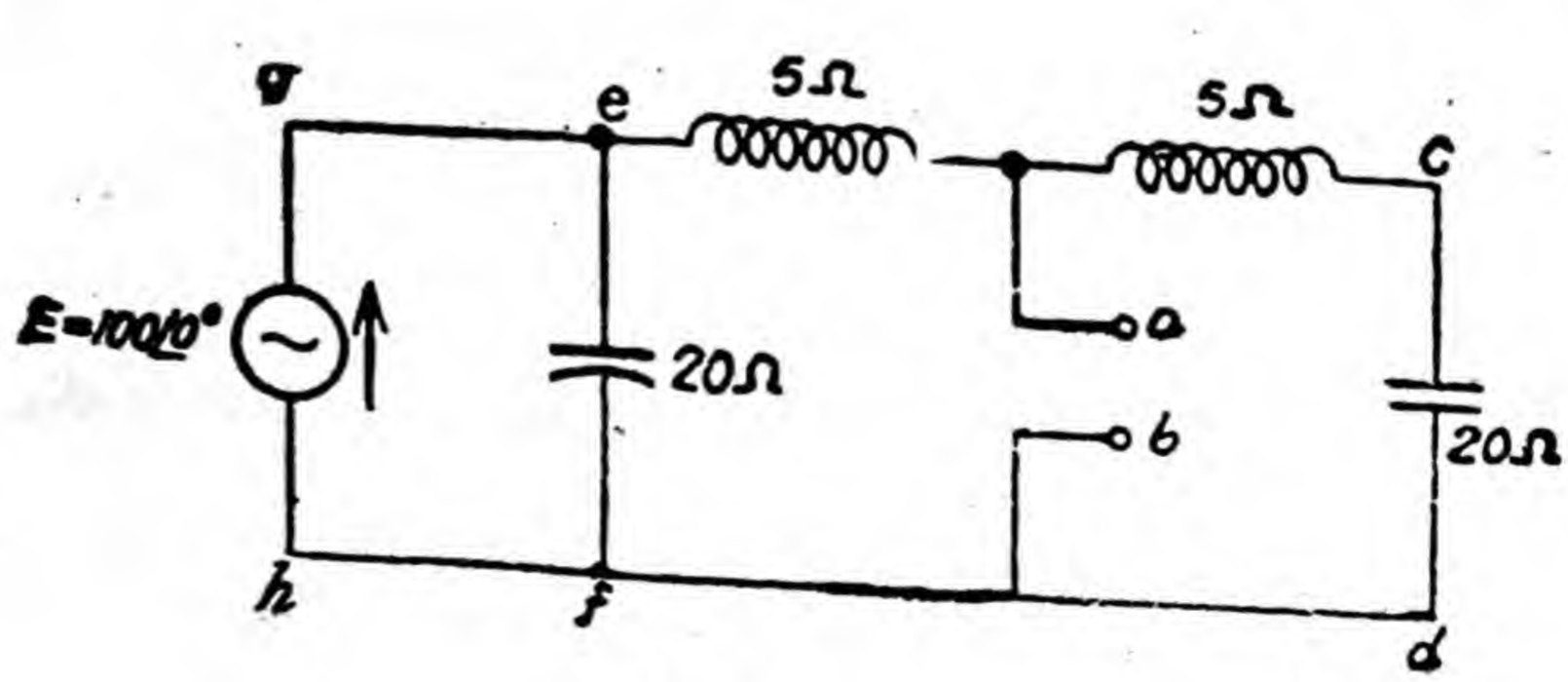


Fig. 24

For calculating the impedance Z_0 looking into *ab*, we short circuit the branch *gh* then

$$Z_0=\frac{(+j5-j20)(j5)}{j5-j20+j5}=\frac{(-j15)(j)}{-j10}=j7{\cdot}5 \text{ ohms}$$

Now $Z_{ocd}=j5+j5-j20=-j10$

So $Z_{fdc}=\frac{(-j20)(-j10)}{-j20-j10}=\frac{(-j20)(-j10)}{(-j30)}$

$=-j6{\cdot}66$ ohms.

So $I_{ge}=\frac{E}{Z_{fdce}}=\frac{100+j0}{-j6{\cdot}66}=j15$ amps.

So $I_{ce}=\frac{j15}{-j30}\times -j20=j10$ amps.

and $I_{ef}=\frac{j15}{-j30}\times -j10=j5$ amps.

Now voltage across *ab i.e.* V_{ab} will be :

voltage across *ef* $\pm$ voltage across $j5\ \Omega$;

$V_{ab}=100+j0\pm j5\times j10$

$=100\pm -50$

$V_{ab}=150\ \angle 0°$.

This V_{ab} can also be calculated as

$V_{od}=-j20\times j10=200$

and voltage across $j5\Omega=j5\times j10=-50$

So $V_{ab}=V_{od}+Vj5=200-50=150\ \angle 0°$ ohms.

and $I_{load}=\frac{V_{ab}}{Z_0+Z_L}$

$=\frac{150+j0}{j7{\cdot}5+10-j7{\cdot}5}=\frac{150+j0}{10}$

$=15\ \angle 0°$ amp.

3·6. NORTON'S THEOREM :

This theorem states that any two terminal network consisting of linear impedances and generators can be replaced by an equivalent circuit containing a current source I′ in parallel with an admittance Y′. The value of I′ is the short circuit current between the terminals of the network, and Y′ is the admittance measured between the terminals with all generators removed (but not their admittances).

This theorem may be easily proved by considering a Thevenin's equivalent network shown in fig. (25*a*). The load impedance Z_R is appearing between two terminals *a* and *b*. Now this Thevenin's equivalent circuit may be easily converted into a circuit containing current source *I′* in parallel with *Y′* and Y_R appearing between two terminals *a* and *b* as shown in fig. (25*b*).

The value of I_R from fig. (23*a*) is given by

$$I_R=\frac{E'}{Z'+Z_R}=\frac{E'}{\frac{1}{Y'}+\frac{1}{Y_R}}=\frac{E'Y'Y_R}{Y'+Y_R}, \qquad \ldots(2)$$

where Y' and Y_R are the reciprocal of Z' and Z_R respectively, known as admittances.

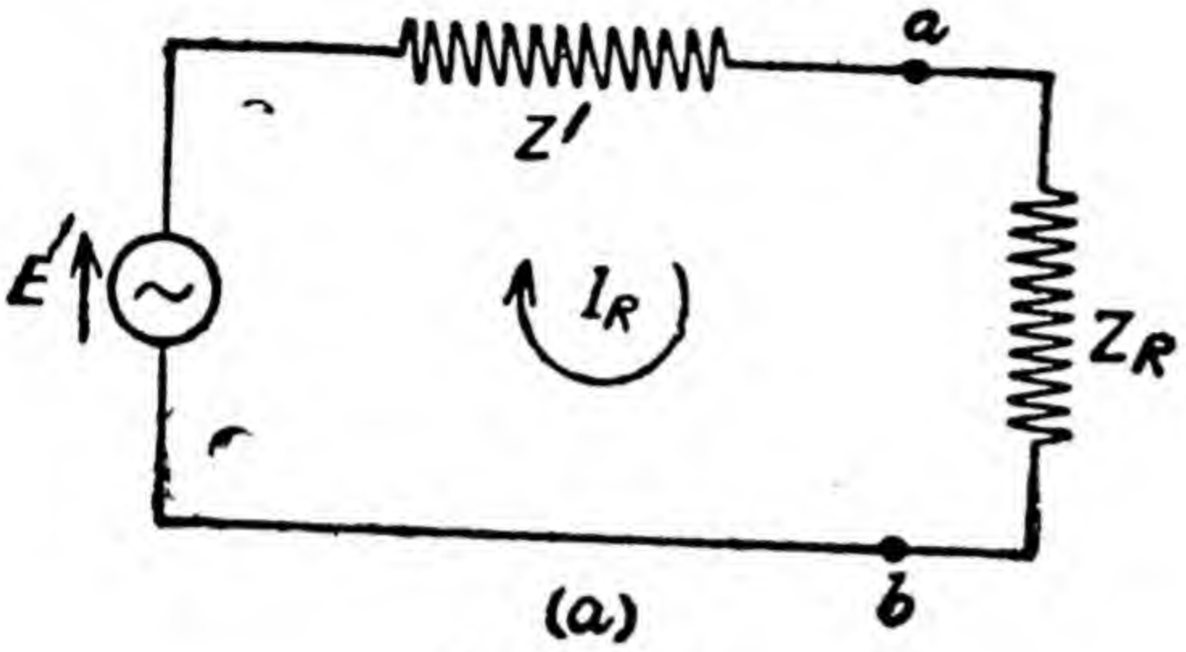


Fig. 25 (a)

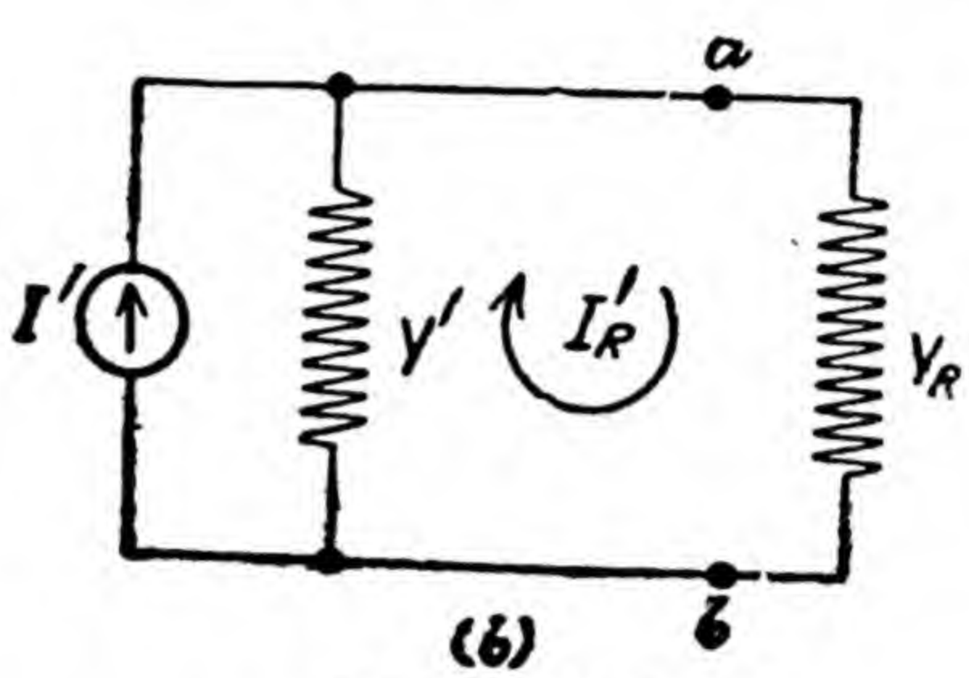


Fig. 25 (b)

Applying the current division law in fig. (25-*b*); we have

$$I_R'=\frac{I'\,Y_R}{Y'+Y_R}. \qquad \ldots(2)$$

The load current I_R' can be made equal to I_R. Then the comparison of equations (1) and (2) gives

$$I'=E'Y'=\frac{E'}{Z'}. \qquad \ldots(3)$$

Equation (3) clearly indicates that circuits (25*a*) and (25*b*) are the same. The current I' is the short circuit current of the Thevenin's equivalent network of fig. (25*b*). Thus we observe that fig. (25*b*) is equivalent to fig. (25*a*) and fig. (25*a*) is equivalent to the original fig. (18*a*).

Thus we see that interchange of voltage and current sources with the help of Thevenin's and Norton's theorems give a method of circuit analysis. As described earlier, voltage source is removed from a circuit by short circuiting its e.m. f. whereas a current source is removed by opening its circuit.

Problem 1. Convert the following linear network Thevenin's equivalent network and then into Norton's equivalent network and show that power delivered to the load R_L in both the cases is same.

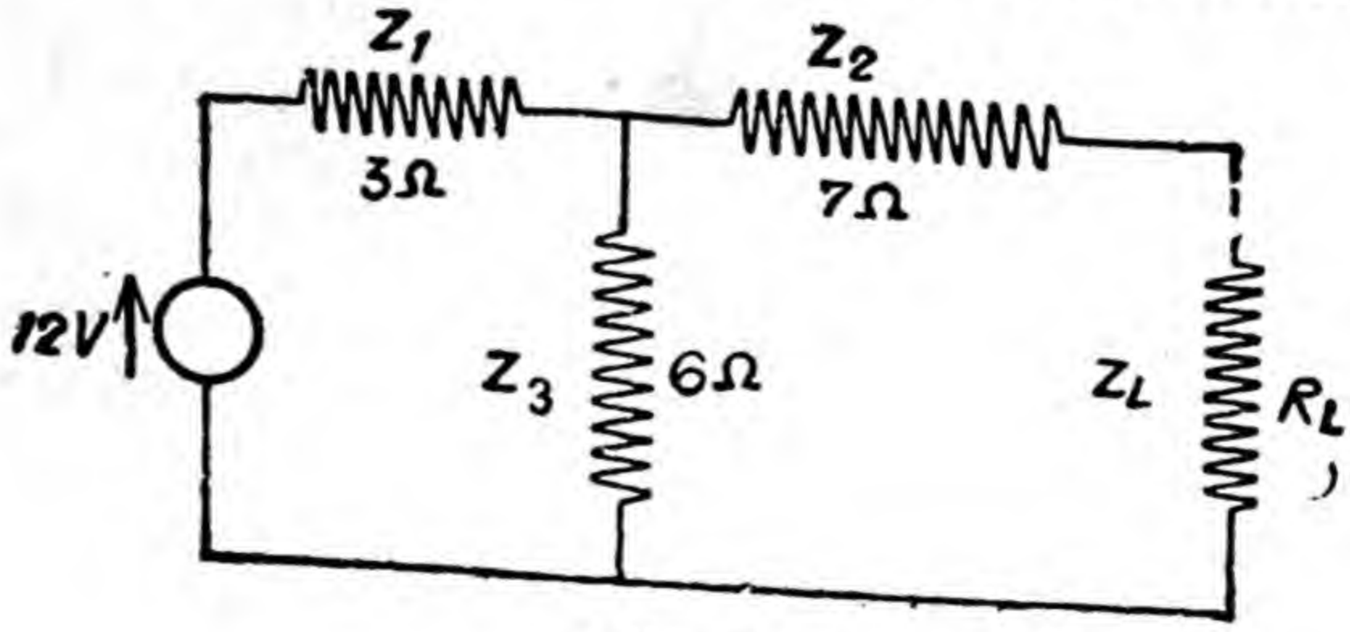


Fig. 26.

For converting this network into Thevenin's equivalent, we have to find out the Thevenin's components given as

$$E' = \frac{EZ_3}{Z_1+Z_3} = \frac{12\times 6}{3+6} = \frac{12\times 6}{9} = 8 \text{ volts,}$$

and $$Z' = Z_2 + \frac{Z_1 Z_3}{Z_1+Z_3} = 7 + \frac{3\times 6}{3+6} = 7 + \frac{18}{9} = 9 \text{ ohms.}$$

Hence the Thevenin's equivalent circuit is shown in fig. 27, and power delivered to the load is

$$P_L = \left(\frac{8}{9+R_L}\right)^2 R_L \text{ watts} \qquad \text{...(A)}$$

since $P_L = I^2 R_L$

Now applying the Norton's theorem in fig. (27), we have

$$I' = E'Y' = \frac{E'}{Z'} = \frac{8}{9} \text{ amps.}$$

and $$Y' = \frac{1}{Z'} = \frac{1}{9} \text{ ohms.}$$

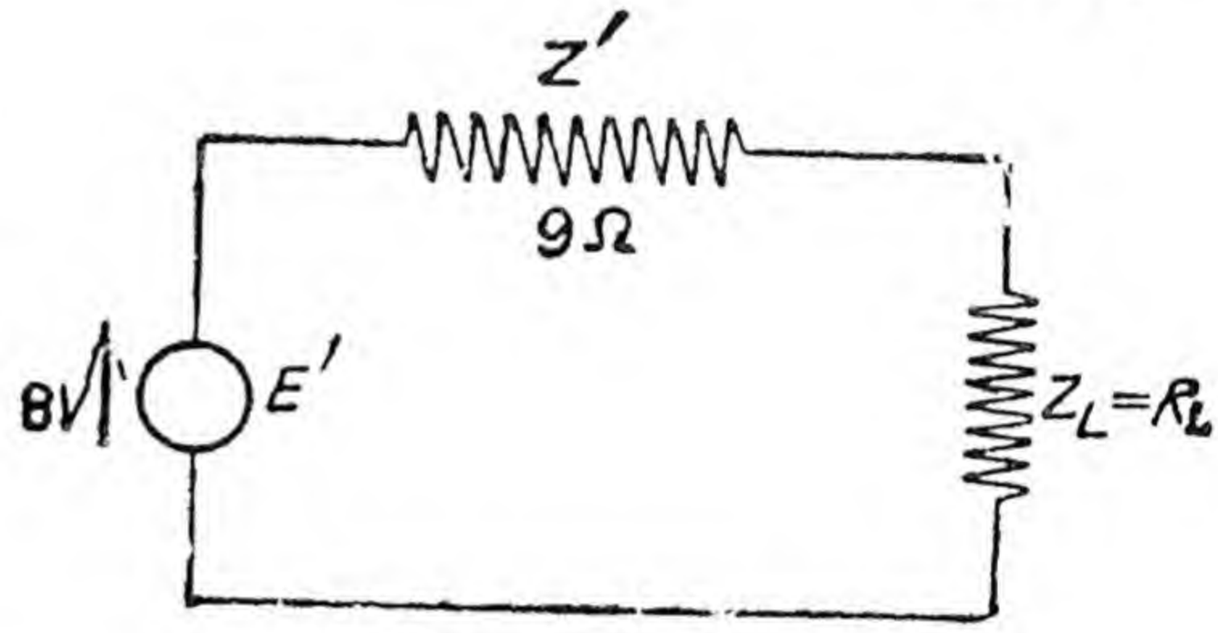


Fig. 27

Hence the Norton's equivalent circuit is shown in fig. (25). Now the power delivered to the load is calculated by applying the current division law.

The current through Y_L is given by

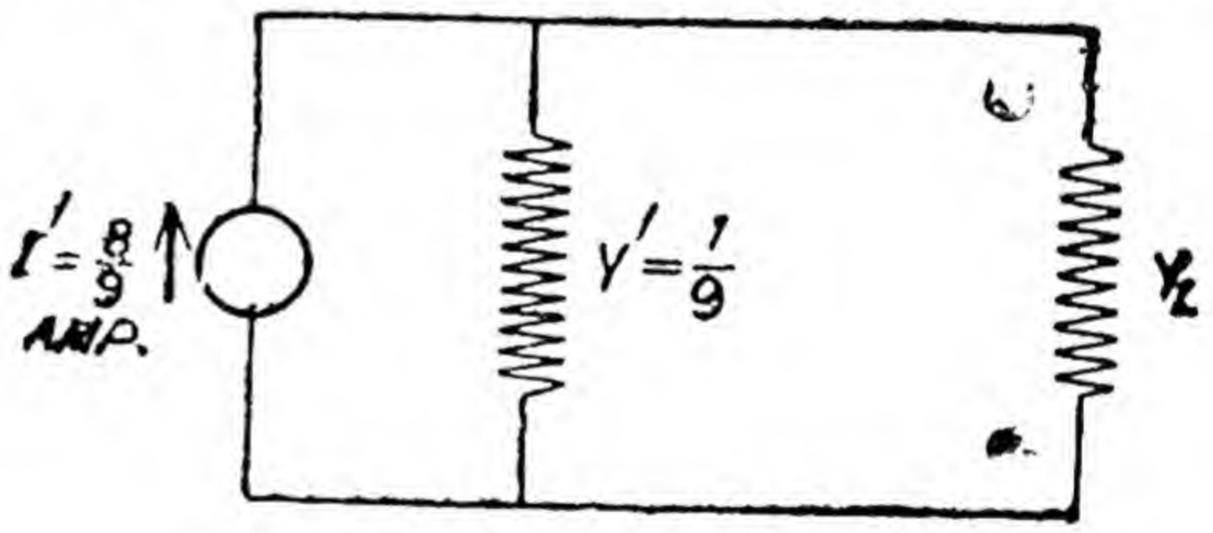


Fig. 28

$$I_L = \frac{8}{9}\cdot\frac{Y_L}{1/9+Y_L}$$

$$P_L = (I_L)^2 \times R_L = (I_L)^2 \frac{1}{Y_L}$$

or $$P_L = \left(\frac{8}{9}\cdot\frac{Y_L}{1/9+Y_L}\right)^2 \frac{1}{Y_L}$$

or $$P_L = \left(\frac{8}{9+R_L}\right)^2 R_L \text{ watts.} \qquad \text{...(B)}$$

From equations (A) and (B) we see that in both the cases power delivered to the load is the same and hence we conclude that both Thevenin's and Norton's circuits are equivalent to the original circuit.

Problem 2. Draw the Thevenin's and Norton's equivalent circuits for the following circuit (fig. 29). Calculate the current in the load in each case.

Sol. First of all we shall calculate the open circuited voltage across AB. The current I flowing in the loop is given by

$$0{\cdot}6I + 0{\cdot}8I + 0\ 2I = 24 \quad i.e.,\ I = 15 \text{ amp.}$$

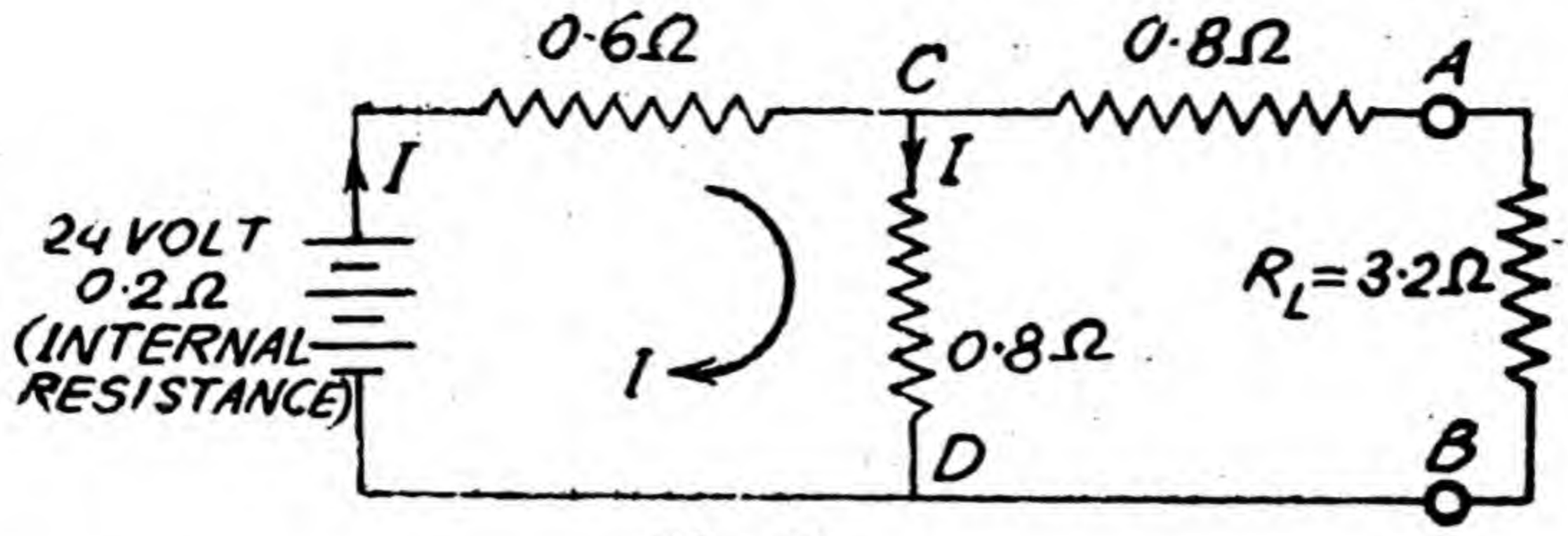


Fig. 29.

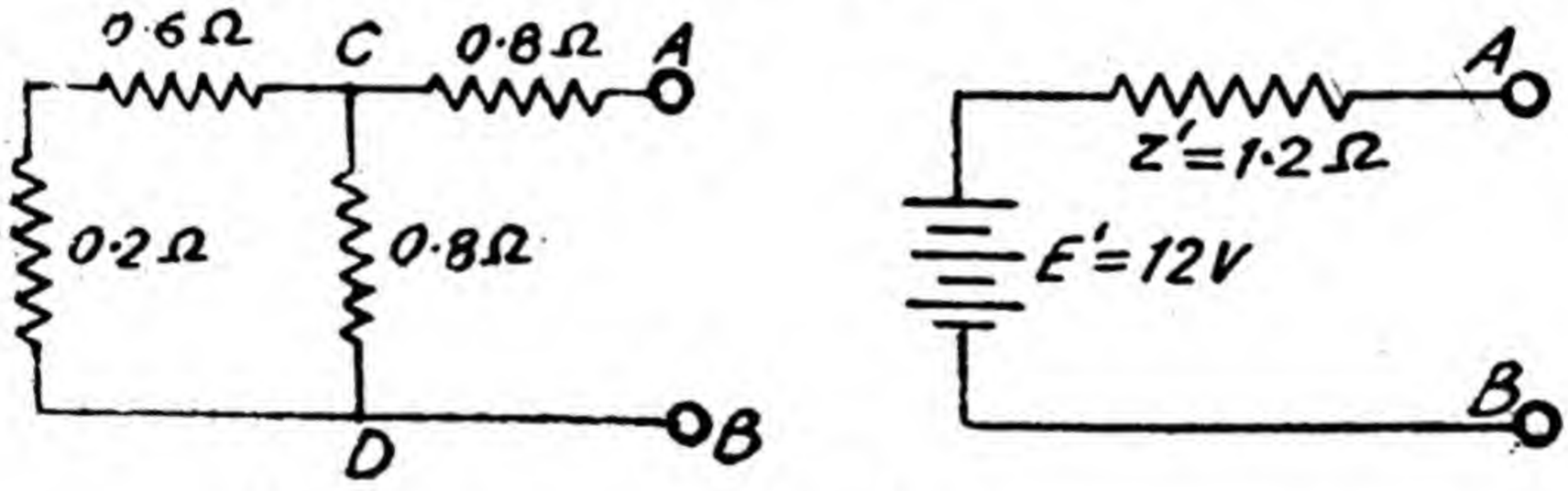
The voltage across AB is the same as voltage across CD, *i.e.*,

E'=voltage across CD=15 amp.×0·8Ω=12 Volt

Now we shall calculate impedance Z' across AB. For this purpose the battery is removed but not its internal resistance. The circuit now becomes as shown in fig. 30.

As viewed from points A and B, the resistors 0·2Ω and 0·6 Ω are in series. Their equivalent resistance is 0·8 Ω. The resistor 0·8 between C and D is in parallel. Hence equivalent resistance is given by

$$\frac{1}{R}=\frac{1}{0\cdot8}+\frac{1}{0\cdot8}=\frac{1+1}{0\cdot8}=\frac{2}{0\cdot8} \quad i.e.\ R=0\cdot4\Omega$$



(*a*) Fig. 30 (*b*)

Now resistor 0·8 Ω between C and A is in series. Hence the imdedance Z'=0·8+0 4=1 2 Ω.

The Thevenin's equivalent circuit is shown in fig. 30b. The current in load R_L when connected between AB is given by

$$I_L=\frac{12\text{ Volt}}{1\cdot2\Omega+3\cdot2\Omega}=\frac{12\text{ Volt}}{4\cdot4\ \Omega}=\frac{30}{11}\text{ amp.}$$

The Norton's equivalent circuit may be found from Thevenin's equivalent circuit. The short circuited current found by shorting the terminals A and B together is

$$I_{AB}=\frac{\text{Voltage}}{\text{Resistance}}=\frac{12\text{ Volt}}{1\cdot2\ \Omega}=10\text{ amp.}$$

The No ton's equivalent ci.cuit is shown in fig. 31.

When the load R_L=3·2 Ω is connected across AB, the current in the load is given by

$$I_L=\frac{1\cdot2\ \Omega}{3\cdot2\Omega+1\cdot2\Omega}\times10\text{ amp.}$$

$$=\frac{12}{4\cdot4}=\frac{30}{11}\text{ amp.}$$

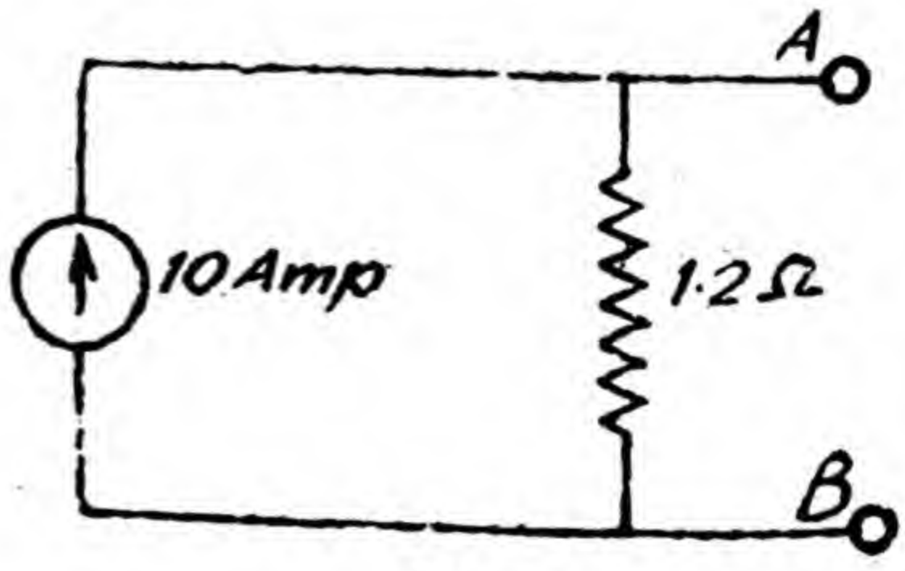


Fig. 31.

Problem. 3. For the network shown in fig. (32), calculate the current through load impedance Z_R using Norton's theorem.

First of all using Thevenin's theorem, we shall calculate the Thevenin's components E' and Z' given as

$$E'=\frac{EZ_3}{Z_1+Z_3}=\frac{(100+j0)\,(-20j)}{j10-j20}=\frac{(100+j0)\,(-j20)}{-j10}$$

or $E'=(200+j0)$ volts ...(A)

and $Z'=Z_2+\dfrac{Z_1Z_3}{Z_1+Z_3}$

$$=j10+\frac{(j10)\,(-j20)}{j10-j20}$$

$Z'=j10+j20=j30.$ ...(B)

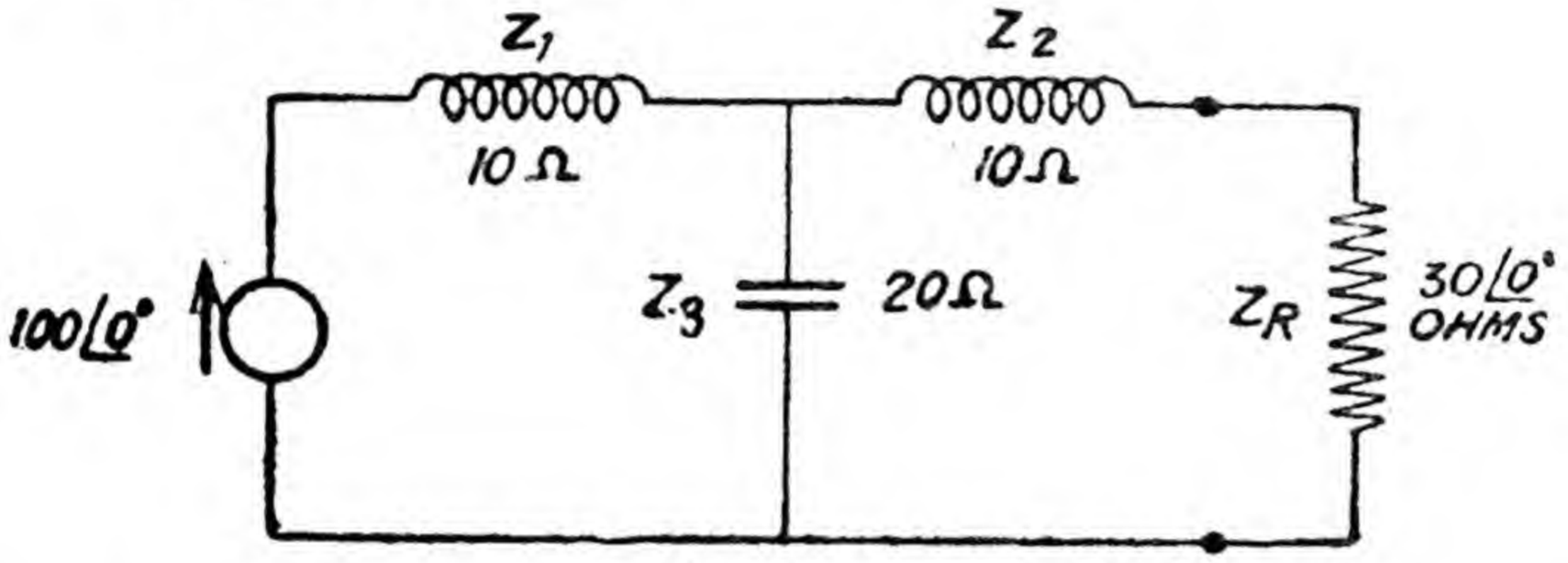


Fig. 32.

Now, we will use Norton's theorem for calculating the current I_R through load impedance Z_R. Current I_R is given by

$$I_R=\frac{I'Y_R}{Y'+Y_R}$$

and $I'=E'Y'$ and $Y'=\dfrac{1}{Z'}$

$$\therefore \quad Y'=\frac{1}{j30}$$

$$I'=(200+j0)\,\frac{1}{j30}=\frac{200}{j30}$$

$$I_R=\frac{\dfrac{200}{j30}\times\dfrac{1}{30+j30}}{\dfrac{1}{j30}+\dfrac{1}{30+j0}}=\frac{200}{30+j30}$$

or $I_R=3{\cdot}33-j3{\cdot}33$ amps.

or $I_R=\sqrt{\{(3{\cdot}33)^2+(3{\cdot}33)^2\}}\,\tan^{-1}\left(-\dfrac{3{\cdot}33}{3{\cdot}33}\right)$

or $I_R=4{\cdot}72\angle -45°$ amps.

Problem 4. Find the open-circuit voltage and Thevenin's resistance of the two terminal network shown in fig. (33) and then reduce the network into Norton's equivalent circuit, and find the current through load resistance $Z_R=2$ ohms. Find also the power delivered to the load.

The Thevenin's resistance Z' can be calculated by short

circuiting *df* branch. In that case *cd* and *cf* will be parallel to each other; and *de* and *fe* will be parallel to each other and both will be in series then

$$Z' = \frac{R_{cd} \times R_{cf}}{R_{cd} + R_{cf}} + \frac{R_{de} \times R_{fe}}{R_{de} + R_{fe}}$$

or $$Z' = \frac{3 \times 7}{3+7} + \frac{1 \times 9}{1+9}$$

$$= \frac{21}{10} + \frac{9}{10} = 3 \text{ ohms.}$$

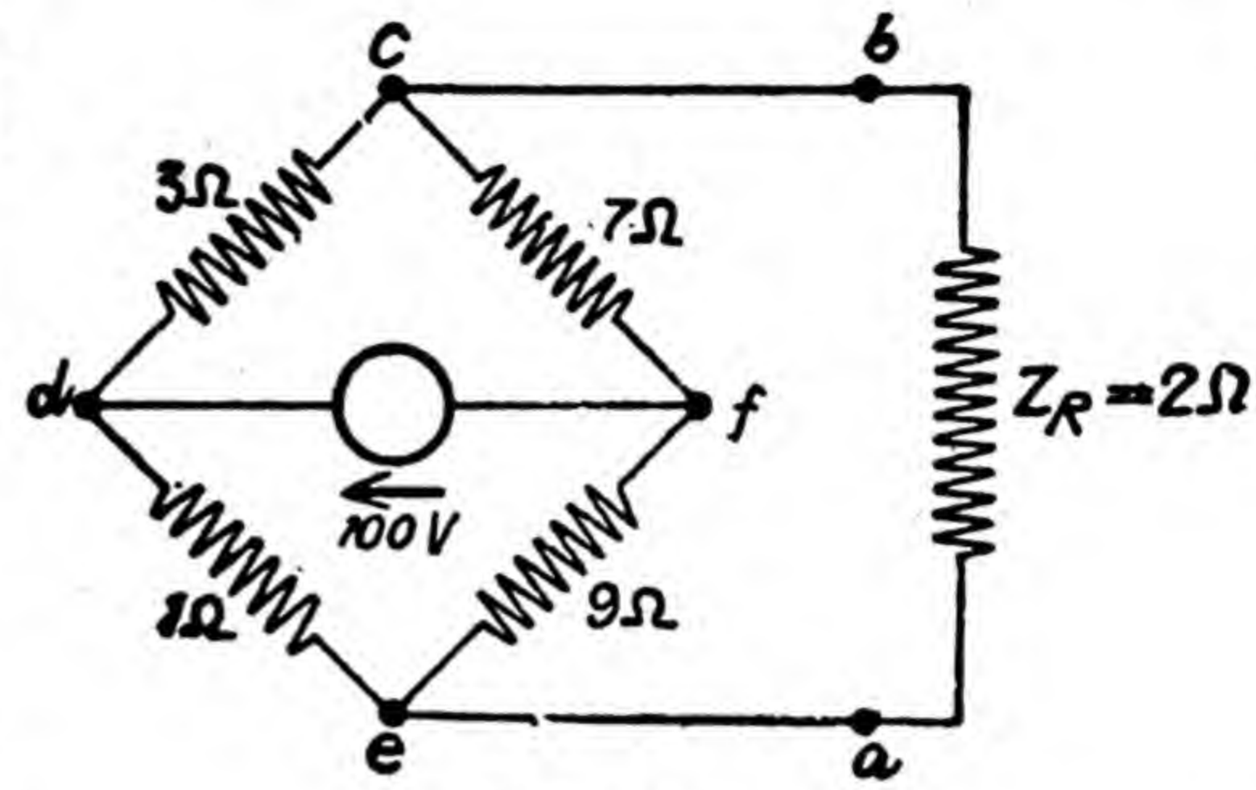


Fig. 33.

Now we have to calculate the open circuited voltage between terminals *a* and *b*. We see from fig. 33 that the voltage drop between *ab* is same as at *c* and *e* points. The current generated from the source in each branch is voltage divided by the total resistance.

i.e. $$\frac{100}{10} = 10 \text{ amps.}$$

P.D. across $cd = -30$ volts
P.D. across $de = 10$ volts.
P.D. across $ce = -20$ volts.

Calculating potential difference across *cf* and *fe* we would get potential difference across *ce* as −20 volts.

$E' = -20$ volts and $Z' = 3$ ohms.

Now the circuit reduces into the Thevenin's equivalent as shown in fig. 34.

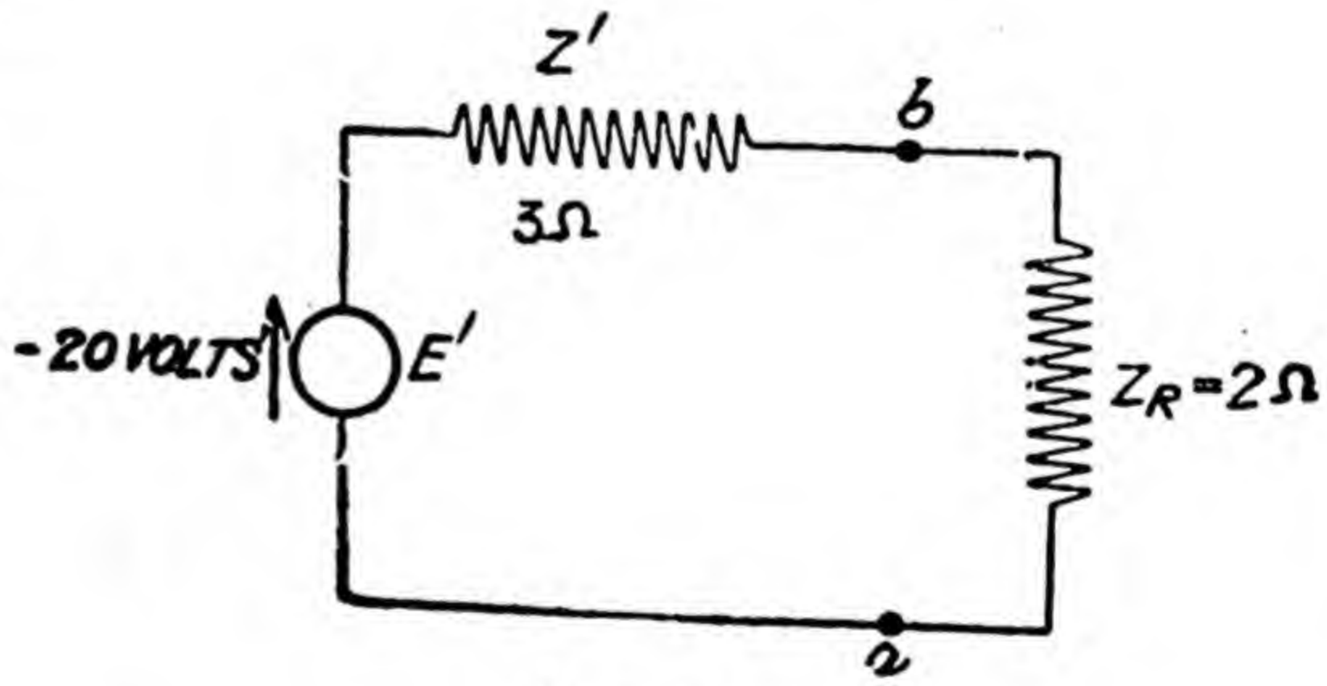


Fig. 34.

Now we have to convert this network into Norton's equivalent circuits. For this let us find out I' and Y' and Y_R

$$Y' = \frac{1}{Z'} = \tfrac{1}{3} \text{ ohms.}$$

$$Y_R = \frac{1}{Z_R} = \tfrac{1}{2} \text{ ohms.}$$

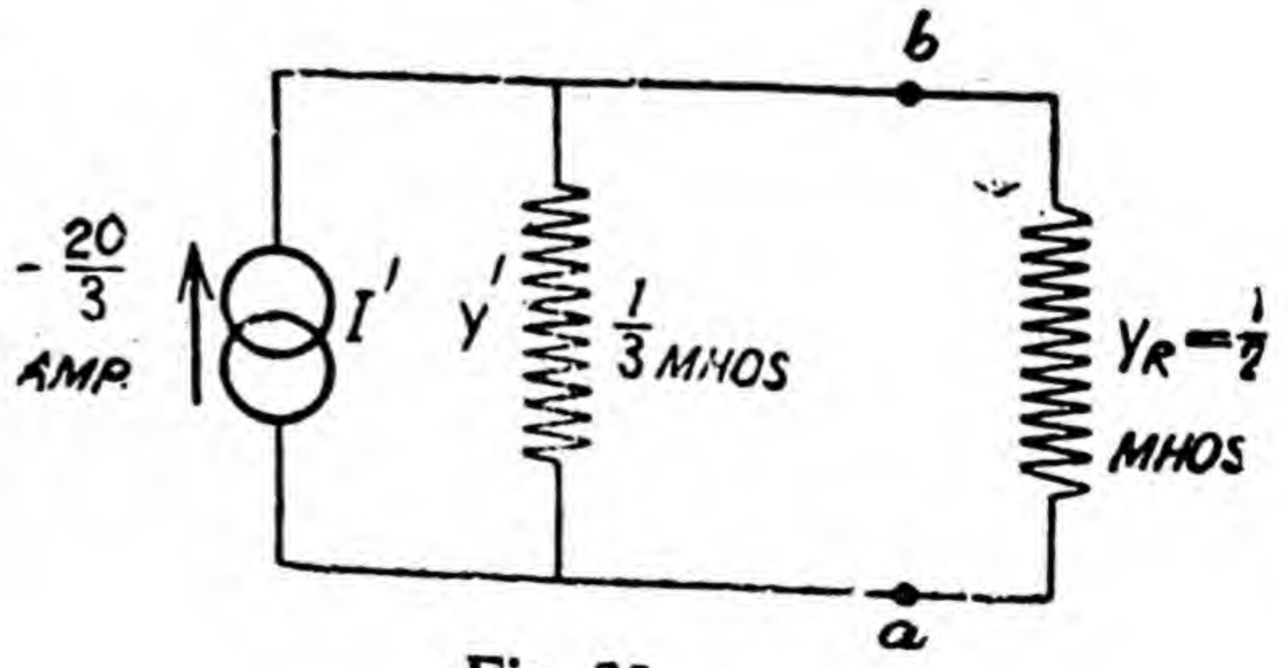
$$I' = E'Y' = -20 \times \tfrac{1}{3}$$

$$= -\tfrac{20}{3} \text{ amp.}$$

So the Norton's circuit is as shown in fig. (35). Current through the load is given by

$$I_R = \frac{I' Y_R}{Y' + Y_R}$$

$$= -\frac{(20/3) \times 1/2}{(1/3) + (1/2)}$$



Fig, 35.

$I_R = -10/3 \times 6/5 = -4$ amp.

This is in *ba* direction, hence current in *ab* direction is 4 amps.

Therefore power delivered to the load is $I^2 Z_R$ or

$$P_L = (I)^2 \frac{1}{Y_R} = (4)^3 \frac{1}{1/2}$$

$$P_L = 16 \times 2 = 32 \text{ watts.}$$

3·7. THE MAXIMUM POWER-TRANSFER THEOREM :

This theorem states that the maximum power will be consumed by a network from another circuit connected to its two terminals when the impedance of the receiving network is varied to make the impedance looking into the network at its two terminals as conjugate to each other.

Let us consider a two terminal active linear network connected to a two terminal passive linear network. In fig. (36) Z_R represents the equivalent impedance of a passive linear network and the network at the left of a, b terminals represents the Thevenin's equivalent active network.

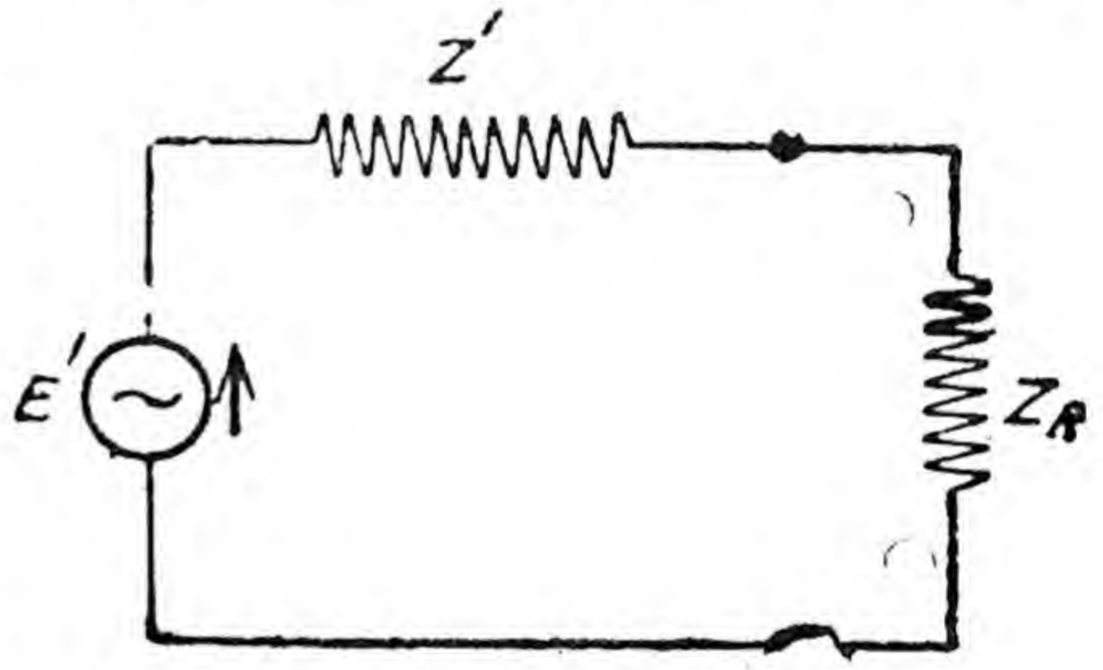


Fig. 36.

The impedance of the active network Z' is equal to the ohmic resistance R' plus the resistance X', or

i.e. $$Z' = R' + jX'. \qquad \ldots(1)$$

Similarly $$Z_R = R_R + jX_R \qquad \ldots(2)$$

We have to prove that Z' and Z_R are conjugate to each other

i.e. $$Z_R = R' - jX'. \qquad \ldots(3)$$

For this, we proceed as follows : let I be the current flowing in the network, then

$$I = \frac{E'}{Z' + Z_R} \qquad \ldots(4)$$

or $$I = \frac{E'}{(R' + R_R) + j\,(X' + X_R)}. \qquad \ldots(5)$$

The power delivered to the load is

$$P_L = \frac{(E')^2\, R_R}{(R_R + R')^2 + (X_R + X')^2} \qquad \ldots(6)$$

If X_R is varying, the maximum power will be calculated by putting $\frac{\partial P_L}{\partial X_R} = 0$. From equatior (6), we get

$$\frac{\partial P_L}{\partial X_R} = \frac{-2\,(E')^2\, R_R\,(X_R + X')}{[(R_R + R')^2 + (X_R - X')^2]^2} = 0 \qquad \ldots(7)$$

or $$X_R+X'=0$$
or $$X_R=-X'. \qquad \ldots(8)$$

Substituting equation (8) in equation (6), we get

$$(P_L)_{max}=\frac{(E')^2\ R_R}{(R_R+R')^2}. \qquad \ldots(9)$$

Now suppose R_R is varying, then maximum power will be calculated by putting $\frac{\partial P_L}{\partial R_R}=0$. From equation (9)

$$\frac{\partial P_L}{\partial R_R}=\frac{(E')^2(R_R+R')^2-2(E')^2\ R_R\ (R_R+R')}{(R_R+R')^4}=0 \qquad \ldots(10)$$

or $$(E')^2\ (R_R+R')-2(E')^2\ R_R=0$$
or $$(E')^2\ [-R_R+R']=0$$
or $$R_R=R' \qquad \ldots(11)$$

From equations (8) and (9), we have

$$Z_R=R_R+jX_R=R'-jX'.$$

Corollary. *If only the absolute magn itude and not the angle of Z_R be varied, then the greatest power output will be delivered from the network if the absolute magnitude of Z_R is made equal to the absolute magnitude of Z'.*

Proof. The power delivered to the load can be written as

$$P_L=\frac{(E')^2\ |\ Z_R\ |\ \cos\theta}{(R'+|\ Z_R\ |\ \cos\theta)^2+(X'+|\ Z_R\ |\ \sin\theta)^2}, \qquad \ldots(12)$$

since, we know that

$$R_R=|\ Z_R\ |\ \cos\theta$$
$$X_R=|\ Z_R\ |\ \sin\theta$$

Equation (12) can be maximized with respect to $|\ Z_R\ |$. We have

$$\frac{\partial P_L}{\partial\ |\ Z_R|}=\frac{\begin{array}{r}(E')^2\cos\theta\ [(R'+|Z_R|\cos\theta)^2+(X'+|Z_R|\sin\theta)^2]\\ -(E')^2\ |Z_R|\cos\theta\ [2(R'+|Z_R|\cos\theta)\cos\theta\\ +2(X'+|Z_R|\sin\theta\ \sin\theta]\end{array}}{[(R'+|\ Z_R\ |\cos\theta)^2+\ X'+|Z_R|\sin\theta)^2]^2}=0$$

or $$(E')^2\cos\theta\ [R'^2+X'^2+|Z_R|^2+2\ |\ Z_R\ |\ R'\cos\theta+2\ |\ Z_R\ |\ X'\sin\theta$$
$$-2\ |\ Z_R\ |\ R'\cos\theta-2\ |\ Z_R\ |^2\cos^2\theta-2\ |\ Z_R\ |\ X'\sin\theta$$
$$-2\ |\ Z_R\ |^2\sin^2\theta]=0$$
or $$[R'^2+X'^2+\ |\ Z_R\ |^2-2\ |\ Z_R\ |^2\ (\cos^2\theta+\sin^2\theta)]=0$$
or $$R'^2+X'^2-|\ Z_R\ |^2=0$$
or $$|\ Z'\ |^2=|\ Z_R\ |^2$$
or $$|\ Z_R\ |=|\ Z'\ |. \qquad \ldots(13)$$

Hence, if the angle of the load impedance is not altered, the magnitude of the load impedance can be made equal to the magnitude of the network impedance to obtain the greatest amount of power from the network.

Now, from equations (9) and (13), the maximum power delivered to the load is

$$(P_L)_{max}=\left(\frac{E'}{2R_R}\right)^2 R_R=\frac{(E')^2}{4R_R}. \qquad ...(14)$$

We have proved above that maximum power is taken out from the network, if and only if the impedance of the receiving network is conjugate to the impedance of the active network. In that case, the circuit (31) reduces to that shown in fig. (37).

Here total resistance is
$R'+jX'+R'-jX'=2R'.$

Then $|I'|^2=\frac{|E'|^2}{4R'^2}$

i.e., **power generated is equal to**

$$|I'|^2(2R')=\frac{|E'|^2}{2R'} \qquad ...(15)$$

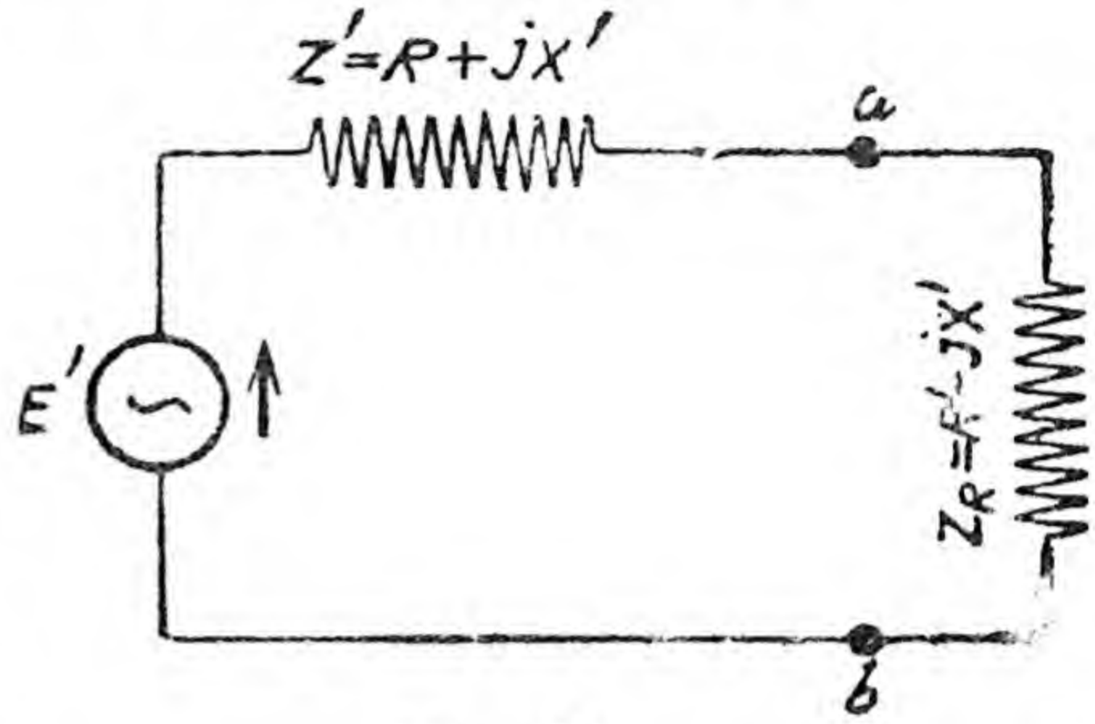


Fig. 37.

Thus, we see from equations (14) *and* (15) *that the power lost in the internal generator is equal to the power delivered to the load and the power efficiency is only* 50 *percent.*

Problem 1. Find the value of R_L which will absorb maximum power, and determine this maximum power in the following network.

For finding out R_L, short circuit. $50V$ source and open circuit the 1 amp. source. We see at-once that $R_L=0$ ohms. Hence $R_L=10$ ohms will absorb maximum power.

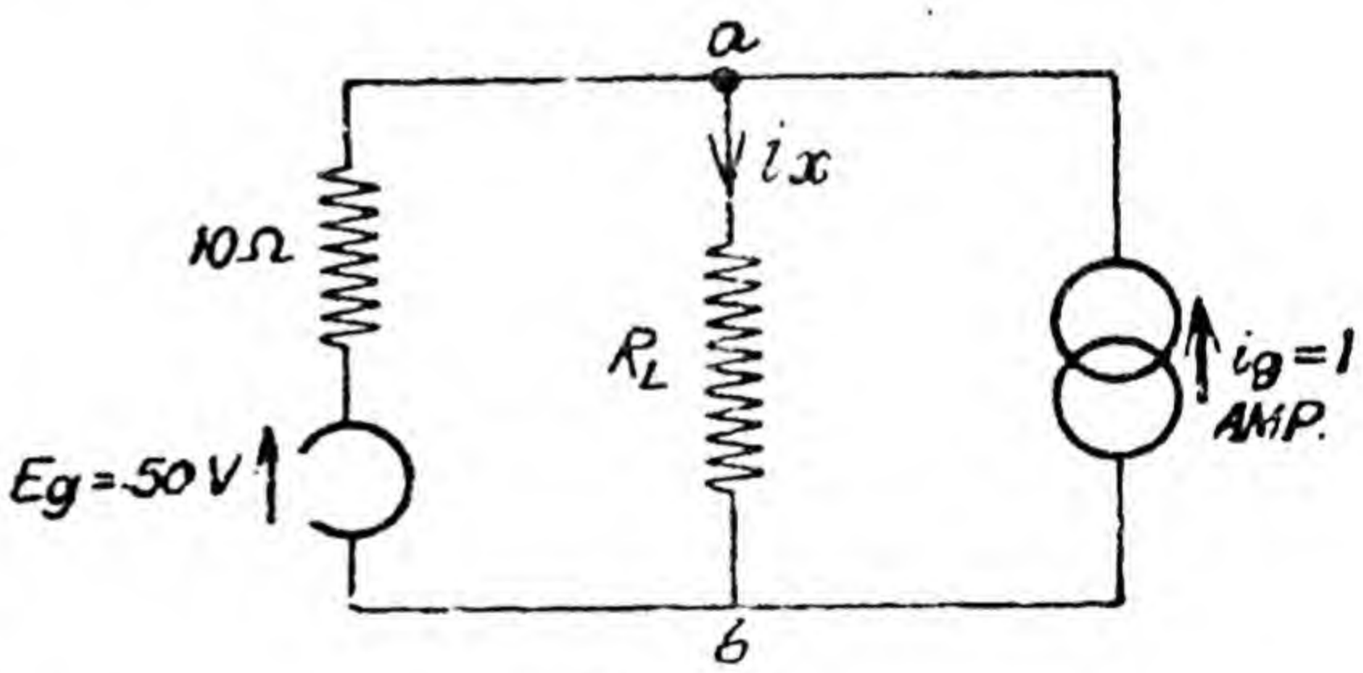


Fig. 38.

For determining the maximum power absorbed by R_L we shall calculate the current flowing through R_L. We can find current through R_L by superposition theorem.

(*i*) Let $i_g=0$, then

$$ix_1=\frac{50}{10+10}=\frac{50}{20}=2{\cdot}5 \text{ amps.}$$

(*ii*) $E_g=3$, then

$$ix_2=\frac{1}{20}\times 10=\frac{10}{20}=0{\cdot}5 \text{ amps.}$$

Hence $i_x=ix_1+ix_2=2{\cdot}5+0{\cdot}5=3{\cdot}0$ amp.

Hence power absorbed by R_L is

$$(i_x)^2 R=(3)^2\times 10=90 \text{ watts}$$

Problem 2. The voltage at two terminals of a network consisting of resistance and reactance is 260 volts. When a resistance of 3 ohms is connected to it, it is found that current of 20 amps. passes through it. When a resistance of 14 ohms is connected, a current of 13 amps is found to flow. Calculate the impedance for which maximum power can be absorbed.

Here $R_R=3$ ohms and 14 ohms respectively and $X_R=0$.

$$|I|=\frac{|E'|}{\sqrt{\{(E_R'+R_R)^2+(X'+X_R)^2\}}}$$

In first case, when $R_R=3$ ohms, we have

$$20=\frac{260}{\sqrt{\{(R'+3)^2+X'^2\}}} \qquad ...(A)$$

In second case, when $R_R=14$ ohms, we have

$$13=\frac{260}{\sqrt{\{(R'+14)^2+X'^2\}}} \qquad ...(B)$$

We see that above equations are satisfied by putting

$R'=2$ ohms and $X'=12$ ohms.

$\therefore \quad Z'=R'+jX'=2+j\ 12$ ohms.

For absorption maximum power the impedance should be conjugate of Z' *i.e.* $2-j\ 12$ ohms.

Problem 3. A generating device has an impedance of 0·5+j 1 ohms and is connected to a load by a line of 1·5+j4 ohms. At what load will maximum power transfer received by the load when adjusted for maximum power ?

For absorb ng maximum power, Z' and Z_R must be conjugate to each other *i.e.*

$Z_R=0{\cdot}5-j1$ ohms

and Z' must be

$1{\cdot}5-j4$ ohms.

Hence maximum power wiil be absorbed when impedance *i.e.* Z_{load} will be

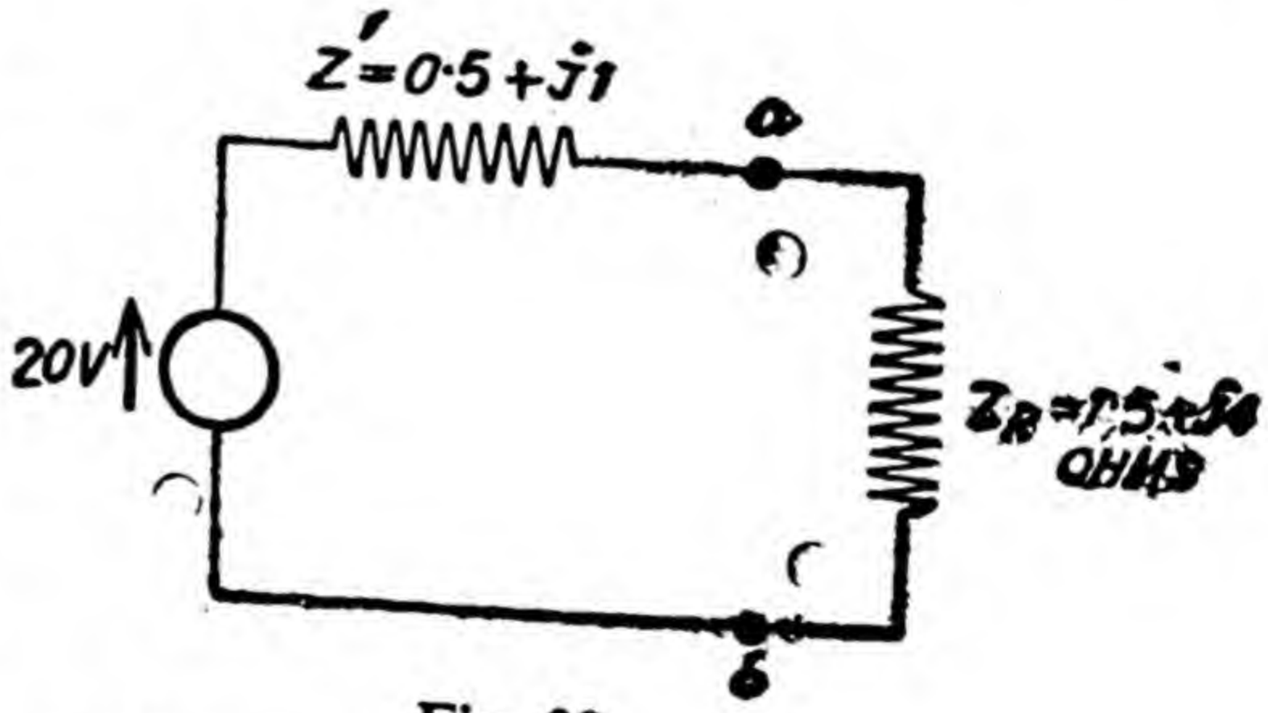


Fig. 39.

$$0{\cdot}5-j\ 1+1{\cdot}5-j\ 4=2-j\ 5 \text{ ohms}$$

and maximum power $P_{max}=\frac{E^2}{4R}=\frac{(20)^2}{4\times 2}$

$=50$ watts at the receiver.

Problem 4. Find the value of R_L which will absorb maximum power, and determine this maximum power of the network shown in fig. (40).

For finding R_L, short circuit 30V source and open circuit 6 amp. source. We find that

$R_L = 10+15=25$ ohms.

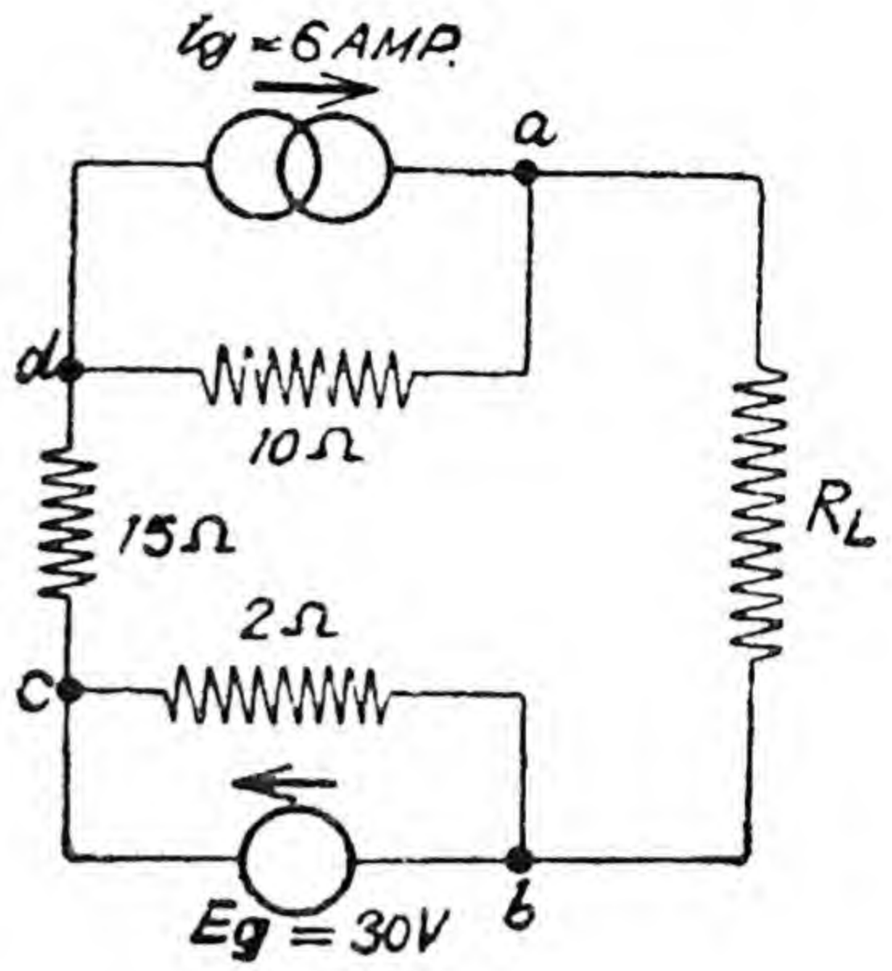


Fig. 40.

2 ohms will not take part in circuit because branch bc is short circuited. Now for finding out the current through R_L, we shall apply super-position theorem.

(i) Put $E_g=0$, bc branch will be short circuited then applying current division law at a node a we have ix_1 through R_L (=25 ohms) :

$$ix_1 = \frac{6\times 10}{25+25} = \frac{60}{50} \text{ amps.}$$

(ii) Put $i_g=0$, then ad branch will be open circuited, in this case branch cb and $cdab$ will be in parallel with each other so equivalent resistance $R_{eq} = \dfrac{R_{cdab}\times R_{cb}}{R_{cdab}+R_{cb}}$.

$$\therefore \quad R_{cdab} = 15+10+25 = 50 \text{ ohms.}$$

$$= \frac{50\times 2}{50+2} = \frac{100}{52}$$

So current generated by the 30 volt source is

$$= \frac{30}{100} \times 52 \text{ amps.}$$

and current flowing through R_L can be calulated by applying current division law at node c.

We have $ix_2 = \dfrac{30\times 52}{100} \times \dfrac{2}{52} = \dfrac{60}{100}$ amps.

So $\quad i_x = ix_1 + ix_2 = \dfrac{60}{50} + \dfrac{60}{100} = \dfrac{90}{50} = \dfrac{9}{5}$ amp.

So power absorbed by R_L is

$$P_L = (i_x)^2 R_L$$

$$\therefore \quad P_L = (9/5)^2 \times 25 = 81 \text{ watts.}$$

EXERCISES AND PROBLEMS

1. What is a four terminal network ? Obtain expression for image parameters in a four terminal network.
2. What do you mean by image transfer constant γ ? Prove that

$$\tanh \gamma = \sqrt{\left(\frac{A_{12}A_{21}}{A_{11}A_{22}}\right)},$$

where the symbols have their usual meanings.

3. What are iterative parameters ? Prove that for a L section the iterative transfer constant is given by the relation

$$e^p = 1 + \frac{Z_{1K}}{Z_b},$$

the symbols have their usual meaning.

4. When the two impedances match together ? Show that the mismatching in *db* can be expressed as

mismatching in $db = 20 \log_{10} (1/k)$.

where k is reflection or transmission factor.

5. State Superposition theorem. Prove this theorem in a case of general network.
6. State Thevenin's theorem and prove it in case of two terminal network.
7. A certain network, including a generator, has an open-circuit voltage of 125 volts, and on short circuit produces a current of 5·59 amp. When a 10 ohm. resistive load is connected, the current is 4·41 amp. By Thevenin's theorem, find an equivalent voltage source circuit for the network. How could the sign of the reactance be determined ?
8. Give the statement of Norton's theorem. Prove that the interchange of voltage and current sources with the help of Thevenin's and Norton's theorems give a method of circuit analysis.
9. What is the maximum power transfer theorem ? Show that power lost in the internal generator is equal to the power delivered to the load, and the power efficiency is only 50%.

4 FILTER CIRCUITS

4·0. FILTER CIRCUITS :

Filters are electrical networks used to separate alternating from direct current components or to separate a group of a c. components included within a particular frequency range from those lying outside this range. To achieve the desired effect, the filter is designed to provide a low attenuation for frequency components within a particular pass band range and a high attenuation at frequencies within other stop band ranges. The networks provide a uniform response over a wide range of frequencies than that obtained with resonant circuits.

Filters are commonly classified inaccordance with their selectivity characteristics as below :

(a) *A low pass filter* transmits all frequencies below a limiting frequency f_o, known as cut off frequency, and stops all those above this frequency.

(b) *A high pass filter* passes frequencies above the cut off frequency and stops all those below this frequency.

(c) *A band pass filter* passes frequencies in a particular band between two cut off frequencies and stops those above and below this band limit.

(d) *A band elimination filter* stops frequencies within a specified band and passes those above and below the limits of this bandb.

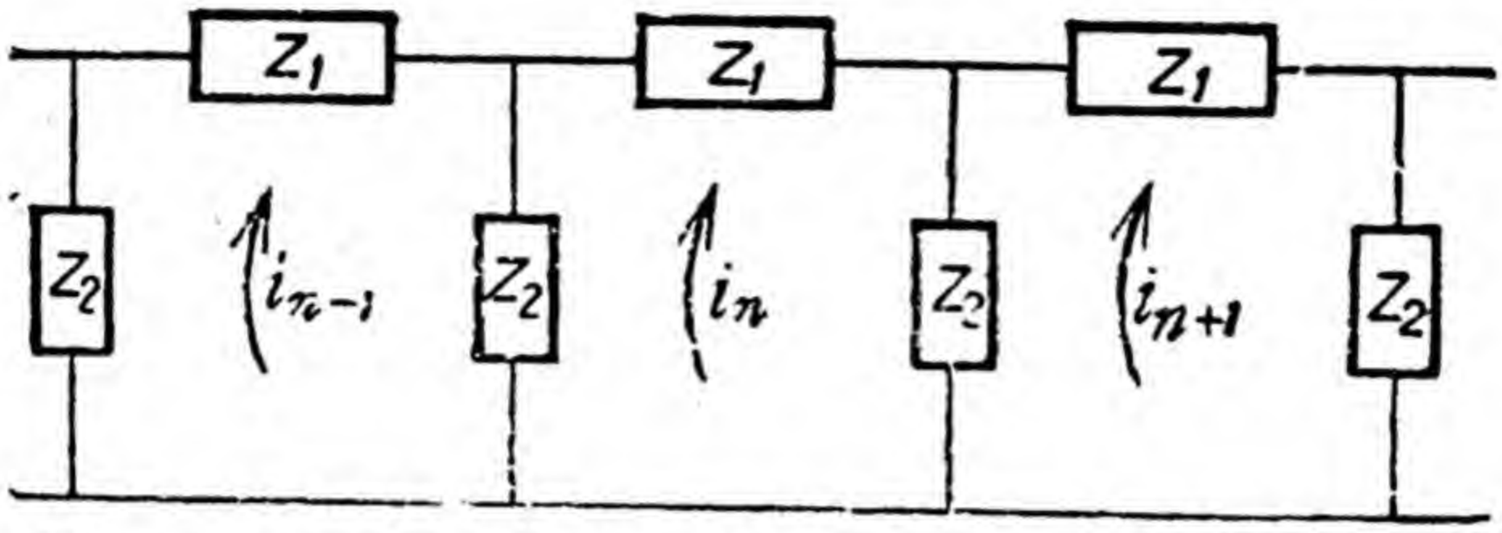
4·1. ELEMENTARY FILTER THEORY :

The network consisting sections each with an impedances Z_1 in series arm and Z_2 in shunt arm constitute the electric wave filter. Z_1 and Z_2 can either be inductances or capacitances or a combination of both.

Because of wave nature of voltages and currents, their evaluation in different elements by the application of Kirchoff's law would be very laborious. For a simple treatment, we proceed in a different way making use of the recurrent nature of the elements.

Let us consider a uniform filter, consisting of a chain of similar section as shown in fig. 1, which are repeated indefinitely

forming an infinite chain. If a generator is applied at some point earlier in the chain, currents will flow in various sections.



Fig, 1. Generalised type of uniform ladder filter.

Let the currents in successive sections be i_{n-1}, i_n, i_{n+1}. Application of Kirchhoff's law to the central section gives the following expression :

$$Z_2\,(i_n-I_{n-1})+Z_1\,i_n+Z_2\,(i_n-i_{n+1})=0$$

or
$$-Z_2 i_{n-1}+(Z_1+2Z_2)\,i_n-Z_2 i_{n+1}=0 \qquad \ldots (1)$$

We may write $i_n=ai_{n-1}$, where a is a real or complex number. Then in an infinite chain where we cannot distinguish between sections, we must have

$$i_{n+1}=ai_n.$$

Then equation (1) gives

$$*\tfrac{1}{2}\left(a+\frac{1}{a}\right)=1+\frac{Z_1}{2Z_2},$$

or
$$a^2-2\left(1+\frac{Z_1}{2Z_2}\right)a+1=0 \qquad \ldots(2)$$

The equation determines the attenuation constant a. We shall confine to the case where $\frac{Z_1}{Z_2}$ is real, which corresponds to Z_1 and Z_2 being both pure resistances or pure reactances. Then a can be either real or complex, but not a purely imaginary quantity. The roots will be real when $\frac{Z_1}{4Z_2}$ lies outside the range 0 to -1. We consider separately the three cases where it is greater than 0, between 0 and -1, and less than -1.

(*i*) $\dfrac{Z_1}{4Z_2}>0.$

a is real and positive. Since network does not contain power generating elements, the currents must decrease as we move away

---

*Let $a=e^{\gamma}$: then from this expression

$$\frac{e^{\gamma}+e^{-\gamma}}{2}=1-\frac{Z_1}{2Z_2},\ i.e.\ \cosh\gamma=1+\frac{Z_1}{2Z_2},$$

which is called *Campbell's Equation*.

If $\gamma=\alpha+j\beta$, $a=e^{\alpha+j\beta}$: a will be real if $\beta=0$.

from the generator attached to one end of the filter. The significance of the positive sign of a is that the wave is attenuated without change of phase. If we write $a=e^{-\alpha}$, where α is the attenuation constant per section, equation (2) becomes

$$\cosh \alpha = 1+\frac{Z_1}{2Z_2} \qquad \ldots(3)$$

(*ii*) $0 > \frac{Z_1}{4Z_2} > -1.$

a is a complex with modulus unity, so that we may write $a=e^{-j\beta}$. The wave is not attenuated at all, but suffers change of phase by an angle β in each section, where

$$\cos \beta = 1+\frac{Z_1}{2Z_2}. \qquad \ldots(4)$$

(*iii*) $\frac{Z_1}{4Z_2} < -1.$

a is then real and negative, so that the wave is attenuated with a phase change of π in successive sections. If we write $a=-e^{-\alpha}$, equation (1) becomes

$$-\cosh \alpha = 1+\frac{Z_1}{2Z_2}. \qquad \ldots(5)$$

Since we are dealing with electric wave filters, the currents must be represented by a time varying function. If the current in one section alternates at a given frequency, so that we can write

$$i_n = I_n e^{j\omega t},$$

then the current in a latter section $(n+m)$ will be

$$i_{n+m} = a^m I_n e^{j\omega t},$$

If we write $a=e^{-h}$. the current in the section $(n+m)$ will be

$$i_{n+m} = I_n e^{-hm} e^{j\omega t}$$
$$= I_n e^{(j\omega t - hm)},$$

which is similar to the expression for a wave motion in a continuous medium except that m, the number of sections which can only be an integer, replaces 'x', the distance travelled in the medium. If h is real, we write it as α and we have an attenuated wave, but if h is a pure imaginary quantity $j\beta$, we have an attenuated wave with a phase change β per section. Values of α and β are given in previous three cases. The third case where the attenuated wave changes sign in successive section would not arise in a continuous medium. The behaviour of the filter is determined immediately from the value of $(Z_1/4Z_2)$. It lies within the range 0 to 1, we have a 'pass band' with no attenuation: outside this range we have a 'stop band'.

The uniform chain of filter can be regarded as made up of similar sections joined together, but two types of section can be obtained by cutting the chain in different ways. These two types

are shown in figs. 2 and 3, named as π section and *T* section filters respectively.

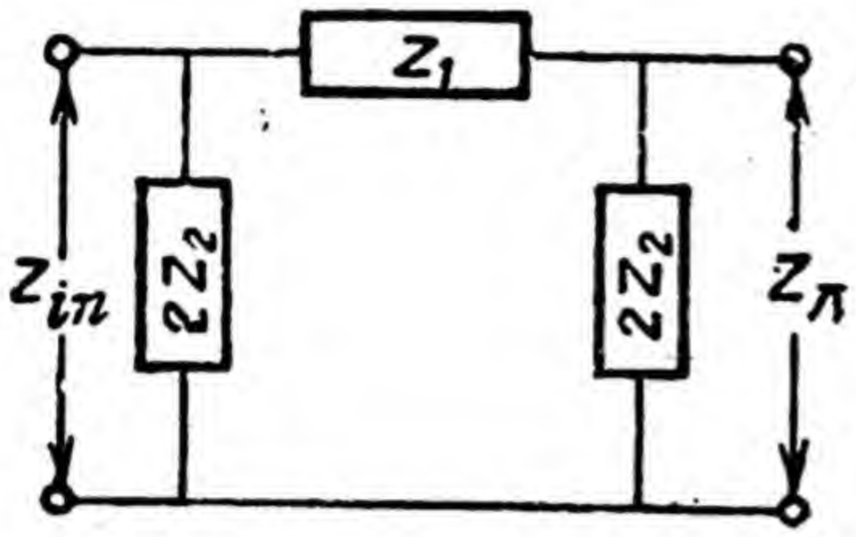


Fig. 2. π-section Filter

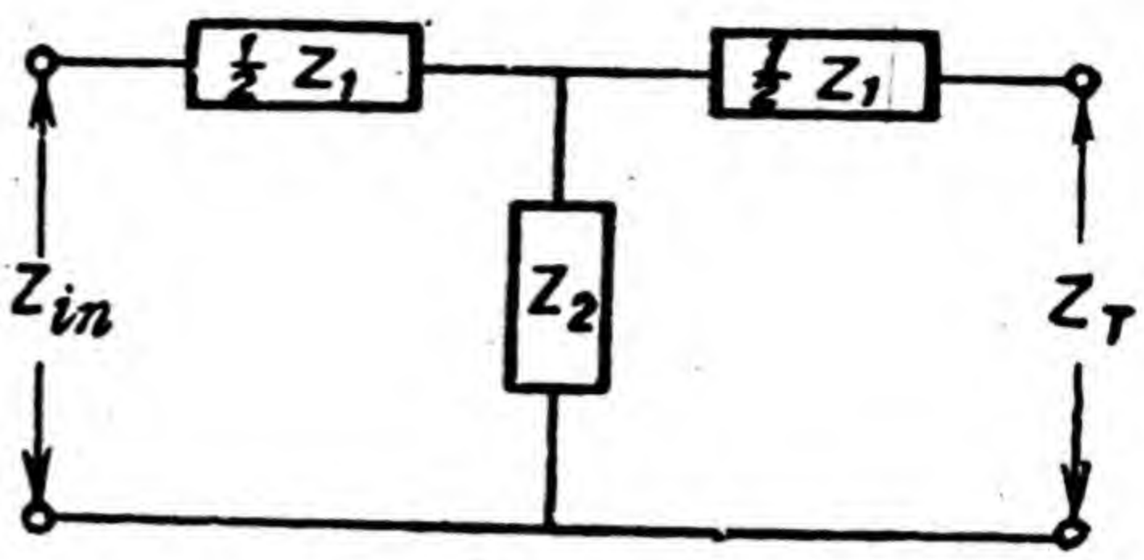


Fig. 3. *T* section Filter.

The series impedance of a symmetrical *T* section is composed of two similar units, each of which is labelled at $\frac{Z_1}{2}$ and shunt impedance is Z_2 as labelled in figure 3. The shunt impedance of symmetrical π section is composed of two equal branches, each labelled $2Z_2$ in fig. 2 and these shunt branches are located on either side of the series impedance Z_1.

4·2 CHARACTERISTIC IMPEDANCE OF T AND π SECTIONS :

(*i*) *T Section.* If the output terminals of *T* section are closed through an impedance Z_T, the impedance across the input terminals is

$$Z_{in}=\frac{Z_1}{2}+\frac{Z_2\left(\frac{Z_1}{2}+Z_T\right)}{Z_2+\frac{Z_1}{2}+Z_T} \qquad ...(1)$$

If $Z_{in}=Z_T=Z_K$, then

$$Z_K=\frac{Z_1}{2}+\frac{Z_2\left(\frac{Z_1}{2}+Z_K\right)}{Z_2+\frac{Z_1}{2}+Z_K}$$

or $$Z_K\left(Z_2+\frac{Z_1}{2}+Z_K\right)=\frac{Z_1}{2}\left(Z_2+\frac{Z_1}{2}+Z_K\right)+Z_2\left(\frac{Z_1}{2}+Z_K\right)$$

$$Z^2_K=Z_1Z_2+\frac{Z_1^2}{4};$$

$$\therefore \quad Z_K=\sqrt{\left(Z_1Z_2+\frac{Z_1^2}{4}\right)}. \qquad ...(2)$$

Z_K is termed as *characteristic* or *iterative impedance* of *T*-section.

(*ii*) *π Section.* If the output terminals of the π section are closed through an impedance Z_π, the impedance across the input terminals is

$$Z_{in}=\frac{2Z_2\left(Z_1+\frac{2Z_2Z_\pi}{2Z_2+Z_\pi}\right)}{2Z_2+Z_1+\frac{2Z_2Z_\pi}{2Z_2+Z_\pi}}. \qquad ...(3)$$

If $Z_{in}=Z_\pi=Z'_K$,

$$Z'_K=\frac{2Z_2\left(Z_1+\frac{2Z_2Z'_K}{2Z_2+Z'_K}\right)}{2Z_2+Z_1+\frac{2Z_2Z'_K}{2Z_2+Z'_K}}$$

or $$Z'_K\left(2Z_2+Z_1+\frac{2Z_2Z'_K}{2Z_2+Z'_K}\right)=2Z_2\left(Z_1+\frac{2Z_2Z'_K}{2Z_2+Z'_K}\right)$$

or $$2Z'_KZ_2(2Z_2+Z'_K)+Z_1Z'_K(2Z_2+Z'_K)+2Z_2Z'_K$$

or $$=2Z_1Z_2(2Z_2+Z'_K)\ 2^2Z_2^2Z'_K$$

or $$(Z_1+4Z_2)\ Z'^2_K=4Z_1Z_2^2.$$

$$Z'_K=\sqrt{\left\{\frac{Z_1Z_2}{(1+Z_1/4Z_2)}\right\}}. \qquad ...(4)$$

Z'_K is called the *iterative impedance* of π section.

Equations (2) and (4) are important relations in filter theory because they define the *characteristic impedance* Z_K and Z'_K in terms of the series and shunt elements that constitute the T and π sections. If a filter section is terminated in its characteristic impedance, the impedance across the input terminals is the same as the receiving end impedance. It can be shown from equations (2) and (4) that

$$Z_KZ'_K=Z_1Z_2. \qquad ...(5)$$

4·3. CHARACTERISTIC IMPEDANCE AS A FUNCTION OF OPEN CIRCUIT AND SHORT CIRCUIT IMPEDANCES :

Let us first consider a T section filter. If Z_0 denotes the open circuit impedance of T section then

$$Z_0=\frac{Z_1}{2}+Z_2$$

When the output terminals are short-circuited, the impedance of T section is

$$Z_S=\frac{Z_1}{2}+\frac{\frac{Z_1Z_2}{2}}{\frac{Z_1}{2}+Z_2}=\frac{\frac{Z_1^2}{4}+Z_1Z_2}{\frac{Z_1}{2}+Z_2}$$

The geometric mean of Z_0 and Z_S is

$$\sqrt{(Z_0Z_S)}=\sqrt{\left(\frac{Z_1^2}{4}+Z_1Z_2\right)}$$

$=Z_K$ from equation (2) art. 4·2

$$\therefore \quad Z_K=\sqrt{(Z_0Z_S)} \qquad ...(1)$$

The fact that Z_K is equivalent to the geometric mean of Z_S and Z_0 provides the basis for a simple experimetal method of determining the characteristic impedance of a given section.

For π section, open circuit impedance is

$$Z_0=\frac{2Z_2(Z_1+2Z_2)}{Z_1+4Z_2}.$$

If the output terminals of the π section are short circuited, then

$$Z_S=\frac{2Z_2Z_1}{Z_1+2Z_2}.$$

Geometric mean of Z_0 and Z_S will be

$$\sqrt{(Z_0.Z_S)}=\sqrt{\left\{\frac{Z_1Z_2}{(1+Z_1/4Z_2)}\right\}}=Z'_K \text{ from equation (4) art. 4·2.}$$

Therefore $$Z'_K=(Z_0Z_S)^{1/2}. \qquad ...(2)$$

Equations (1) and (2) indicate that the characteristic impedance of either T or π section is equal to the geometric mean of their respective open and short-circuit impedances.

4·4. LOW PASS FILTER :

The simplest type of low pass filter is shown in fig. 5 with condenser in the shunt arms and inductance in the series arms.

Then $Z_1=j\omega L$

and $Z_2=1/j\omega C.$

giving $Z_1/4Z_2=-\omega^2LC/4.$

The cut off frequency ω is obtained by setting this equal to -1, *i.e.* $\omega_0^2LC/4=1$

$$\therefore \quad \omega_0=2(LC)^{-1/2}. \qquad ...(1)$$

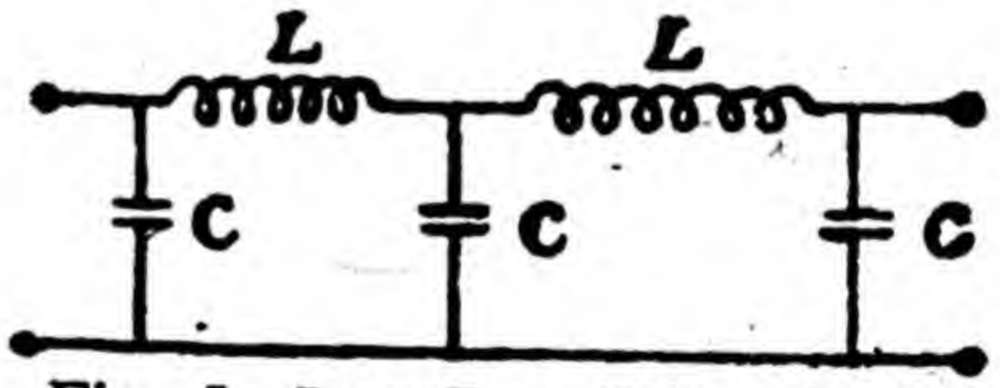


Fig. 5. Low Pass Filter Section

Frequencies below this are passed without attenuation, while higher frequencies are attenuated. In the pass band, the phase change is given by

$$\cos\beta=1-\omega^2LC/2=1-2(\omega/\omega_0)^2, \qquad ...(2)$$

showing that β changes from 0 to π as the frequency changes from zero to the cut off value. The characteristic impedance of a T section is

$$Z_T=\left(\frac{L}{C}-\frac{\omega^2L^2}{4}\right)^{1/2}=\left(\frac{L}{C}\right)^{1/2}\left(1-\frac{\omega^2}{\omega_c^2}\right)^{1/2}.$$

which is resistive in the pass band, but varies from $(L/C)^{1/2}$ to zero at the cut off frequency. In the stop band Z_T is a pure reactance and the attenuation is given by

$$-\cosh\alpha=1-\frac{\omega^2LC}{2},$$

which rises with frequency from zero at $\omega=\omega_0$. The negative sign shows that currents in successive sections are reversed in direction

From relation (5) art. 4·2, the variation of Z_π can be found because $Z_\pi . Z_T = Z_K Z'_K = Z_1 Z_2$,

$$\therefore \quad Z_\pi = \frac{Z_1 Z_2}{Z_T} = \frac{L}{C Z_T}.$$

Variations of α, β and Z_T are shown in figure 6. It should be noted that product of series arm and shunt impedance is constant, *i.e.*, $Z_1 Z_2 = L/C = K^2$ (constant),

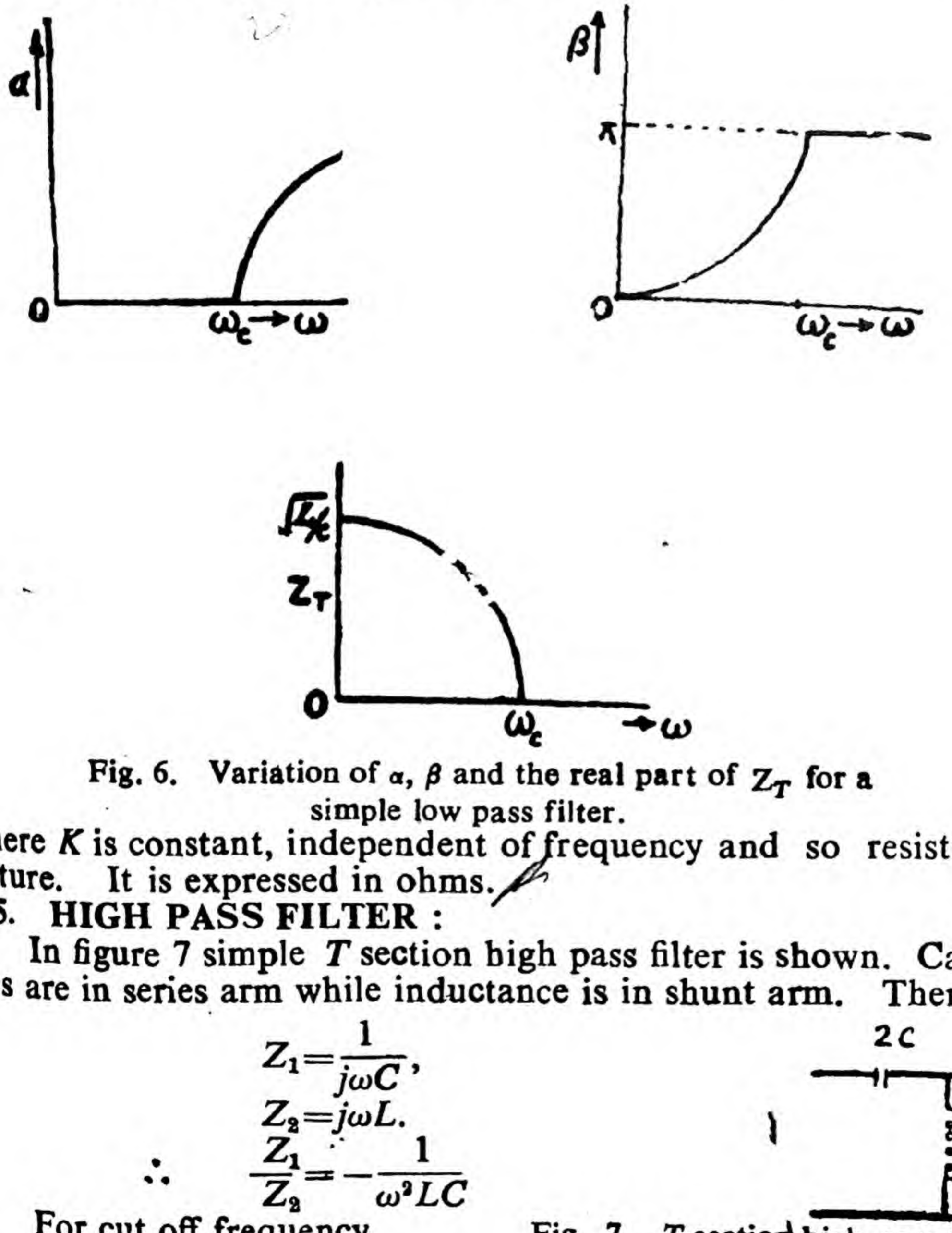


Fig. 6. Variation of α, β and the real part of Z_T for a simple low pass filter.

where K is constant, independent of frequency and so resistive in nature. It is expressed in ohms.

4·5. HIGH PASS FILTER :

In figure 7 simple T section high pass filter is shown. Capacities are in series arm while inductance is in shunt arm. Then

$$Z_1 = \frac{1}{j\omega C},$$

$$Z_2 = j\omega L.$$

$$\therefore \quad \frac{Z_1}{Z_2} = -\frac{1}{\omega^2 LC}$$

For cut off frequency

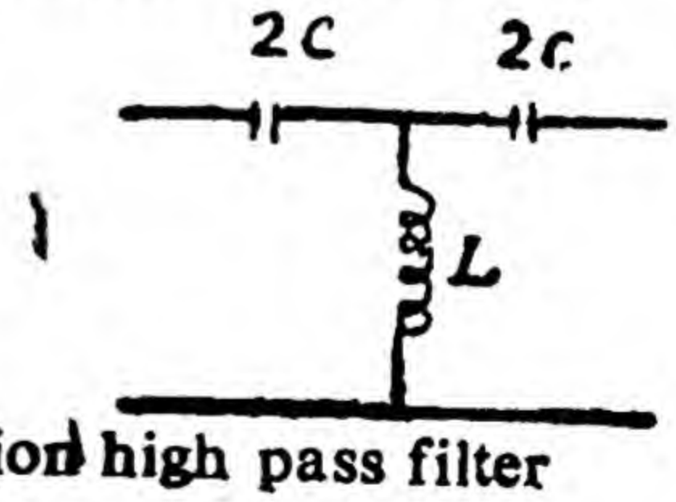


Fig. 7. T-section high pass filter

$$\frac{Z_1}{4Z_2} = -\frac{1}{4\omega^2 LC} = -1, 0.$$

$$\omega_c = \frac{1}{2\sqrt{(LC)}}, 0$$

The phase ehange per section in the passband is given by

$$\cos\beta = 1 - \frac{1}{2\omega^2 LC} = 1 - 2\left(\frac{\omega_o}{\omega}\right)^2$$

and the modulus of this expression gives also the value of cosh α. in the stopband.

The iterative impedance is

$$Z_T = \left(\frac{L}{C} - \frac{1}{4\omega^2 C^2}\right)^{1/2}$$

$$= \left(\frac{L}{C}\right)^{1/2} \left(1 - \frac{\omega_o^2}{\omega^2}\right)^{1/2},$$

which is imaginary in the stopband and rises from zero at the critical frequency to a limiting value of $(L/C)^{1/2}$ in the pass band. The behaviour of α, β and Z_T is illustrated in fig. 8. In this type of filter we again see that

$$Z_1 Z_2 = \frac{L}{C} = K^2 \text{ (constant).}$$

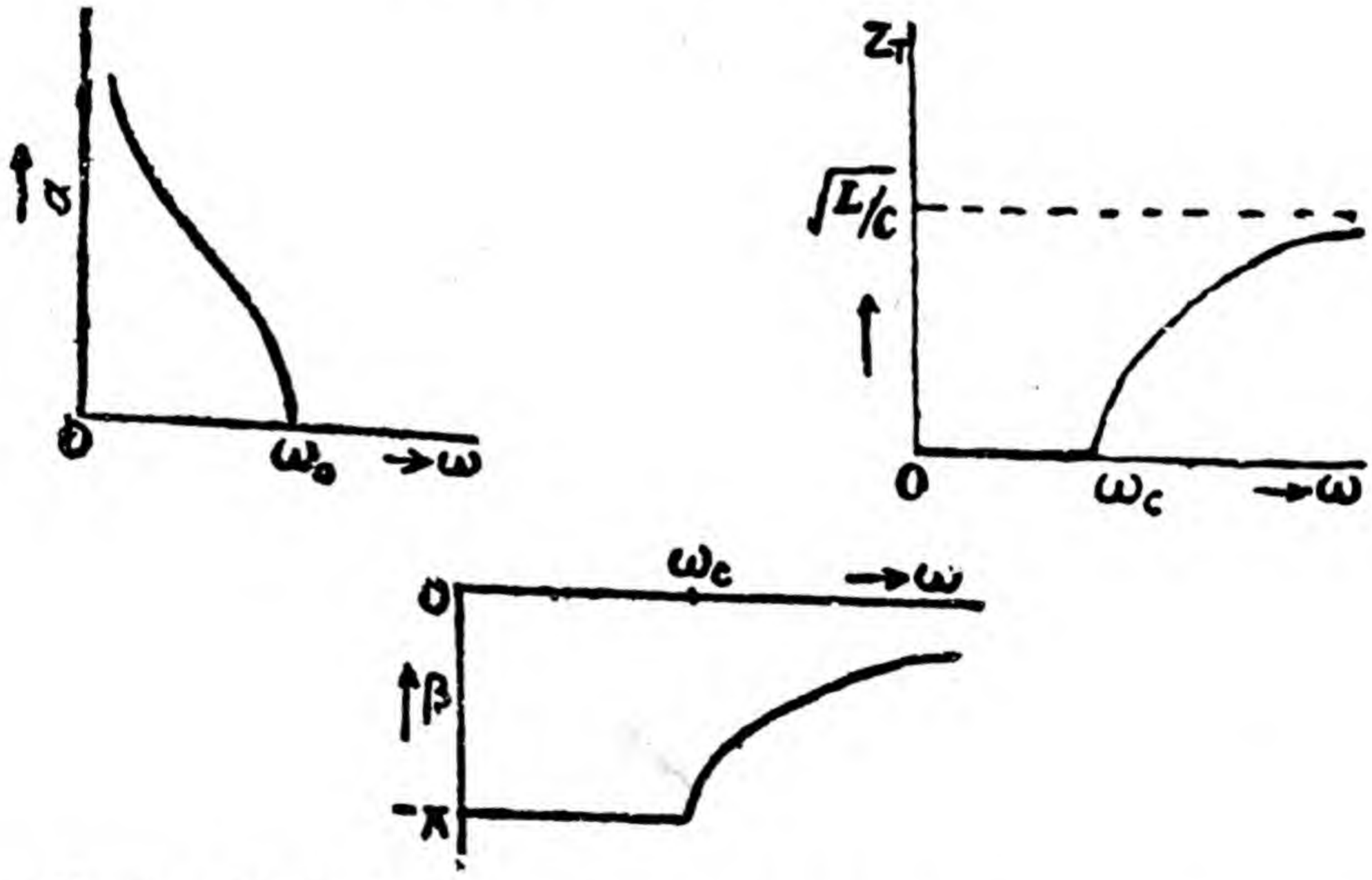


Fig. 8. Illustrating the variation of α, β and real part of Z_T.

4·6. BAND PASS FILTER :

A simple type of band pass filter is shown in figure 9. Its behaviour can be seen as follows : At very low frequencies the impedance in the series will be dominated by the capacity so that the section will act as a simple capacity type attenuator. At frequencies higher than the resonant frequency of the combination L and C_1, $\omega L > (\omega C_1)^{-1}$ and the impedance of the series arm is inductive; then the section behaves as a low pass filter, the highest frequencies again being stopped.

Quantitatively, the analysis is

$$Z_1 = j\omega L + \frac{1}{j\omega C_1}.$$

$$Z_2=\frac{1}{j\omega C_2},$$

$$\frac{Z_1}{4Z_2}=\frac{1}{4}\left[\frac{C_2}{C_1}-\omega^2 LC_2\right],$$

which is positive at zero frequency but tends to $-\infty$ as $\omega\to\infty$. When it lies between 0 and -1, we have a pass band whose lowest frequency is

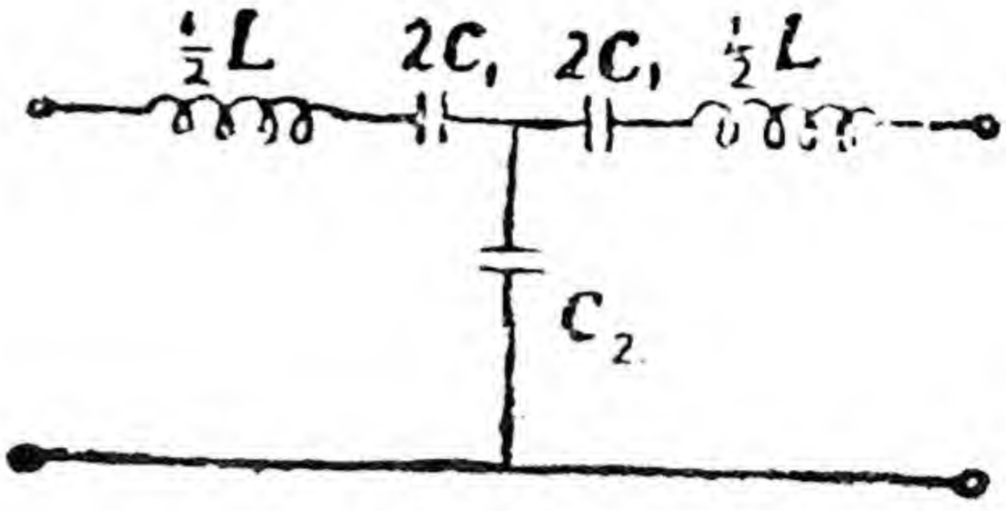


Fig. 9.

$$\omega_1=\frac{1}{\sqrt{(LC_1)}} \text{ where } \frac{Z_1}{4Z_2}=0,$$

and highest frequency

$$\omega=\left[\frac{C_2+4C_1}{LC_2C_1}\right]^{1/2} \text{ where } \frac{Z_1}{4Z_2}=-1.$$

The behaviour of α, β and Z_T is shown in fig. 10

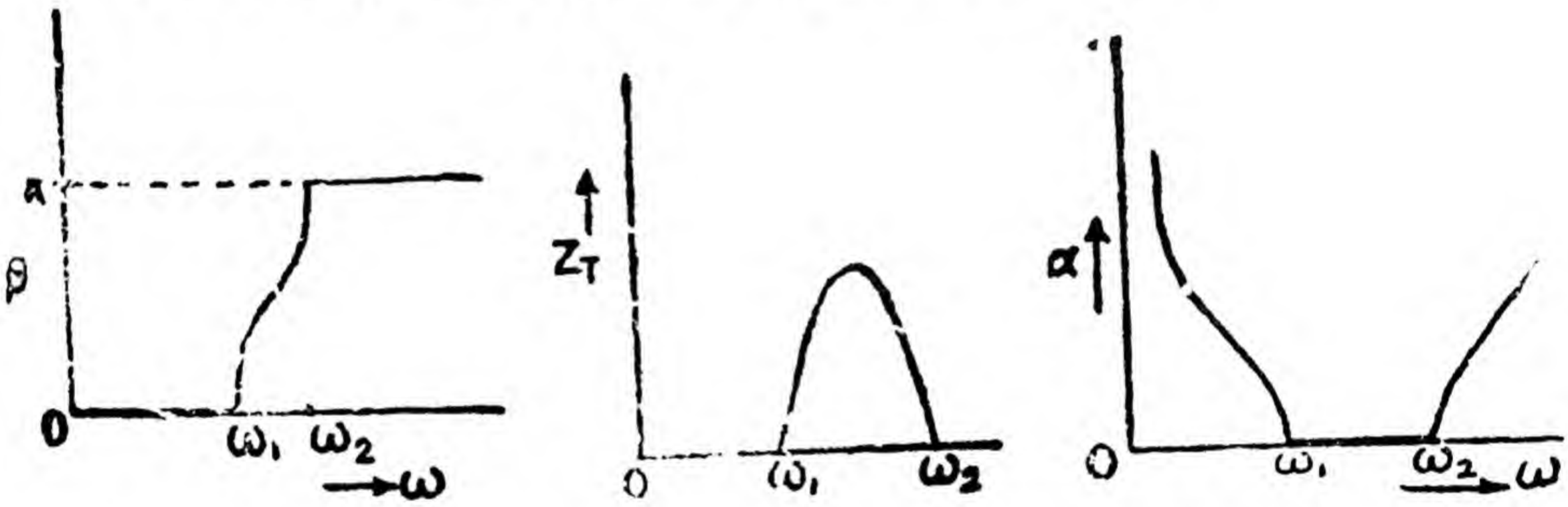


Fig. 10. Representing the behaviour of β. real of Z_T and α.

Now we shall describe a second type of band pass filter in which

$$\frac{L_2}{C_1}=\frac{L_1}{C_2}=K^2 \text{ (constant)}.$$

The T section of such band pass filter is shown in fig. 1. The series arm contains a series resonant circuit while the shunt arm contains a parallel resonant. If the resonance in series arm occurs at a frequency at which the shunt arm of parallel circuit is resonant, the section will have band pass characterisic.

If Z_1 and Z_2 repsesent the whole series and shunt arm impedance, then

$$Z_1=j\left(\omega L_1-\frac{1}{\omega C_1}\right)$$

$$=j\left(\frac{\omega^2 L_1C_1-1}{\omega C_1}\right)$$

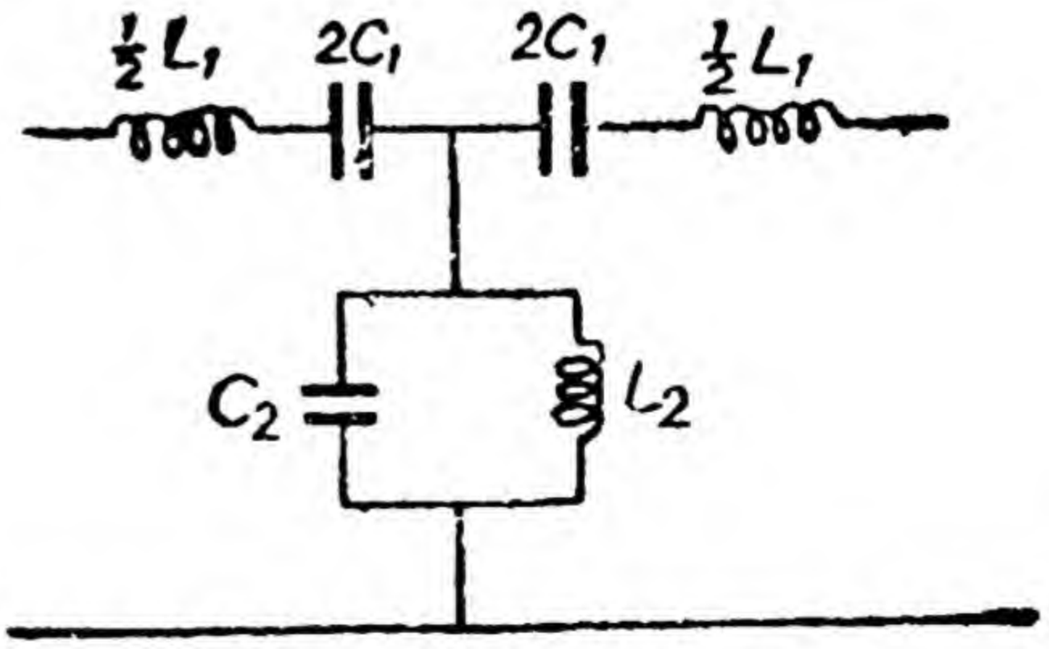


Fig. 11. T-section band pass filter.

$$Z_2=\frac{L_2/C_2}{j\left(\omega L_2-\frac{1}{\omega C_2}\right)}=\frac{\omega L_2}{j(\omega^2 L_2C_2-1)}$$

$$\therefore \quad Z_1Z_2=\frac{L_2}{C_1}\left[\frac{\omega^2L_1C_1-1}{\omega^2L_2C_2-1}\right]=\frac{L_2}{C_1} \text{ if } L_1C_1=L_2C_2.$$

Now we calculate

$$\frac{Z_1}{Z_2}=-\frac{(\omega^2L_1C_1-1)(\omega^2L_2C_2-1)}{\omega^2L_2C_1}.$$

Since $\omega_0^2=\frac{1}{L_1C_1}=\frac{1}{L_2C_2}$,

$$\therefore \quad \frac{Z_1}{Z_2}=-\frac{\left(\frac{\omega^2}{\omega_0^2}-1\right)^2}{\omega^2L_2C_1}$$

For pass band $\frac{Z_1}{4Z_2}=0$ or -1.

Putting $\frac{Z_1}{4Z_2}=0$, we get $\omega=\omega_0$ which is not the critical frequency.

If we put $\frac{Z_1}{4Z_2}=-1$, we get

$$\frac{\left(\frac{\omega^2}{\omega_0^2}-1\right)^2}{4\omega^2L_2C_1}=1$$

or

$$\frac{\omega^4}{\omega_0^4}-2\frac{\omega^2}{\omega_0^2}+1=4\omega^2C_1L_2$$

$$\frac{\omega^4}{\omega_0^4}-2\frac{\omega^2}{\omega_0^2}+1=4\frac{\omega^2}{\omega_0^2}\frac{C_1}{C_2}$$

$$\omega^4-2\omega^2\omega_0^2\left(1+\frac{2C_1}{C_2}\right)+\omega_0^4=0.$$

This equation gives the critical frequencies :

$$\omega^2=\omega_0^2\left(1+\frac{2C_1}{C_2}\right)\pm\sqrt{\left\{\omega_0^2\left(1+\frac{2C_1}{C_2}\right)^2-\omega_0^4\right\}}$$

$$=\omega_0^2\left[\left(1+\frac{C_1}{C_2}\right)+\frac{C_1}{C_2}\pm2\sqrt{\left\{\frac{C_1}{C_2}\left(1+\frac{C_1}{C_2}\right)\right\}}\right]$$

$$\therefore \quad \omega=\omega_0\left[\sqrt{\left(1+\frac{C_1}{C_2}\right)}\pm\sqrt{\left(\frac{C_1}{C_2}\right)}\right].$$

If the roots of this equation be ω_1 and ω_2, then

$$\omega_1=\omega_0[\sqrt{(1+C_1/C_2)}+\sqrt{(C_1/C_2)}].$$

$$\omega_2=\omega_0[\sqrt{(1+C_1/C_2)}-\sqrt{(C_1/C_2)}]. \quad \ldots(1)$$

Multiplying above equations, we get $\quad \ldots(2)$

$$\omega_1\omega_2=\omega_0^2.$$

Hence resonant frequency is the geometric mean of two cut off frequencies. The cut off angular frequencies are given by equations (1) and (2).

The *iterative impedances* can be calculated by putting values of Z_1 and Z_2 in the relation

$$Z_T = \sqrt{\left\{ Z_1 Z_2 \left(1 + \frac{Z_1}{4Z_2} \right) \right\}}.$$

Putting values of Z_1 and Z_2,

$$Z_T = \sqrt{\left\{ \frac{L_2}{C_1} \left(1 - \frac{(\omega^2/\omega_0^2 - 1)^2}{4\omega^2 L_2 C_1} \right) \right\}}$$

$$= \sqrt{\left(\frac{L_2}{C_1} \right)} \sqrt{\left\{ 1 - \frac{(\omega^2 - \omega_0^2)^2}{4\omega^2 \omega_0^2 . C_1/C_2} \right\}}$$

$$= \left(\frac{L_2}{C_1} \right) \sqrt{\left\{ 1 - \frac{(\omega/\omega_0 - \omega_0/\omega)^2}{4C_1/C_2} \right\}}.$$

It can be shown from equations (1) and (2) that

$$\left(\frac{\omega_1}{\omega_0} - \frac{\omega_0}{\omega_1} \right)^2 = \left(\frac{\omega_2}{\omega_0} - \frac{\omega_0}{\omega_2} \right)^2 = 4 \frac{C_1}{C_2}.$$

Therefore expression of Z_T becomes

$$Z_T = \sqrt{\left(\frac{L_2}{C_1} \right)} . \sqrt{\left\{ 1 - \frac{(\omega/\omega_0 - \omega_0/\omega)^2}{(\omega_1/\omega_0 - \omega_0/\omega_1)^2} \right\}}.$$

$$\text{also } Z_T = \left(\frac{L_2}{C_1} \right) \sqrt{\left\{ 1 - \frac{(\omega/\omega_0 - \omega_0/\omega)^2}{(\omega_2/\omega_0 - \omega_0/\omega_2)^2} \right\}}.$$

Therefore when $\omega = \omega_1$, $Z_T = 0$

$\omega = \omega_2$, $Z_T = 0$

$\omega = \omega_0$, $Z_T = \sqrt{(L_2/C_1)} = K$

The variations of attenuation constant α and Z_T with frequency are shown in figures (12) and (13).

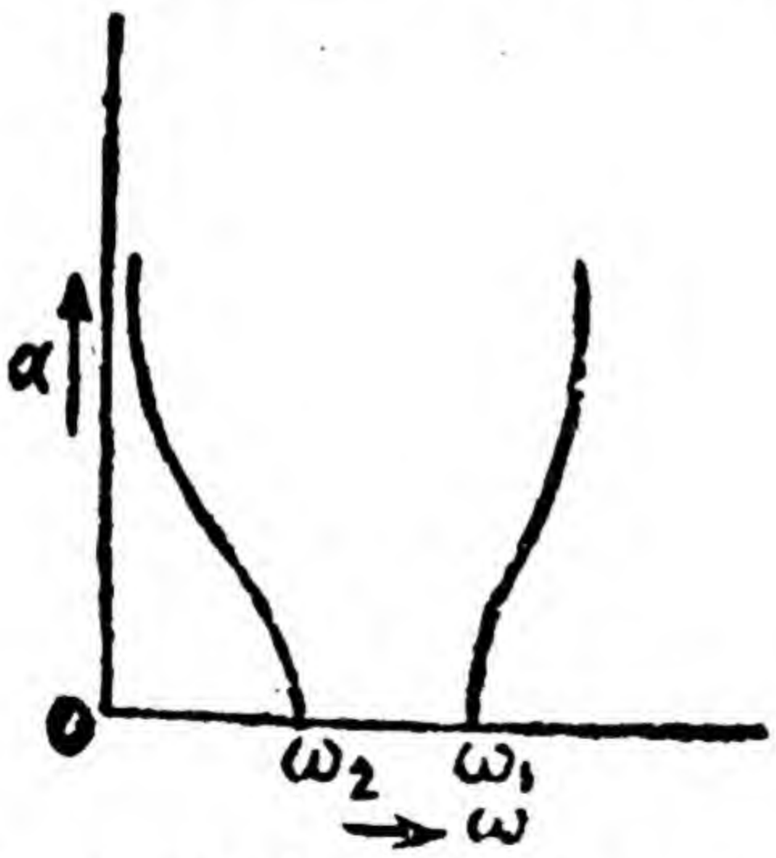


Fig. 12. Variation of attenuation constant.

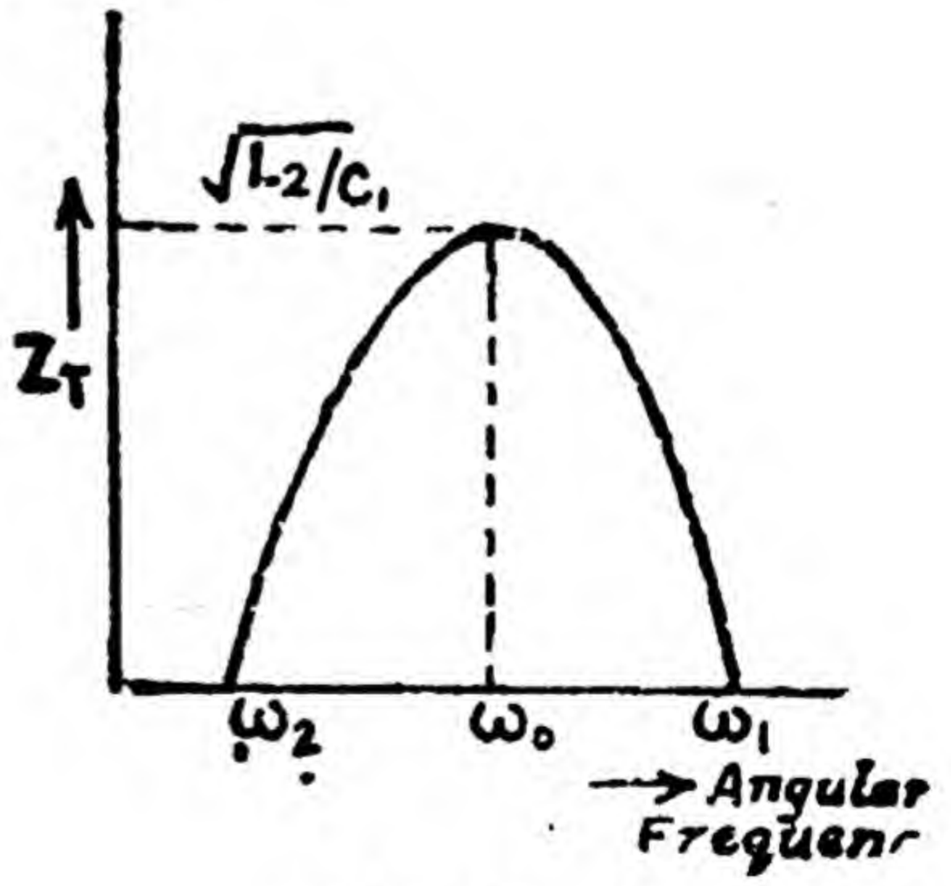


Fig. 13. Variation of impedance, Z_T.

Element values for design : The values of the elements may be determined in terms of two cutoff frequencies and design impedance K.

We have shown previously that

$$\omega_1\omega_2=\omega_0^2=\frac{1}{L_1C_1}=\frac{1}{L_2C_2}. \qquad ...(3)$$

$$K^2=\frac{L_1}{C_1}=\frac{L_1}{C_2}$$

From equations (1) and (2), we have

$$\omega_1-\omega_2=\omega_0.2\sqrt{\left(\frac{C_1}{C_2}\right)}=\frac{1}{\sqrt{(L_1C_1)}}.2.\sqrt{\left(\frac{C_1}{C_2}\right)}$$

$$=\frac{2}{L_1}\sqrt{\left(\frac{L_1}{C_1}\right)}=\frac{2K}{L_1}$$

or $$L_1=\frac{K}{\omega_1-\omega_2}=\frac{K}{\pi.(f_1-f_2)}$$

From (3), we see that

$$C_1=\frac{1}{L_1\omega_1\omega_2}=\frac{\pi\,(f_1-f_2)}{K\omega_1\omega_2}=\frac{\pi\,(f_1-f_2)}{K.4\pi f_1 f_2}$$

$$=\frac{(f_1-f_2)}{4\pi K f_1 f_2}$$

From relation (4), values of L_2 and C_2 can be determined

i.e. $$L_2=K^2.C_1=\frac{K\,(f_1-f_2)}{\pi f_1 f_2},$$

$$C_2=\frac{1}{\pi K\,(f_1-f_2)}$$

Thus the values of L_1, L_2, C_1 and C_2 can be calculated in terms of K, f_1 and f_2.

4·7. BAND-ELIMINATION FILTER :

Band stop filters are networks which are designed to pass all frequency except those in a specified band.

A T section band stop filter is shown in fig. 14. The series arm consist of a parallel resonant circuit and shunt arm consists of series resonant circuit.

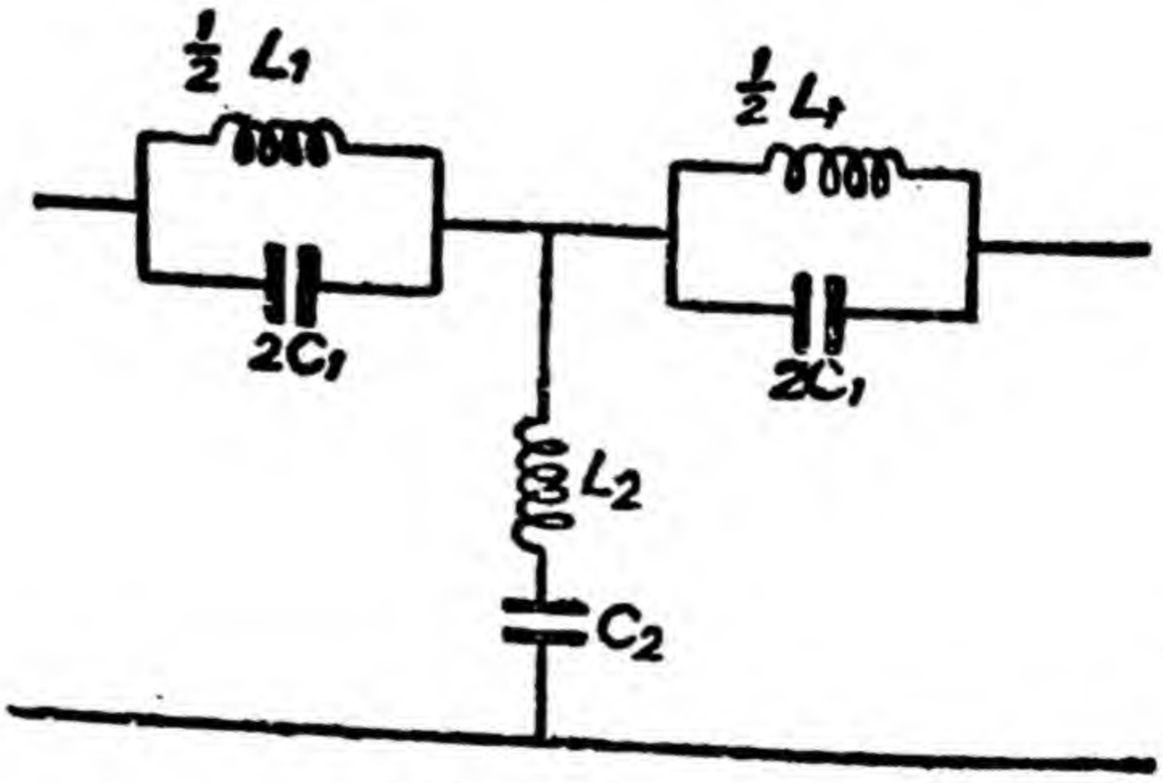


Fig. 14.
T section band stop filter.

If Z_1 represents total series arm impedance and Z_2, total shunt arm impedance then

$$Z_1=\frac{j\omega L_1\,\frac{1}{j\omega C_1}}{j\omega L_1+\frac{1}{j\omega C_1}}=\frac{\omega L_1}{j\,(\omega^2L_1C_1-1)}$$

$$Z_2=j\omega L_2+\frac{1}{j\omega C_2}=j\left[\frac{\omega^2L_2C_2-1}{\omega C_2}\right].$$

Now $Z_1Z_2=Z_2/C_1$ (If $L_2C_2=L_2C_1$),

The arrangement in this type of filter is also same as band pass filters, *i.e.*

$$Z_1.Z_2=\frac{L_2}{C_1}=\frac{L_1}{C_2}=K^2 \text{ (constant).}$$

From the condition of passband $\frac{Z_1}{4Z_2}=0$ or -1.

From $\frac{Z_1}{4Z_2}=0$, we get

$$-\frac{\omega^2L_1C_2}{(\omega^2L_1C_1-1)(\omega^2L_2C_2-1)}=0$$

$$(\omega^2L_1C_1-1)(\omega^2L_2C_2-1)=\infty$$

From $\frac{Z_1}{4Z_2}=-1$, we get

$$\frac{\omega^2L_1C_2}{(\omega^2L_1C_1-1)(\omega^2L_2C_2-1)}=4.$$

Putting $L_1C_1=L_2C_2=\frac{1}{\omega_0^2}$,

$$\frac{\omega^2L_1C_2}{\left(\frac{\omega^2}{\omega_0^2}-1\right)\left(\frac{\omega^2}{\omega_0^2}-1\right)}=4$$

or $$\omega^2L_1C_2=4\left(\frac{\omega^2}{\omega_0^2}-1\right)^2$$

or $$\frac{\omega^2}{\omega_0^2}\cdot\frac{C_2}{C_1}=4\left(\frac{\omega^4}{\omega_0^4}-2\frac{\omega^2}{\omega_0^2}+1\right)$$

or $$4\omega^4-8\omega^2\omega_0^2-\omega^2\omega_0^2\frac{C_2}{C_1}+4\omega_0^4=0.$$

or $$4\omega^4-\omega^2\left(8\omega_0^2+\omega_0^2\frac{C_2}{C_1}\right)+4\omega_0^2=0$$

The value of ω^2 will be given by

$$\omega^2=\frac{\left(8\omega_0^2+\omega_0^2\frac{C_2}{C_1}\right)\pm\sqrt{\left\{\left(8\omega_0^2+\omega_0^2\frac{C_2}{C_1}\right)^2-64\omega_0^4\right\}}}{8}$$

$$=\frac{\omega_0^2}{8}(8+C_2/C_1)\pm\frac{\sqrt{(C_2/C_1)}.\sqrt{(C_2/C_1+16)}}{8}\times\omega_0^2$$

$$\omega^2=\frac{\omega_0^2}{16}[(16+C_2/C_1)+C_2/C_1\pm2\sqrt{(C_2/C_1)}.\sqrt{(16+C_2/C_1)}]$$

$$=\frac{\omega_0^2}{16}[\sqrt{(16+C_2/C_1)}\pm\sqrt{(C_2/C_1)}]^2.$$

$$\therefore\quad \omega=\frac{\omega_0}{4}[\sqrt{(16+C_2/C_1)}\pm\sqrt{(C_2/C_1)}].$$

Let the two values of ω be represented by

$$\omega_1=\frac{\omega_0}{4}\left[\sqrt{(C_2/C_1+16)}+\sqrt{(C_2/C_1)}\right] \quad \ldots(1)$$

$$\omega_2=\frac{\omega_0}{4}\left[\sqrt{(C_2/C_1+16)}-\sqrt{(C_2/C_1)}\right]. \quad \ldots(2)$$

If we multiply both the above equations, then we get

$$\omega_1\omega_2=\omega_0^2.$$

Thus the resonant frequency is the geometric mean of two cut off frequencies given by equations (1) and (2).

Variation of attenuation constant with frequency is shown in figure 15. Thus the frequencies lying between $(\omega_1-\omega_2)$ frequency band are stopped while those lying outside this band are allowed to pass.

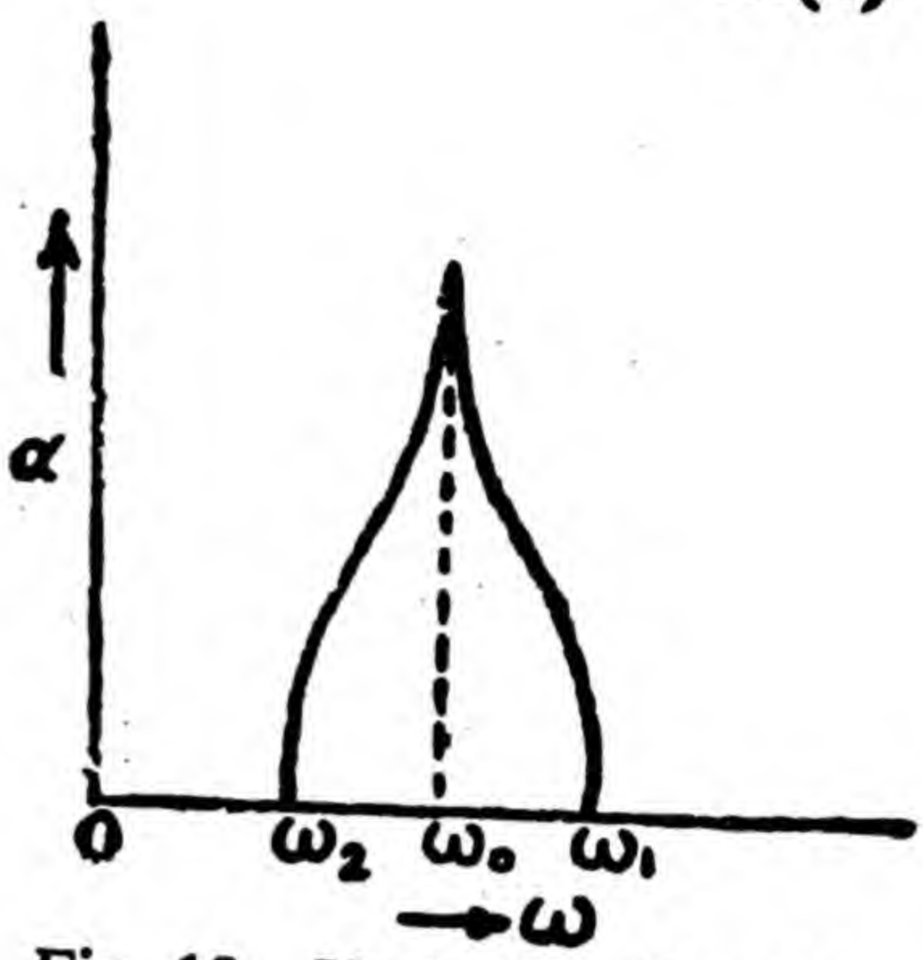


Fig. 15. Variation of α with angular frequency ω

Iterative Impedance of T section :

$$Z_T=\sqrt{\left\{Z_1Z_2\left(1+\frac{Z_1}{4Z_2}\right)\right\}}=K\sqrt{\left(1+\frac{Z_1}{4Z_2}\right)}.$$

Putting $\frac{Z_1}{Z_2}=-\frac{\omega^2L_1C_2}{\left(\frac{\omega^2}{\omega_0^2}-1\right)^2}$, we get

$$Z_T=K\sqrt{\left\{1-\frac{\omega_0^2L_1C}{4(\omega^2/\omega_0^2-1)^2}\right\}}=K\sqrt{\left\{1-\frac{C_2/C_1}{4(\omega/\omega_0-\omega_0/\omega)^2}\right\}}.$$

It can be shown from equations (1) and (2) that

$$\frac{C_2}{C_1}=4\left(\frac{\omega_1}{\omega_0}-\frac{\omega_0}{\omega_1}\right)=4\left(\frac{\omega_2}{\omega_0}-\frac{\omega_0}{\omega_2}\right)^3.$$

Putting these values in the expression for Z_T, we have

$$Z_T=K\sqrt{\left\{1-\frac{(\omega_1/\omega_0-\omega_0/\omega_1)^2}{(\omega/\omega_0-\omega_0/\omega)^2}\right\}}, \quad \ldots(3)$$

and

$$Z_T=K\sqrt{\left\{1-\frac{(\omega_2/\omega_0-\omega_0/\omega_2)^2}{(\omega/\omega_0-\omega_2/\omega)^2}\right\}}. \quad \ldots(4)$$

Therefore, when

$\omega=0$, $Z_T=K$,

$\omega=\omega_1$, $Z_T=0$,

$\omega=\omega_2$, $Z_T=0$,

$\omega=\infty$, $Z_T=K$,

$\omega=\omega_0$, Z_T is imaginary.

This variation of Z_T with frequency is shown in figure 16.

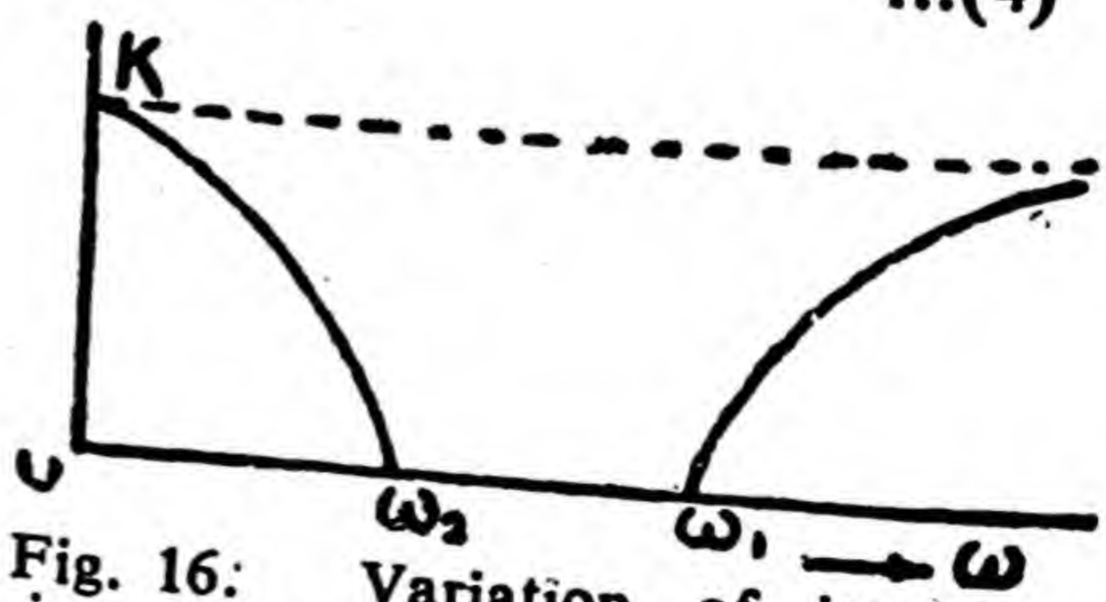


Fig. 16. Variation of iterative impedance with angular frequency.

The values of various elements for bandstop filter can be calculated in a similar manner as for band pass filter.

4·8. DISADVANTAGES OF SIMPLE TYPES OF FILTER :

Filters considered earlier belong to the class of '*K* derived' or 'constant *K*' filters since their impedances obey the relation

$$Z_1 Z_2 = K^2,$$

where *K* is a constant, independent of frequency This constant *K* type filter suffers from two principal disadvantages :

(*i*) Its characteristic impedance is not sufficiently constant over the pass band but varies with frequency. The filter, therefore, not be terminated correctly throughout the pass band.

(*ii*) The attenuation does not rise very abruptly at the boundary to the transmission band.

In order to overcome the inherent limitations of constant *K* type, Zobel devised a filter section which he called the *m*-derived type filter Such types of filter give practically uniform characteristic impedance over a large part of the pass band and at the same time increase the abruptness with which cut off occurs.

A high degree of attenuation beyond the cut off or a constant impedance in the pass band demands a more complicated structure. If the constant *K* section is regarded as the prototype, it is possible to design a section to have the same impedance and hence the same pass and attenuation bands but with a different degree of attenuation outside the pass band. This may be done as shown in the following articles :

4·9. m-DERIVED T-SECTION (Series derived filter):

Suppose that *T* section, as shown in figure 17 has the series arm modified by some constant *m*. Then if this new section is to have the same impedance Z_T as the prototype, the shunt impedance must be modified in the same way.

For the prototype section,

$$Z_T = \sqrt{\left\{Z_1 Z_2 \left(1 + \frac{Z_1}{4Z_2}\right)\right\}}$$

If for this new section $Z_1' = mZ_1$ and impedance is Z_T', then

$$Z_T' = \sqrt{\left\{mZ_1 Z_2' \left(1 + \frac{mZ_1}{4Z_2''}\right)\right\}},$$

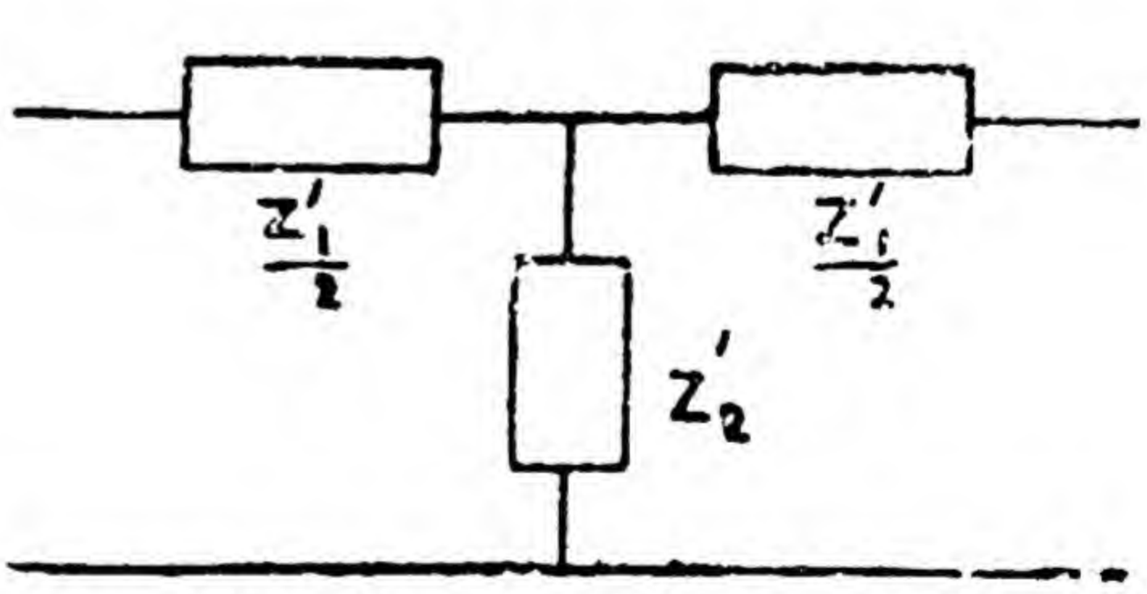


Fig. 17. Modified prototype *T* section.

where Z_2' is the modified shunt impedance in the new section.

If Z_T' is equal to Z_{T1} then Z_2' must have the value determined by

$$Z_1 Z_2 \left(1 + \frac{Z_1}{4Z_2}\right) = mZ_1 Z_2' \left(1 + \frac{mZ_1}{4Z_2'}\right),$$

i.e.
$$Z_2' = \frac{Z_2}{m} + \left(\frac{1-m^2}{4m}\right).Z_1 \qquad ...(1)$$

and so the derived T section as shown in fig. 18 has the same impedance as the prototype T section, m must always be chosen such that

$$0 < m < 1.$$

Now we shall discuss the pass band limit in this type of filter.

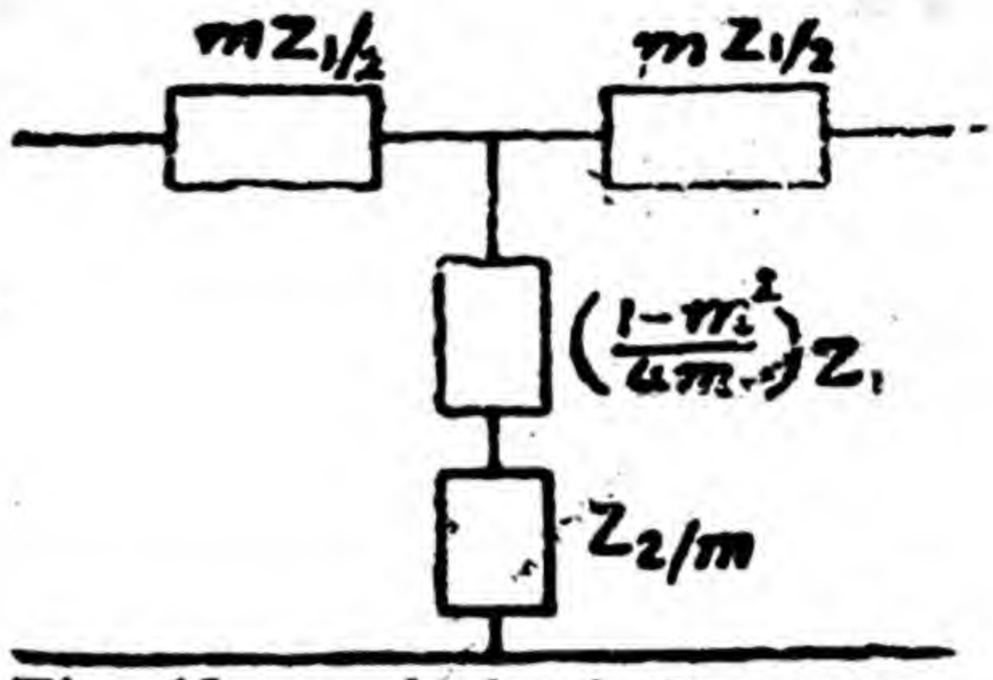


Fig. 18. m-derived T section

Pass band limit : For pass band in prototype T section, we know that $\frac{Z_1}{4Z_2} = 0$ or -1. Considering m derived T section, we write

$$\frac{Z_1'}{4Z_2'} = -\frac{mZ_1}{4\left[\frac{Z_2}{m} + \frac{1-m^2}{4m} Z_1\right]}$$

$$= \frac{mZ_1/Z_2}{4\left[\frac{1}{m} + \frac{1-m^2}{m} \cdot \frac{Z_1}{4Z_2}\right]}.$$

If $\frac{Z_1}{4Z_2} = 0$, then $\frac{Z_1'}{4Z_2'} = 0$.

Also if $\frac{Z_1}{4Z_2} = -1$, then $\frac{Z_1'}{4Z_2'} = -\frac{4m}{4\left[\frac{1}{m} - \frac{1-m^2}{m}\right]} = -1.$

Thus pass band in m derived T section will be determined by

and
$$\left.\begin{aligned} \frac{Z_1'}{4Z_2'} &= 0, \\ \frac{Z_1'}{4Z_2'} &= -1. \end{aligned}\right\} \qquad ..(2)$$

Since the pass band limit in the case of m-derived T section is same as in the case of constant K type, the critical frequency of prototype and m-derive d filter is the same.

The shunt arm is to be chosen in such a way that it is resonant at some frequency of infinite or high attenuation called f_∞ above f_0. This means that at the resonant frequency

$$\left|\frac{Z_2}{m}\right| = \left|\frac{1-m^2}{4m} Z_1\right|,$$

and for *low pass filter*

$$\frac{1}{2\pi f_\infty mC} = \frac{1-m^2}{4m} 2\pi f_\infty L,$$

or $$f_\infty=\frac{1}{\pi\sqrt{\{(1-m^2)\,LC\}}}. \qquad \ldots(3)$$

We know that the cut off frequency for the low-pass filter is

$$f_c=\frac{1}{\pi\sqrt{(LC)}} \qquad \ldots(4)$$

so that the frequency of infinite attenuation will be

$$f_\infty=\frac{f_c}{\sqrt{(1-m^2)}}.$$

from which $$m=\sqrt{\{1-(f_c/f_\infty)^2\}}.$$

This equation determines the value of m to be used for a particular f_∞.

Similar relations for the *high pass* filter can derived; they are

$$f_\infty=f_c\sqrt{(1-m^2)} \qquad \ldots(5)$$

and

$$m=\sqrt{\{1-(f_\infty/f_c)^2\}} \qquad \ldots(6)$$

The m derived section is designed by following the design of the prototype T section.

The variation of attenuation over the attenuation band for *low pass m derived* section in the stop band is dependent on the sign of reactances or

$$\alpha=2\cosh^{-1}\sqrt{\left(\left|\frac{Z_1}{4Z_2}\right|\right)} \quad \text{or} \quad \alpha=2\sinh^{-1}\sqrt{\left(\left|\frac{Z_1}{4Z_2}\right|\right)}$$

$$f_c<f<f_\infty \qquad\qquad f_\infty<f$$

For $Z_1=j\omega L$ and $Z_2=-j\omega C$ for the prototype, then

$$\left|\frac{Z_1}{4Z_2}\right|=\frac{m\omega L}{4[1/m\omega C-\omega L(1-m^2)/4m]}$$

so that for $f_c < f < f_\infty$

$$\alpha=2\cosh^{-1}\frac{mf/f_c}{\sqrt{(1-f^2/f_\infty^2)}}$$

and for $f_\infty < f$,

$$\alpha=2\sinh^{-1}\frac{mf/f_\infty}{\sqrt{(f^2/f_\infty^2-1)}}$$

From the above expression α may be determined. Figure 19 shows a plot of α against f/f_c for $m=0{\cdot}6$. It is observed that f_∞ is equal to 1·25 times the cut off frequency f_c. The large increase in sharpness of cut off for the m derived section over the prototype is apparent.

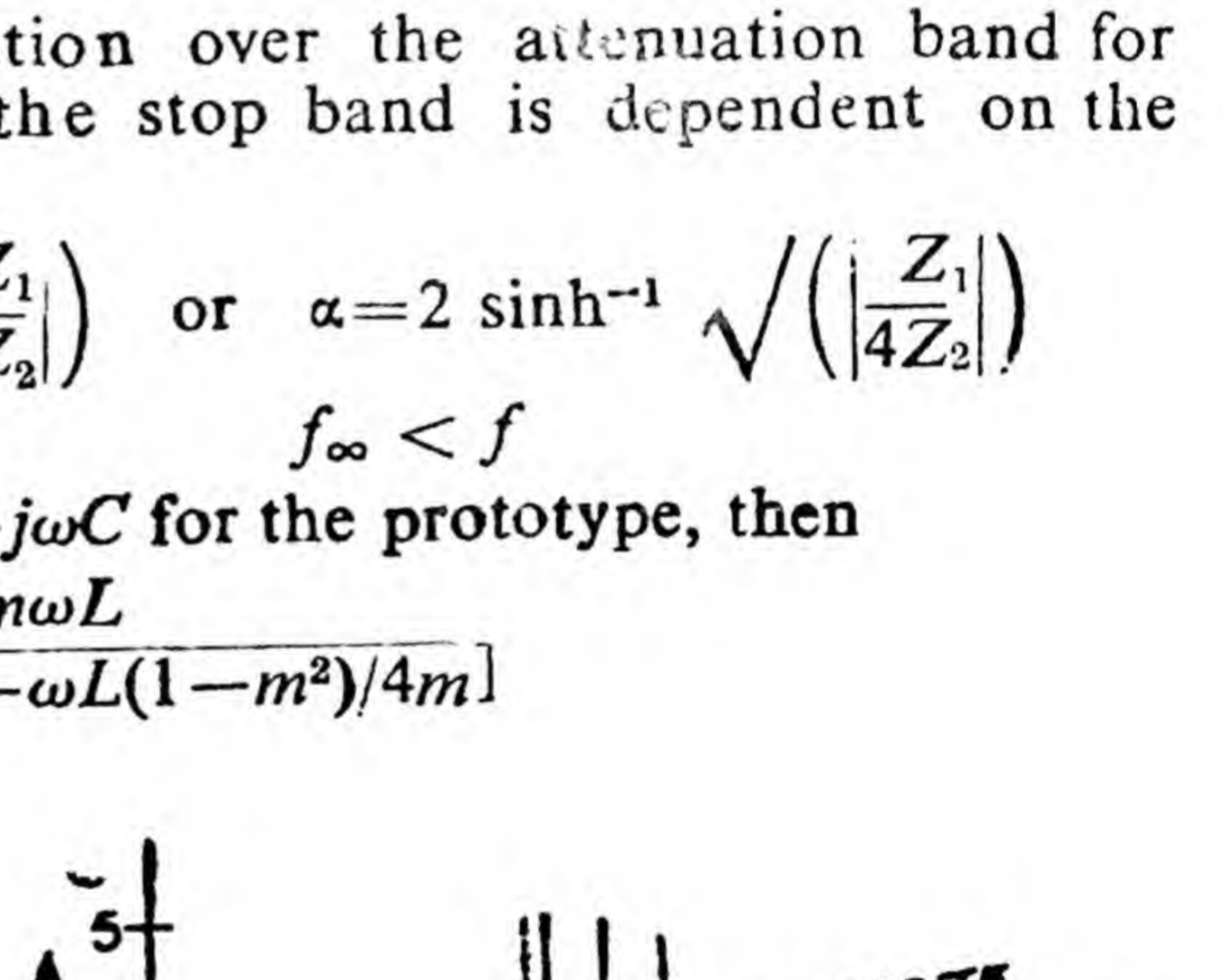


Fig. 19. Variation of attenuation for the prototype and m derived section, and the composite result of the two in series.

The constant β may be

determined in the pass band from

$$\beta = 2 \sin^{-1} \sqrt{\left(\left| \frac{Z_1}{4Z_2} \right| \right)}$$

$$= 2 \sin^{-1} \frac{m.f/f_c}{\sqrt{\{1-(f^2/f_c^2)\,(1-m^2)\}}}. \qquad \ldots(7)$$

In the attenuation band, upto f_∞, β has the value π. Above f_∞ the value of β drops to zero, because the shunt arm becomes inductive above resonance. The phase shift of the m derived section is plotted as a function of f/f_c in the figure 20.

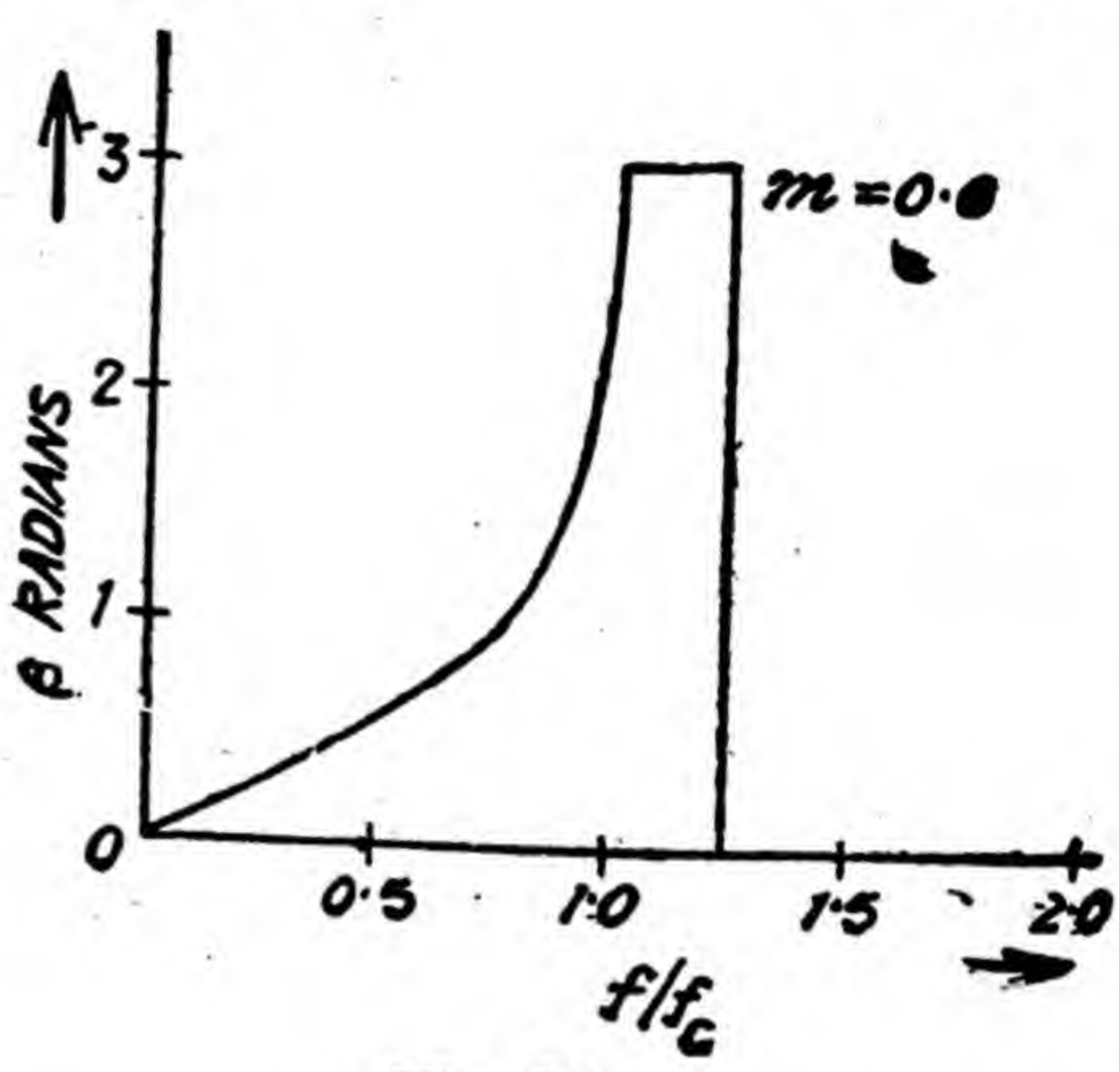


Fig. 20.

The sharpness of cut off increases for small values of m, the attenuation beyond the point of peak attenuation becomes smaller for small m. This emphasizes the necessity of supplementing the m derived section with a prototype section in series to raise attenuation for frequencies well remote from cut off.

4·10. m-DERIVED π SECTION FILTER (Shunt derived filter) :

Here we assume $Z_2' = \frac{Z_2}{m}$ and find the value of Z_1', so that the iterative impedance of m derived π-section is the same as that of a prototype π-section.

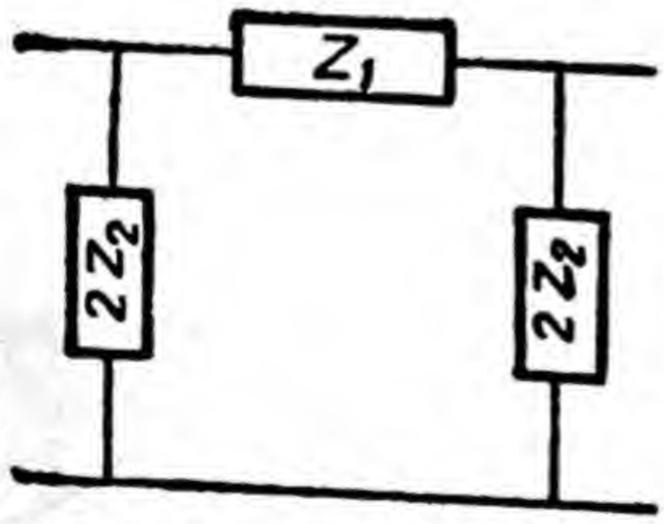


Fig. 21 (*a*) Prototype π section.

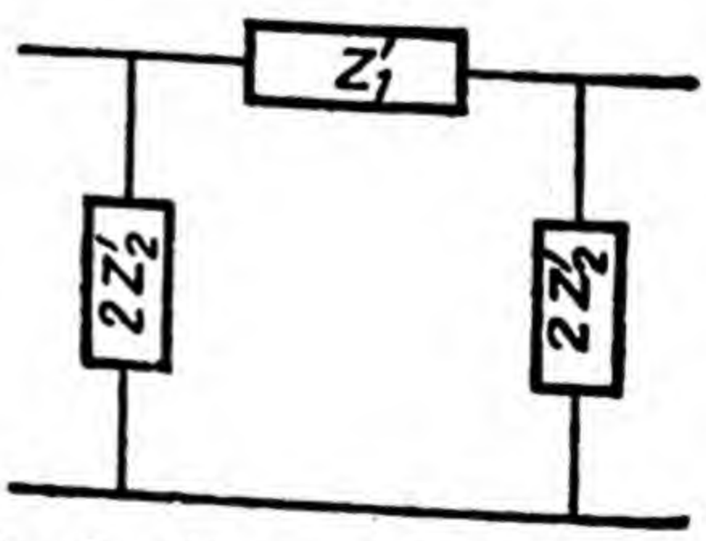


Fig. 21 (*b*) Modified π-section.

Now if the two types of π-section have the same iterative impedance, then

$$Z_\pi = \sqrt{\left(\frac{Z_1 Z_2}{1+Z_1/4Z_2} \right)} = Z_\pi' = \sqrt{\left(\frac{Z_1' Z_2/m}{1+m.Z_1'/4Z_2} \right)}$$

or

$$\frac{Z_1'}{m}\left(1+\frac{Z_1}{4Z_2}\right) = Z_1\,(1+mZ_1'/4Z_2),$$

$$Z_1'\left(\frac{Z_2}{m}+\frac{Z_1}{4m}\right) = Z_1Z_2 + Z_1'\left(\frac{mZ_1}{4}\right),$$

$$Z_1' \left[\frac{Z_2}{m}+\frac{Z_1}{4m}(1-m^2)\right]=Z_1Z_2.$$

$$Z_1'=\frac{mZ_1.4Z_2}{4Z_2+Z_1(1-m^2)}.$$

Multiplying above and below by $\frac{m}{1-m^2}$,

$$Z_1'=\frac{mZ_1\left(\frac{4m}{1-m^2}\right)Z_2}{\frac{4m}{(1-m^2)}.Z_2+mZ_1} \quad \ldots (1)$$

The above expression represents a parallel combination of mZ_1 and $\left(\frac{4m}{1-m^2}\right)Z_2$. Therefore series and shunt arm impedances of filter section can be shown as in fig. 22.

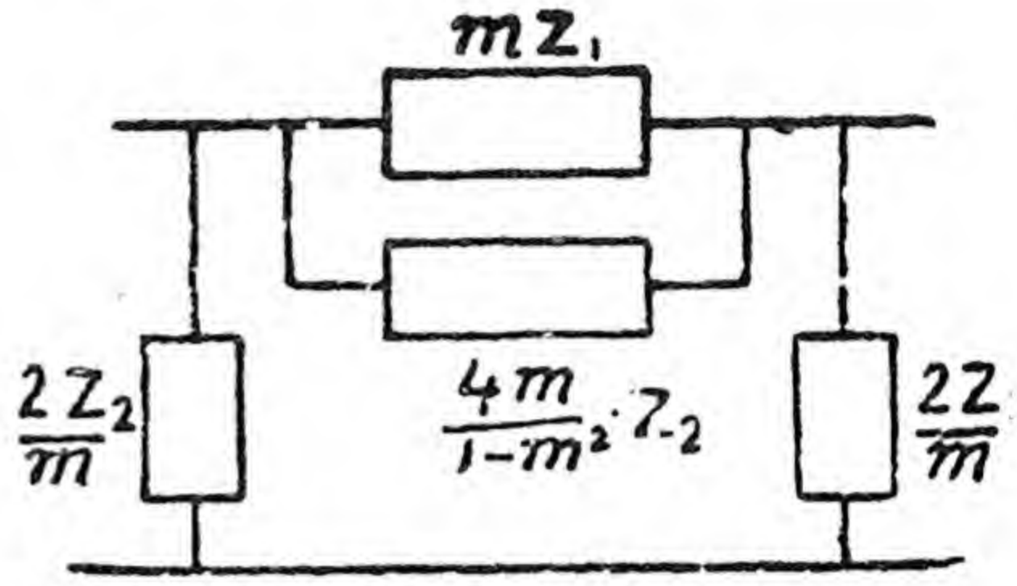


Fig. 22. m-derived π section.

Pass band condition :

To find the pass band condition, we determine the ratio $\frac{Z_1'}{4Z_2'}$.

$$\frac{Z_1'}{4Z_2'}=\frac{1}{\left(\frac{1}{mZ_1}+\frac{1-m^2}{4mZ_2}\right)\Big/\frac{4Z_2}{m}}$$

$$=\frac{m}{4Z_2}\,\frac{2}{\frac{1}{Z_1}\left(\frac{1}{m}+\frac{1-m^2}{m}.\frac{Z_1}{4Z_2}\right)}$$

$$=\frac{m.Z_1/4Z_2}{\frac{1}{m}+\left(\frac{1-m^2}{m}\right)\frac{Z_1}{4Z_2}}.$$

$$\left.\begin{aligned}&\text{If } \frac{Z_1}{4Z_2}=0, \text{ then } \quad \frac{Z_1'}{4Z_2'}=0.\\ &\text{If } \frac{Z_1}{4Z_2}=-1, \text{ then } \quad \frac{Z_1'}{4Z_2'}=-1.\end{aligned}\right\} \quad \ldots(2)$$

We see that the condition of pass band in the case of T section and π section of m derived filter is the same.

Condition of Infinite attenuation : For infinite attenuation $Z_1=\infty$ and $Z_2'=0$. From $Z_1'=\infty$, we get

$$\frac{1}{\frac{1}{mZ_1}+\frac{1-m^2}{4mZ_2}}=\infty,$$

$$\frac{1}{mZ_1}+\frac{1-m^2}{4mZ_2}=0.$$

$$\frac{Z_1}{4Z_2}=\frac{1}{1-m^2}$$

for $-\frac{Z_1}{4Z_2}>1$, m should be less than one.

Therefore the condition of infinite attenuation in the case of T and π sections is the same.

Thus m-derived sections may be employed to give increased attenuation near the cut off point and by proper adjustment of the parameter m they can be made to meet any practical attenuation requirement. m-derived section in conjunction with 'K-derived' section overcomes practically all the disadvantages.

4·11. m-DERIVED BAND PASS FILTERS:

The m-derived band pass network is shown in fig. 23.

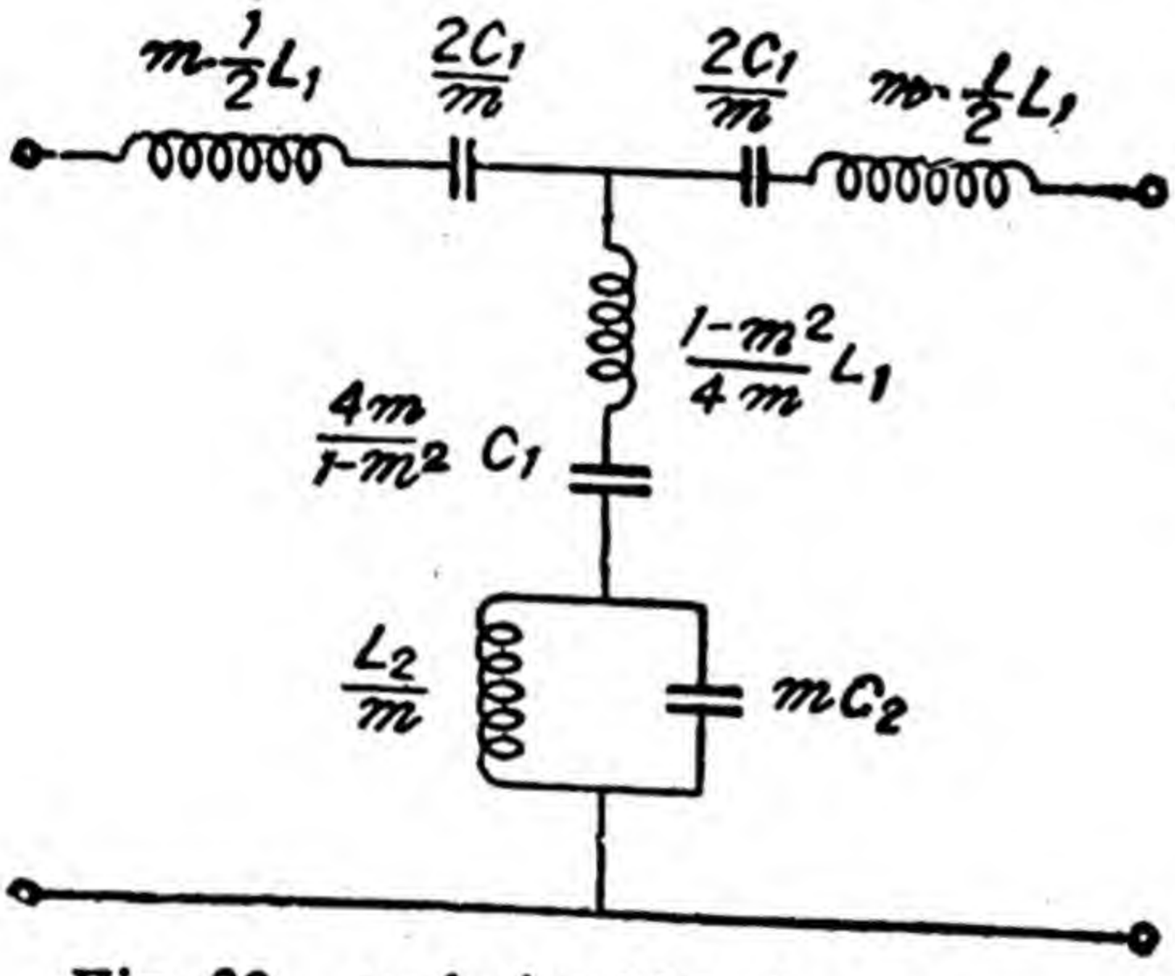


Fig. 23. m-derived band pass section.

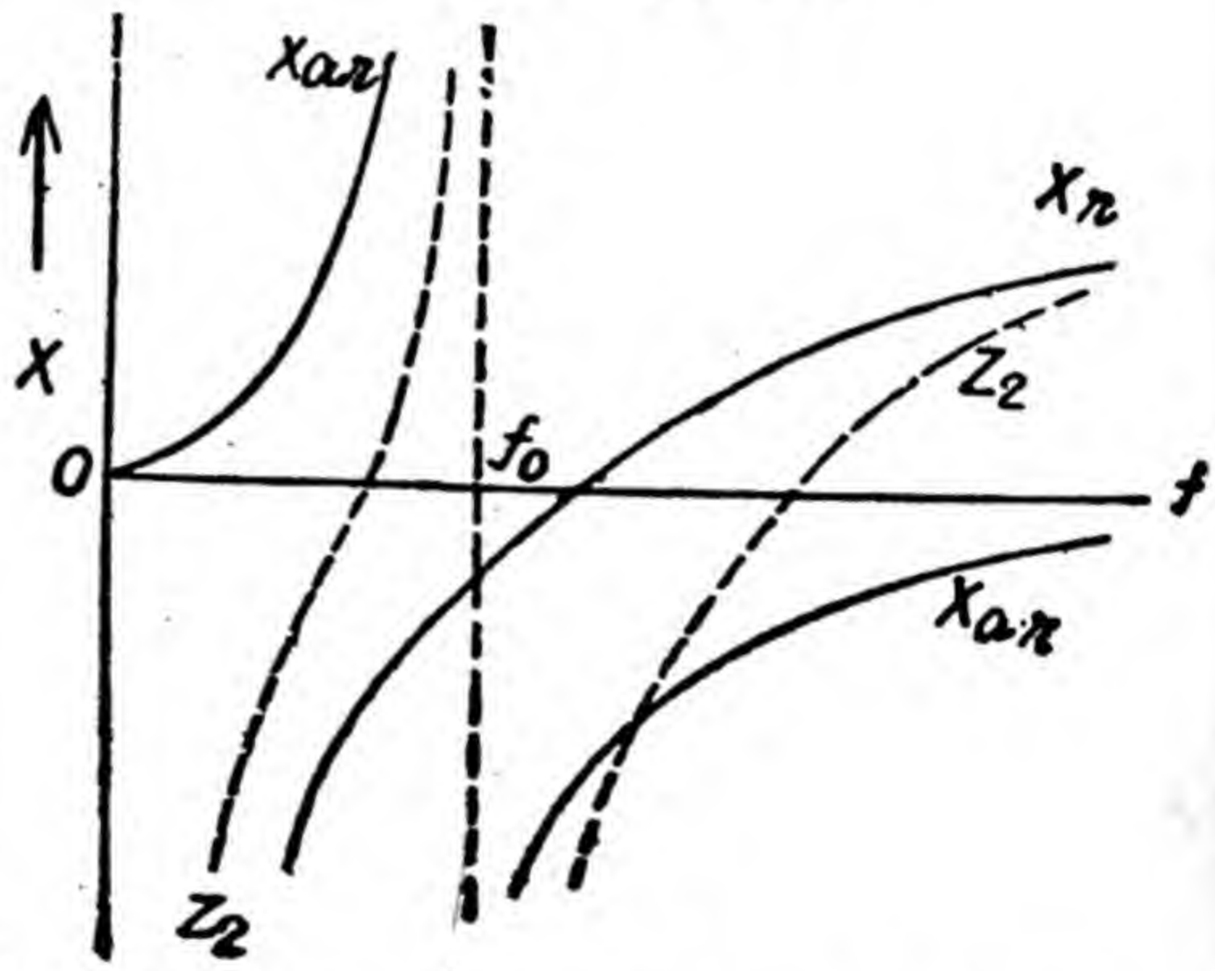


Fig. 24. Reactance curves for shunt arm of the m-derived bandpass section.

The shunt arm consists of series resonant and antiresonant circuits, both in series. Plot of reactance curves for these two circuits, and the reactance variation of the shunt arm, Z_2, of filter is shown in fig. 24.

Here f_0 is antiresonant frequency of the arm as a whole. The reactance curve for Z_2 shows that the shunt arm is resonant at a frequency above f_0 and against at a frequency below f_0. At these frequencies the network is short circuited, and thus they are frequencies of high attenuation f_∞.

At one f_∞, the reactance X_r, and X_{ar} are equal and opposite so that

$$j\omega_\infty\left(\frac{1-m^2}{4m}\right)L_1\frac{j}{\omega_\infty\,[4m/(1-m^2)]\,C_1}=\frac{(j\omega_\infty L_2/m)\,(-j\omega_\infty mC_2)}{j\,(1/\omega_\infty.mC_2-\omega_\infty L_2/m)}$$

$$\frac{1-m^2}{4}\,(\omega_\infty^2L_1C_1-1)=\frac{\omega_\infty^2\,L_2C_1}{\omega_\infty^2L_2C_2-1}. \qquad \text{.. (1)}$$

In view of the fact

$$L_1C_1=L_2C_2=\frac{1}{{\omega_0}^2},$$

equation (1) becomes

$$\frac{1-m^2}{4}\left(\frac{{f_\infty}^2}{{f_0}^2}-1\right)^2=4\pi^2 {f_\infty}^2 L_2C_1. \qquad \ldots (2)$$

The values of L_2 and C_1 have been calculated in K type band pass filter, they are

$$L_2=\frac{K(f_1-f_2)}{4\pi f_1 f_2} \text{ and } C_1=\frac{(f_1-f_2)}{4\pi K f_1 f_2}$$

$$\therefore \quad L_2C_1=\frac{K(f_1-f_2)}{4\pi f_1 f_2}\times\frac{(f_1-f_2)}{4\pi K f_1 f_2}=\frac{(f_1-f_2)^2}{16\pi^2 {f_0}^4}$$

where $$f_0^2=f_1 f_2$$

Substituting the value of L_2C_1 in equation (2), we have

$$\frac{(1-m^2)}{4}\left(\frac{{f_\infty}^2}{{f_0}^2}-1\right)^2=4\pi^2 {f_\infty}^2\times\frac{(f_1-f_2)^2}{16\pi^2 {f_0}^4}$$

or $$(1-m^2)({f_\infty}^2-f_1 f_2)^2={f_\infty}^2(f_1-f_2)^2 \qquad \ldots(3)$$

$${f_\infty}^2-\frac{(f_1-f_2)f_\infty}{\sqrt{(1-m^2)}}-f_1 f_2=0. \qquad \ldots(4)$$

Solving for the values of the frequencies of peak attenuation

$$f_\infty=\frac{(f_1-f_2)}{2\sqrt{(1-m^2)}}\pm\sqrt{\left(\frac{(f_1-f_2)^2}{4(1-m^2)}+f_1 f_2\right)}. \qquad \ldots(5)$$

The first factor of equation (5) on right hand side is less than the second factor hence one value of f_∞ will be negative which has no physical significance. Thus the expression for f_∞ should be reversed so that the two frequencies of peak attenuation are

$$f_{\infty 1}=\sqrt{\left(\frac{(f_1-f_2)^2}{4(1-m^2)}+f_1 f_2\right)}-\frac{(f_1-f_2)}{2\sqrt{(1-m^2)}} \qquad \ldots(6)$$

and

$$f_{\infty 2}=\sqrt{\left(\frac{(f_1-f_2)^2}{4(1-m^2)}+f_1 f_2\right)}+\frac{(f_1-f_2)}{2\sqrt{(1-m^2)}} \qquad \ldots(7)$$

Solving equation (3), we get

$$m=\sqrt{\left\{1-\left[\frac{f_\infty(f_1-f_2)}{{f_\infty}^2-f_1 f_2}\right]^2\right\}}$$

$$=\frac{\sqrt{\{({f_\infty}^2-{f_2}^2)({f_\infty}^2-{f_1}^2)\}}}{{f_\infty}^2-f_1 f_2}. \qquad \ldots(8)$$

The value of m may be chosen to place either one of the two frequencies of peak attenuation at a desired point, the other frequency of peak attenuation then being definitely fixed.

$$f_{\infty 1} f_{\infty 2}=\frac{f_1^2+2f_1 f_2+f_2^2-4m^2 f_1 f_2-f_1^2+2f_1 f_2-f_2^2}{4(1-m^2)}=f_1 f_2,$$

$$\sqrt{(f_{\infty 1} f_{\infty 2})}=f_0. \qquad \ldots(9)$$

Thus f_0 is the geometric mean of the frequencies of peak

attenuation. If m is selected to place $f_{\infty 1}$ at a desired point, $f_{\infty 2}$ is automatically fixed and vice versa.

4·12. ATTENUATORS :

Attenuators are resistance networks used for the specific purpose of producing loss, in known amounts, for all frequencies. Since the resistance offered by a resistor to the currents of different frequencies in it is independent of their frequencies (except at very high frequencies at which skin effect comes into action), they are best suited for producing same loss for all frequencies. Such networks can be designed to match two unequal impedances and an attenuation greater than a limiting minimum value.

Let us consider an L-section attenuator and take that R_1 and R_2 are the two impedances to be matched by it.

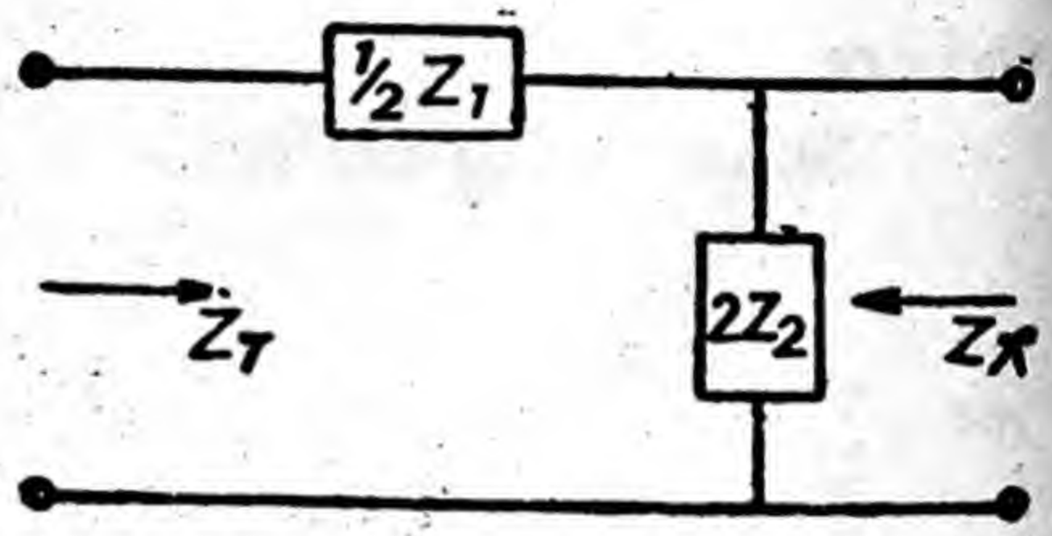


Fig. 25.

We know, from filter theory, that

$$Z_T = \sqrt{(Z_1 Z_2 + \tfrac{1}{4} Z_1^2)},$$

$$Z_\pi = \frac{Z_1 Z_2}{Z_T}.$$

When Z_1 and Z_2 are pure resistances then $Z_T > Z_\pi$. Now if R_1 is greater than R_2, then

$$R_1 = Z_T, \qquad R_2 = Z_\pi.$$

Thus $$Z_1 Z_2 = Z_\pi Z_T = R_1 R_2$$

Putting $Z_T = R_1$ and $Z_1 Z_2 = R_1 R_2$ in the expression for Z_T, we get

$$R_1 = \sqrt{(R_1 R_2 + \tfrac{1}{4} Z_1^2)},$$

$$\therefore \quad \tfrac{1}{4} Z_1^2 = R_1^2 - R_1 R_2.$$

$$\therefore \quad \text{Series arm impedance} = S = \tfrac{1}{2} Z_1$$

$$= \sqrt{[R_1 (R_1 - R_2)]}. \qquad \ldots(1)$$

$$\text{And parallel arm impedance} = P = 2Z_2 = 2 . \frac{R_1 R_2}{Z_1}, \qquad \ldots(2)$$

$$P = \frac{R_1 R_2}{\tfrac{1}{2} Z_1} = \frac{R_1 R_2}{S}.$$

Thus, simply from R_1 and R_2, the shunt arm 'P' and series arm S can be calculated from equations (1) and (2).

Now the loss produced by the L-section will be

$$\frac{\text{Input power}}{\text{Output power}} = \frac{E_1 I_1}{E_2 I_2} = e^{2\gamma}.$$

$$\text{Thus, Loss} = e^{2\gamma} = 1 + \frac{Z_1}{2Z_2} + \sqrt{\left\{\frac{Z_1}{Z_2} + \left(\frac{Z_1}{2Z_2}\right)^2\right\}},$$

(from equation 12 art. 4·3).

Since $\frac{Z_1}{Z_2}=\frac{Z_1^2}{Z_1Z_2}=\frac{4(R_1^2-R_1R_2)}{R_1R_2}=4\left(\frac{R_1}{R_2}-1\right)$,

hence

$$\text{Loss}=1+2\left(\frac{R_1}{R_2}-1\right)+\sqrt{\left\{4\left(\frac{R_1}{R_2}-1\right)+4\left(\frac{R_1^2}{R_2^2}+1-\frac{2R_1}{R_2}\right)\right\}}$$

$$=\frac{2R_1}{R_2}-1+2\sqrt{\left\{\frac{R_1}{R_2}\left(\frac{R_1}{R_2}-1\right)\right\}}. \qquad \ldots(4)$$

Because L section has only two adjustable elements, it can not match two importances and produce required loss, though in matching R_1 and R_2, attenuation produced is minimum.

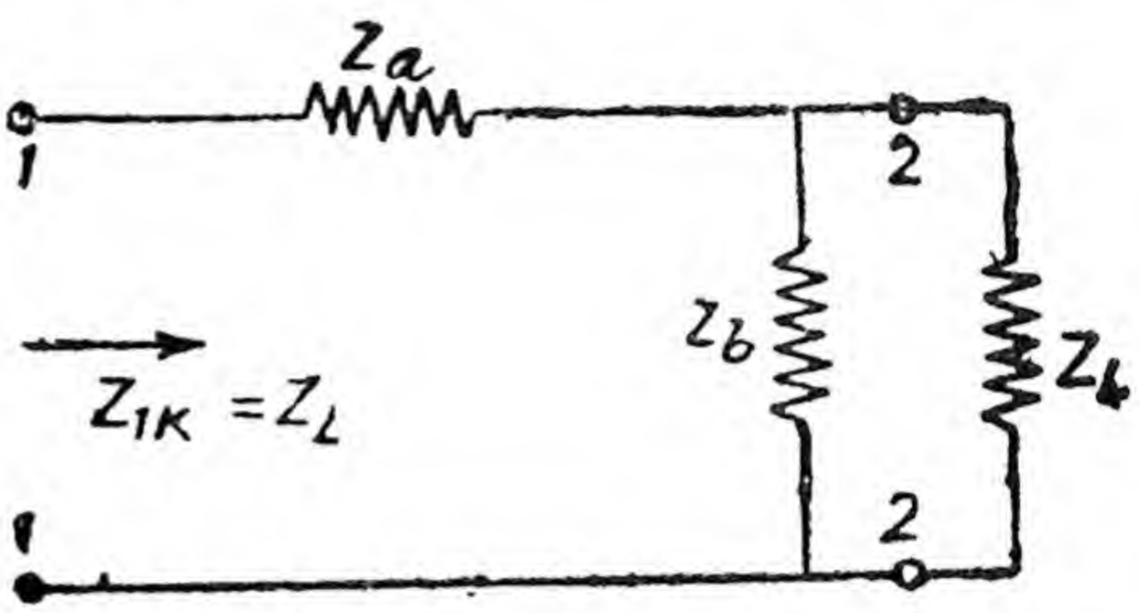


Fig. 26.

To produce any required loss by L-section iterative impedance termination is adopted. To secure required values, in this method, generally input impedance and attenuation are adjusted.

From equation (15), art. 3·0 iterative transfer constant p is given by

$$e^p=\frac{Z_b+Z_{1K}}{Z_b}$$

or
$$Z_b=\frac{Z_{1K}}{e^p-1}=\frac{Z_L}{e^p-1}\cdot \qquad \ldots(5)$$

Also
$$Z_{1K}=Z_a+\frac{Z_bZ_{1K}}{Z_b+Z_{1K}}$$

or
$$Z^2_{1K}-Z_{1K}.Z_a-Z_aZ_b=0$$

or
$$Z_a\,(Z_{1K}+Z_b)=Z^2_{1K}.$$

Putting
$$Z_{1K}+Z_b=Z_b.\epsilon^p$$

and $Z^2_{1K}=Z_L^2$ in above relation, we get $Z_a.Z_be^p=Z_L^2$.

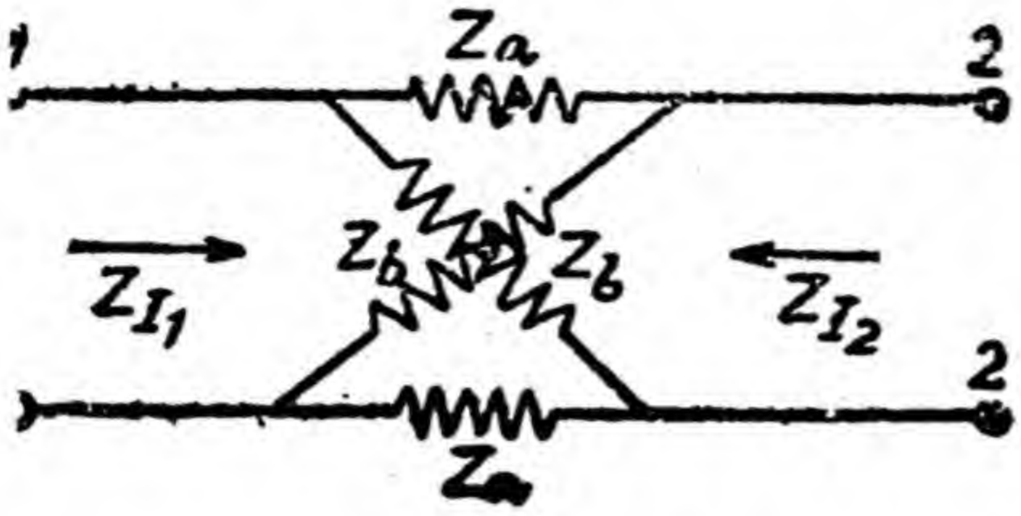
But $Z_b=\frac{Z_L}{e^p-1}$ from (5), therefore

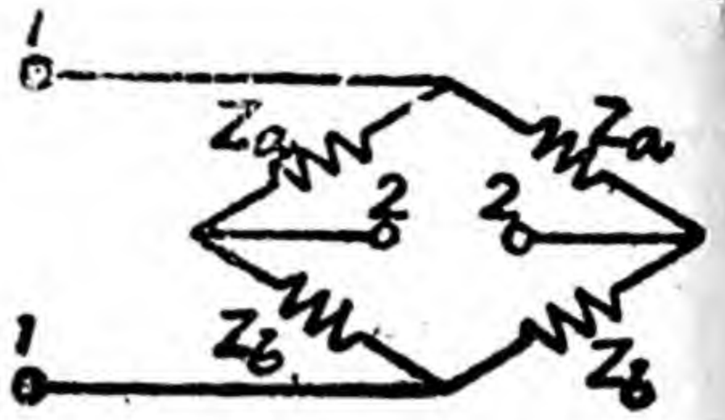
$$Z_a.\frac{Z_L}{e^p-1}.e^p=Z_L^2 \qquad \text{or} \quad Z_a=Z_L\,(1-e^{-p}) \qquad \ldots(6)$$

From equations (5) and (6) $Z_a.Z_b$ can be calculated and input impedance is also Z_L so that we can easily design this attenuator.

Lattice type attenuator. A lattice type attenuator is shown in figure 27. It is a symmetrical balanced four terminal network composed of two pairs of impedances arranged as shown in the

figure 27 (*a*). A lattice is like a bridge with input applied at one diagonal and output taken from other diagonal fig. 27 (*b*).





$$Z_{I_1}=Z_{I_2}-Z_I.$$

Fig. 27 (*a*). Fig. 27 (*b*).

If terminals 2-2 are open, then the resultant impedance Z_0 will be

$$\frac{1}{Z_0}=\frac{1}{Z_a+Z_b}+\frac{1}{Z_a+Z_b}$$

$$\therefore\ Z_0=\frac{Z_a+Z_b}{2}.$$

While with 2-2 shorted, resultant impedance Z_s will be

$$\frac{1}{Z_s}=\frac{1}{2}\left(\frac{1}{Z_a}+\frac{1}{Z_b}\right)$$

$$\therefore\ Z_s=\frac{2Z_aZ_b}{Z_a+Z_b}.$$

Now image impedances

$$Z_{I_1}=Z_{I_2}=Z_I=\sqrt{(Z_0Z_s)}$$

$$=\sqrt{\left(\frac{Z_a+Z_b}{2}\cdot\frac{2Z_aZ_b}{Z_a+Z_b}\right)}$$

$$\therefore\qquad Z_I=\sqrt{(Z_aZ_b)}. \qquad \ldots(7)$$

And image transfer constant γ is given by

$$\tanh^2\gamma=\frac{Z_s}{Z_0}=\frac{4Z_aZ_b}{(Z_a+Z_b)^2},$$

from which it can be seen that

$$\cosh\gamma=\frac{Z_a+Z_b}{Z_b-Z_a}$$

and

$$\sinh\gamma=\sqrt{\left\{\frac{4Z_aZ_b}{(Z_b-Z_a)^2}\right\}},$$

Further, $\cosh\gamma=1+2\sinh^2\frac{\gamma}{2}=\frac{Z_a+Z_b}{Z_b-Z_a}$.

$$\therefore\quad \sinh\frac{\gamma}{2}=\sqrt{\left\{\frac{Z_a}{Z_b+Z_a}\right\}}$$

$$\cosh\frac{\gamma}{2}=\sqrt{\left(\frac{Z_b}{Z_b-Z_a}\right)}$$

Therefore $\tanh \frac{\gamma}{2}=\sqrt{\left(\frac{Z_a}{Z_b}\right)}$ ...(8)

From equations (7) and (8), it is obvious that

$$Z_a=Z_I \tanh \tfrac{1}{2}\gamma$$

and $$Z_b=Z_I/\tanh \tfrac{1}{2}\gamma.$$

Thus if we are given γ, image transfer constant, and Z_I, values Z_a and Z_b can be calculated. The above relations are, thus, design formulae for lattice type attenuator.

NUMERICAL EXAMPLES

1. *Design a low pass constant K filter for* 600 *ohms line to pass speech frequencies upto* 2500 *c/s, employing,*

(*a*) *T section* (*b*) π *section.*

For a low pass constant K filter, the product of Z_1 and Z_2 is independent of frequency, *i.e.*

$$K^2=Z_1Z_2=j\omega L.\frac{1}{j\omega C}=\frac{L}{C}$$

The cut off frequency of a low pass constant K type filter is

$$f_o=\frac{1}{\pi\sqrt{(LC)}}$$

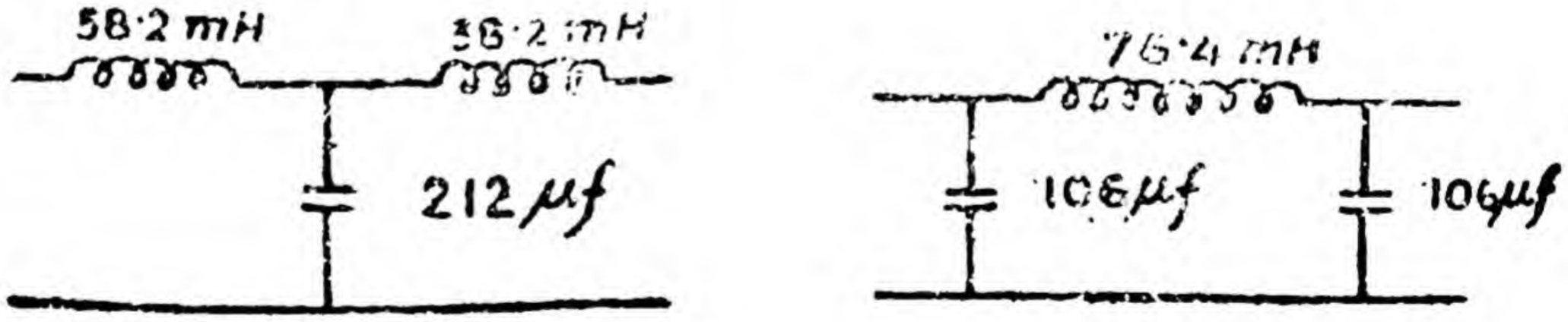
Thus the values of L and C in terms of K and f_c can be given as

$$L=\frac{K}{\pi f_o}=\frac{600}{3\cdot14.250}=\frac{6}{78\cdot5}=\mathbf{76\cdot4\ mH}$$

and

$$C=\frac{1}{\pi K f_c}=\frac{1}{15.10^3.3\cdot14}=\mathbf{0\cdot212\ \mu F}$$

With these elements two sections can be represented as



(*a*) *T*-section (*b*) π section

Fig. 28.

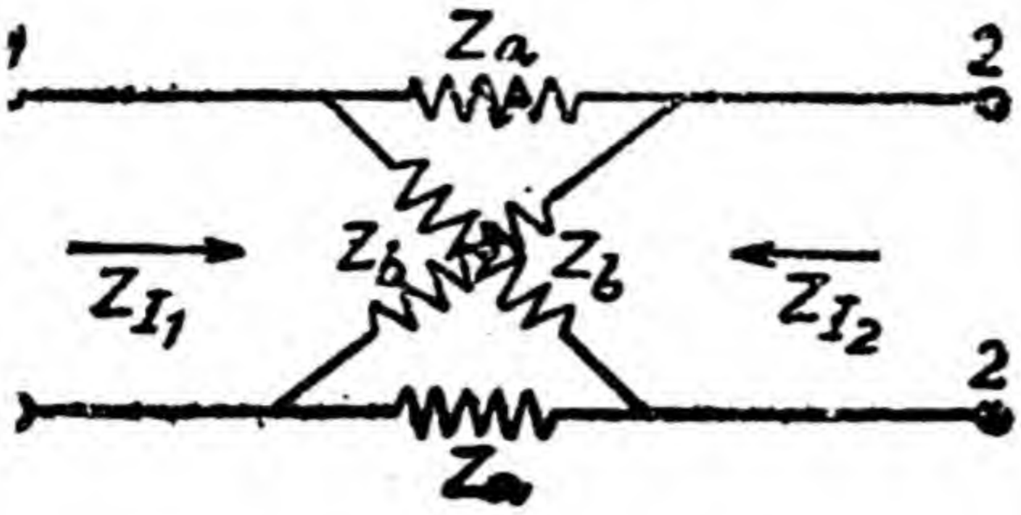
2. *In a constant K band pass filter the ratio of the capacitance in the shunt and series arms is* 100 : 1 *and the resonant frequency in both arms is* 1000 *c/s. Calculate pass bandwidth of filter.*

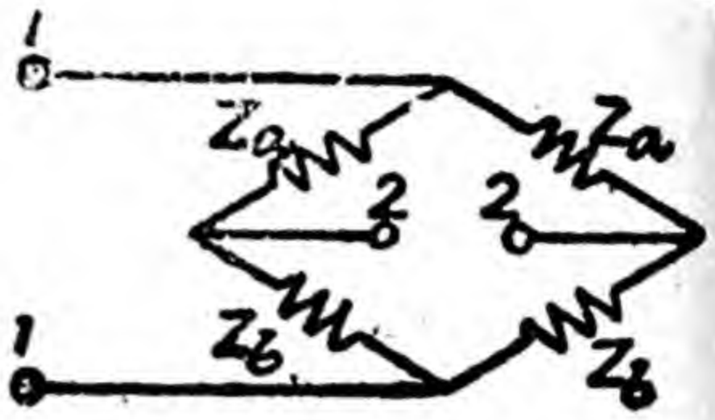
Given $\frac{C_2}{C_1}=\frac{100}{1}$ and $f_2=1000$ c/s.

We know that in case of constant K band pass filter,

$$f_1=f_0\left[\sqrt{(1+C_1/C_2)}+\sqrt{(C_1/C_2)}\right].$$

figure 27 (*a*). A lattice is like a bridge with input applied at one diagonal and output taken from other diagonal fig. 27 (*b*).





$$Z_{I_1}=Z_{I_2}-Z_I.$$

Fig. 27 (*a*). Fig. 27 (*b*).

If terminals 2-2 are open, then the resultant impedance Z_0 will be

$$\frac{1}{Z_0}=\frac{1}{Z_a+Z_b}+\frac{1}{Z_a+Z_b}$$

$$\therefore\ Z_0=\frac{Z_a+Z_b}{2}.$$

While with 2-2 shorted, resultant impedance Z_s will be

$$\frac{1}{Z_s}=\frac{1}{2}\left(\frac{1}{Z_a}+\frac{1}{Z_b}\right)$$

$$\therefore\ Z_s=\frac{2Z_aZ_b}{Z_a+Z_b}.$$

Now image impedances

$$Z_{I_1}=Z_{I_2}=Z_I=\sqrt{(Z_0Z_s)}$$

$$=\sqrt{\left(\frac{Z_a+Z_b}{2}\cdot\frac{2Z_aZ_b}{Z_a+Z_b}\right)}$$

$$\therefore\qquad Z_I=\sqrt{(Z_aZ_b)}. \qquad \text{...(7)}$$

And image transfer constant γ is given by

$$\tanh^2\gamma=\frac{Z_s}{Z_0}=\frac{4Z_aZ_b}{(Z_a+Z_b)^2},$$

from which it can be seen that

$$\cosh\gamma=\frac{Z_a+Z_b}{Z_b-Z_a}$$

and

$$\sinh\gamma=\sqrt{\left\{\frac{4Z_aZ_b}{(Z_b-Z_a)^2}\right\}},$$

Further, $\cosh\gamma=1+2\sinh^2\frac{\gamma}{2}=\frac{Z_a+Z_b}{Z_b-Z_a}$.

$$\therefore\quad \sinh\frac{\gamma}{2}=\sqrt{\left\{\frac{Z_a}{Z_b+Z_a}\right\}}$$

$$\cosh\frac{\gamma}{2}=\sqrt{\left(\frac{Z_b}{Z_b-Z_a}\right)}$$

Therefore $\tanh \frac{\gamma}{2}=\sqrt{\left(\frac{Z_a}{Z_b}\right)}$ ...(8)

From equations (7) and (8), it is obvious that

$$Z_a=Z_I \tanh \tfrac{1}{2}\gamma$$

and $$Z_b=Z_I/\tanh \tfrac{1}{2}\gamma.$$

Thus if we are given γ, image transfer constant, and Z_I, values Z_a and Z_b can be calculated. The above relations are, thus, design formulae for lattice type attenuator.

NUMERICAL EXAMPLES

1. *Design a low pass constant K filter for* 600 *ohms line to pass speech frequencies upto* 2500 *c/s, employing,*

(a) T section (b) π section.

For a low pass constant K filter, the product of Z_1 and Z_2 is independent of frequency, *i.e.*

$$K^2=Z_1Z_2=j\omega L.\frac{1}{j\omega C}=\frac{L}{C}$$

The cut off frequency of a low pass constant K type filter is

$$f_o=\frac{1}{\pi\sqrt{(LC)}}$$

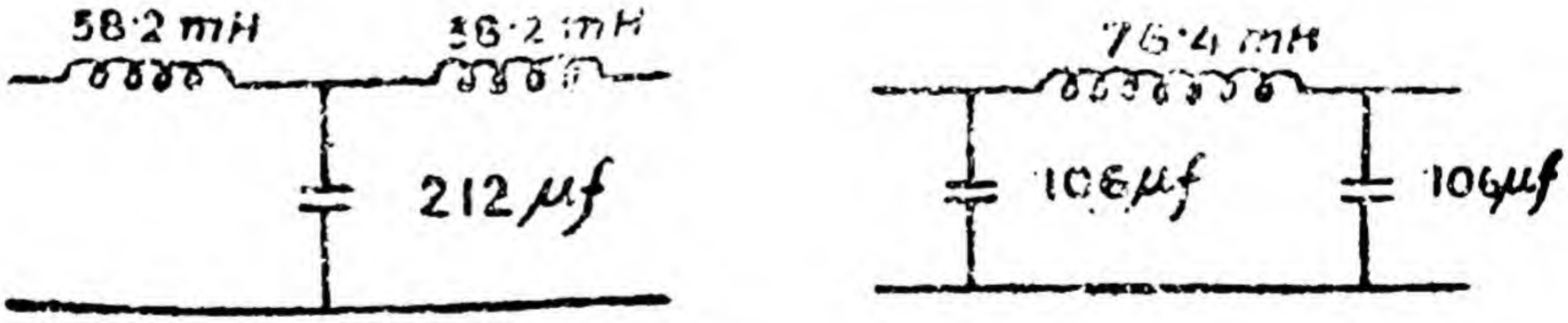
Thus the values of L and C in terms of K and f_c can be given as

$$L=\frac{K}{\pi f_o}=\frac{600}{3\cdot14.250}=\frac{6}{78\cdot5}=\mathbf{76\cdot4\ mH}$$

and

$$C=\frac{1}{\pi K f_c}=\frac{1}{15.10^3.3\cdot14}=\mathbf{0\cdot212\ \mu F}$$

With tnese elements two sections can be represented as



(*a*) *T*-section (*b*) π section

Fig. 28.

2. *In a constant K band pass filter the ratio of the capacitance in the shunt and series arms is* 100 : 1 *and the resonant frequency in both arms is* 1000 *c/s. Calculate pass bandwidth of filter.*

Given $\frac{C_2}{C_1}=\frac{100}{1}$ and $f_2=1000$ c/s.

We know that in case of constant K band pass filter,

$$f_1=f_0\,[\sqrt{(1+C_1/C_2)}+\sqrt{(C_1/C_2)}].$$

and $$f_2=f_0\,[\sqrt{(1+C_1/C_2)}-\sqrt{(C_1/C_2)}]$$

Therefore pass band is

$$f_1-f_2=2.\,f_0\sqrt{(C_1/C_2)}$$

$$=2.1000.\sqrt{\left(\frac{1}{100}\right)}$$

$$=\mathbf{200\ c/s.}$$

3. ***Calculate values of the elements of a T-section low pass filter having a characteristic impedance of*** 600 ***ohms and a cut off frequency*** 3 ***kilocycles per second.***

In a low pass filter the cut off frequency f_o is given by

$$f_o=\frac{1}{\pi\sqrt{(LC)}}. \qquad \ldots(1)$$

and characteristic impedance

$$K^2=Z_1Z_2. \qquad \ldots(2)$$

$$K^2=\omega L\times\frac{1}{\omega C}=\frac{L}{C},$$

$$(600)^2=\frac{L}{C}$$

or $$L=(600)^2.C. \qquad \ldots(3)$$

The cut off frequency $f_1=3\times10^2$ c/s : hence from equation (1),

$$3.10^3=\frac{1}{3{\cdot}14\sqrt{\{(600)^2\,C^2\}}}=\frac{1}{3{\cdot}14\times600\times C}$$

$$\therefore\ C=\frac{1}{3\times10^3\times3{\cdot}14\times600}=\frac{10^{-3}}{5652}$$

$$\mathbf{C=0{\cdot}177\times10^{-6}\ Farad.}$$

Putting the value of C in equation (3)

$$L=600\times600\times0{\cdot}177\times10^{-6}$$

$$=\mathbf{63{\cdot}72\ mH.}$$

4. ***In a high pass filter for*** 600 ***ohms termination is required to cut off below*** 20,000 ***c/s. Compute the value of the shunt and series element for T type filter.***

For a T type high pass filter,

$$f_c=\frac{1}{4\pi\sqrt{(LC)}} \qquad \ldots(1)$$

and $$K=\sqrt{\left(\frac{L}{C}\right)} \qquad \ldots(2)$$

From equation (1),

$$20{,}000=\frac{1}{4\pi\sqrt{(LC)}} \qquad \ldots(3)$$

From equation (2)

$$600=\sqrt{\left(\frac{L}{C}\right)}. \qquad \ldots(4)$$

From equations (3) and (4), we have

$$\frac{1}{4\pi C}=600\times 20{,}000$$

$$\frac{1}{C}=4\times 3{\cdot}14\times 600\times 20{,}000$$

$$C=\frac{1}{4\times 3{\cdot}14\times 600\times 20{,}000}$$

$$=\frac{1\times 10^{-6}}{15172}.$$

$$\mathbf{C=0{\cdot}0066\ \mu F.}$$

But each series element in T-section is $2C$; hence the series element is $0{\cdot}0132\ \mu F$.

Now $$\frac{K}{f_c}=4\pi L$$

or $$L=\frac{K}{4\pi\times f_c}=\frac{600}{4\times 3{\cdot}14\times 20{,}000}$$

$$=0{\cdot}00239$$

$$=\mathbf{2{\cdot}39\ mH.}$$

5. *A constant K band pass filter terminated in* 600 *ohms has a lower cut off frequency of* 12 *kc/s and an upper cut off frequency of* 123 *kc/s. Calculate the values of the elements for a T-section.*

We know that

$$\omega_1=\omega_0\left[\sqrt{\left(\frac{C_1}{C_2}+1\right)}+\sqrt{\left(\frac{C_1}{C_2}\right)}\right] \qquad \ldots(1)$$

and $$\omega_2=\omega_0\left[\sqrt{\left(\frac{C_1}{C_2}+1\right)}-\sqrt{\left(\frac{C_1}{C_2}\right)}\right] \qquad \ldots(2)$$

$$2\pi.120\times 10^3=\omega_0\left[\sqrt{\left(\frac{C_1}{C_2}+1\right)}+\sqrt{\left(\frac{C_1}{C_2}\right)}\right] \qquad \ldots(3)$$

$$2\pi.123\times 10^3=\omega_0\left[\sqrt{\left(\frac{C_1}{C_2}+1\right)}-\sqrt{\left(\frac{C_1}{C_2}\right)}\right] \qquad \ldots(4)$$

Now from equations (1) and (2),

$$\frac{\omega_1+\omega_2}{\omega_2-\omega_1}=\frac{\sqrt{\{(C_1/C_2)+1\}}}{\sqrt{(C_1/C_2)}},$$

$$\frac{243}{3}=\sqrt{\left(1+\frac{C_2}{C_1}\right)},$$

$$81=\sqrt{\left(1+\frac{C_2}{C_1}\right)},$$

or $$6561 = 1 + \frac{C_2}{C_1};$$

$$\therefore \quad \frac{C_2}{C_1} = 6560.$$

We also know that

$$L_1C_1 = L_2C_2 = \frac{1}{\omega_0^2} = \frac{1}{123.120.4.\pi^2.10^6} \qquad \ldots(5)$$

or $$\frac{L_1}{L_2} = \frac{C_2}{C_1} = 6560 \qquad \ldots(6)$$

$$\frac{L_1}{C_2} = K^2 = 3600. \qquad \ldots(7)$$

From equations (5) and (7),

$$\frac{L_1^2C_1}{C_2} = \frac{3600}{123 \times 120 \times 4\pi^2 \times 10^6}$$

or $$L_1^2 = \frac{3600 \times C_2/C_1}{123 \times 120 \times 4\pi^2 \times 10^6} = \frac{3600 \times 6560}{123 \times 120 \times 4\pi^2 \times 10^6}$$

$$\mathbf{L = 6451\ mH.}$$

$$L_2 = \frac{L_1}{C_2/C_1} = \frac{L_1}{6590} = \frac{6{\cdot}451}{6560} = \mathbf{0{\cdot}98\ \mu H}$$

$$C_2 = \frac{L_1}{K^2} = \frac{6{\cdot}451}{3600} = \mathbf{1{\cdot}792\ \mu F.}$$

$$C_1 = \frac{C_2}{6560} = \mathbf{53{\cdot}9\ \mu\mu F.}$$

6. ***Design a L type attenuator operating with a load resistance of 100 ohms and producing a loss of 20 db.***

The iterative attenuation constant in nepers

$$= \frac{20}{8{\cdot}686} = 2{\cdot}303$$

$$e^p = 10$$

Thus $$Z_a = R_a = Z_L\ (1 - e^{-p})$$

$$= 100\ (1 - \tfrac{1}{10}) = \mathbf{90\ ohms.}$$

$$Z_b = R_b = \frac{Z_L}{e^p - 1} = \frac{100}{10 - 1} = \mathbf{11{\cdot}1\ ohms.}$$

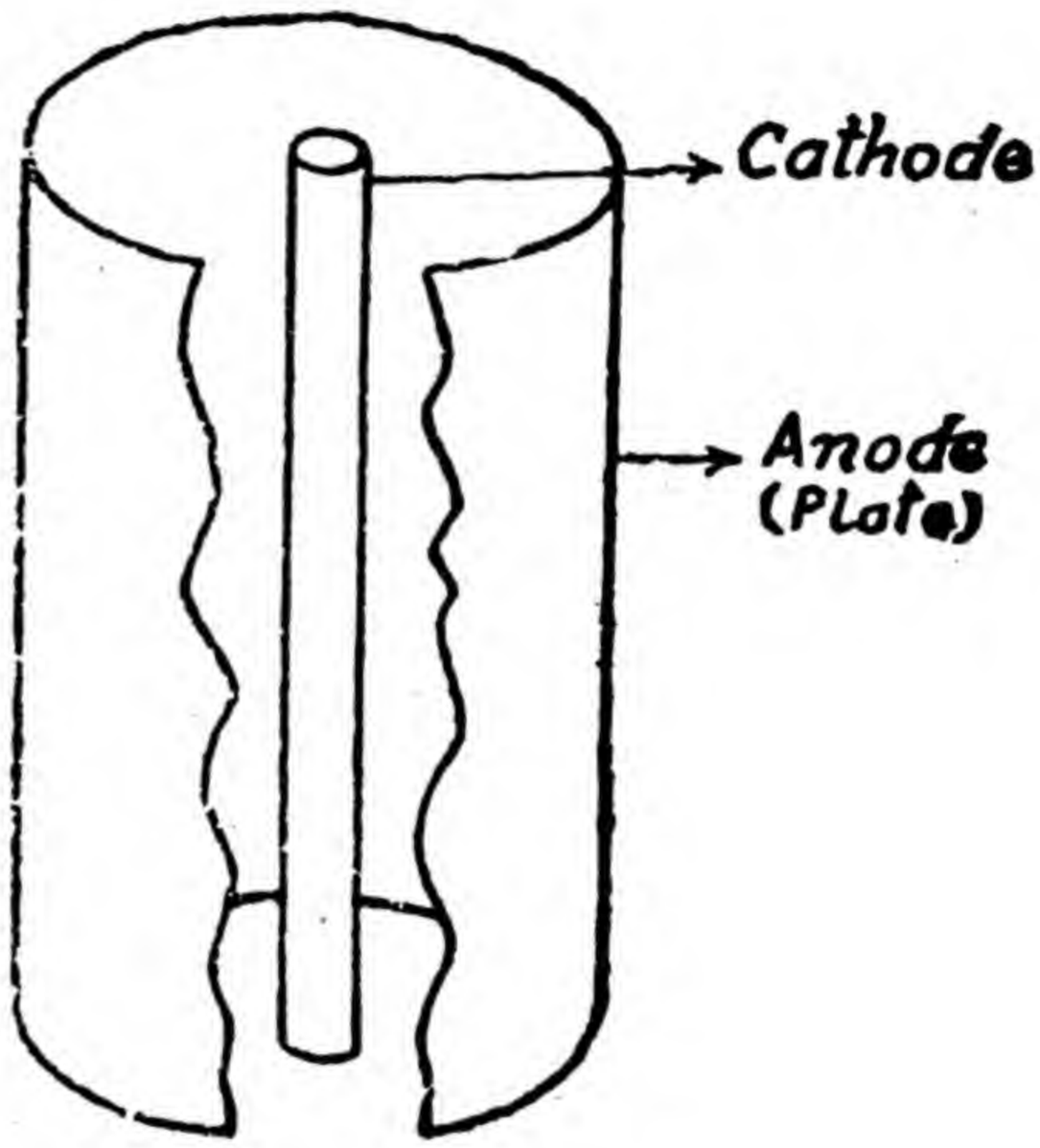
EXERCISES AND PROBLEMS

1. What do you mean by filter circuit ? How the filters are classified in accordance with their selectivity characteristics ?
2. Describe the elementary filter theory.
3. Obtain expressions for the characteristic impedance of T and π sections.
4. Prove that the characteristic impedance of either T or π section is equal to the geometric mean of their respective open and short circuit impedances.

5. What do you mean by low pass, high pass and band filters. Calculate the iterative impedance in case of band pass filter.
6. What are the disadvantages of simple types of filter ? How are these removed in *m* derived filters ?
7 What are attenuators ? Describe lattice type attenuator.
8. A constant *K* band elimination filter is required to suppress frequencies between 1,000 c/s, to 2000 c/s and to have a terminal impedance of 500Ω. Calculate the value of the elements of (*a*) *T*-section and (*b*) a π-section.
9. A symmetrical filter section of the *T*-type has series inductance of 100*mH* and a shunt capacitance of 1μ*F*. Determine the cut-off frequency of the filter.
10. Calculate the component values of a band pass filter having cut-off frequencies of 3,000 c/s. and 6,000 c/s. and a terminating resistance of 10,00Ω. Draw a circuit diagram showing the component values for (*a*) a *T*-section ; and (*b*) a π-section,

5 THERMIONIC TUBES

Thermionic tubes play important part in the operation of electronic equipments. Their development has facilitated advances in the fields of power and transportation. In radio circuits vacuum tubes are used as (*i*) voltage amplifiers, (*ii*) power amplifiers, (*iii*) detectors, (*iv*) oscillators, (*v*) frequency convertors, (*vi*) rectifiers, (*vii*) regulators, (*viii*) modulators. We shall describe all these functions, performed by the vacuum tubes in the next chapters of this book. The purpose of this chapter is to list the basic types of thermionic tubes and their fundamental characteristics.



VACUUM TUBES

5·0. VACUUM DIODE :

The vacuum diode is a two electrode vacuum tube. One electrode, called cathode, acts as an emitter of electrons. The emitter may be either directly heated or indirectly heated. The second electrode is called anode or plate. Structure is shown in figure 1 with its symbol.

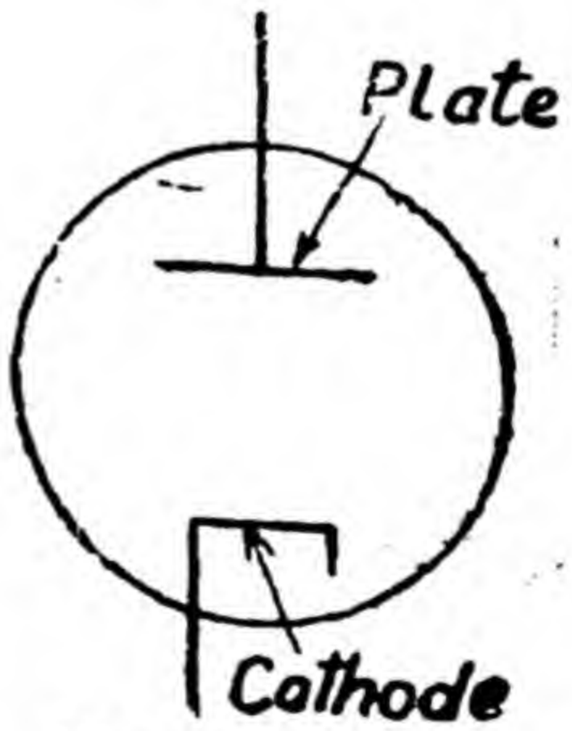


Fig. 1. Diode : structure and symbol.

If a positive voltage is applied to the plate, the electrons emitted from the cathode, being negative, will be attracted towards the plate. These electrons will flow through the external plate battery circuits, constituting the plate current. If the polarity of the plate is changed, electrons will be repulsed to the cathode and the current becomes zero. The polarity of the plate changes when an

alternaitng voltage is applied to the plate. The plate current flows only for the positive half cycles of the applied voltage. This plate current, the flow being in one direction only, is called the rectified current. This property makes the diode useful as a detector and a rectifier for d.c. power supplies.

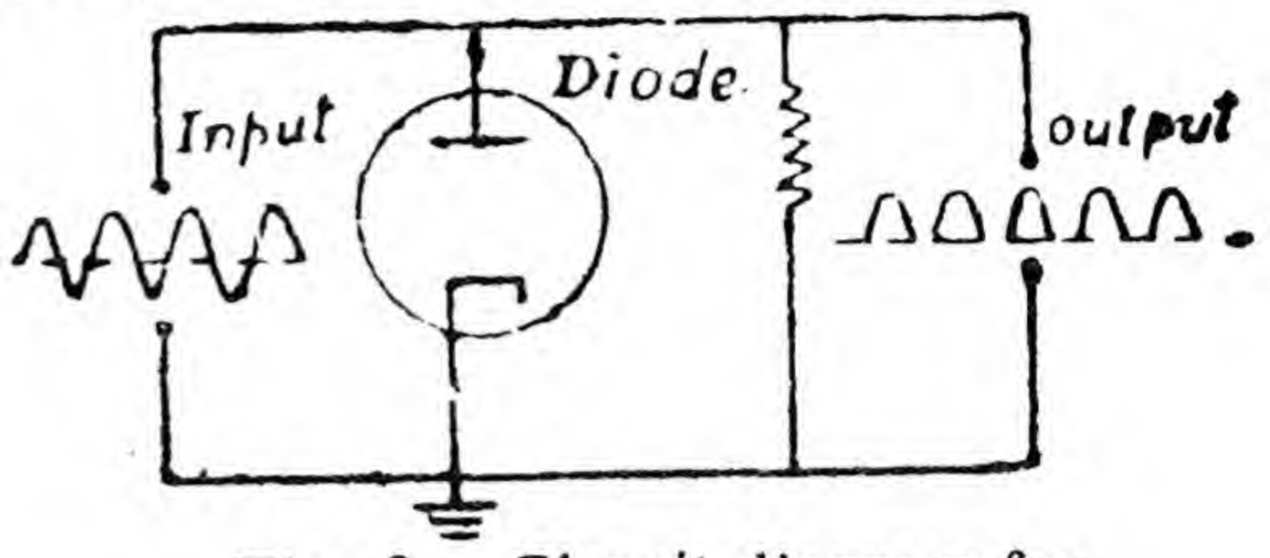


Fig. 2. Circuit diagram for vacuum diode.

For full wave rectification, tube with two plates and one or two cathodes, called duodiode, is used. One of the plates is always positive with respect to the cathodes and thus both halves of the applied voltage are rectified.

Figure 3 shows the circuit arrangement for finding the characteristic curves of thermionic vacuum diode.

In order to draw the characteristics, the filament current is adjusted at a fixed value and the plate voltage e_b is increased in regular steps. The plate current is noted at each steps The observations are repeated for other values of filament current Now the plate current is plotted against plate voltage as shown in figure 4.

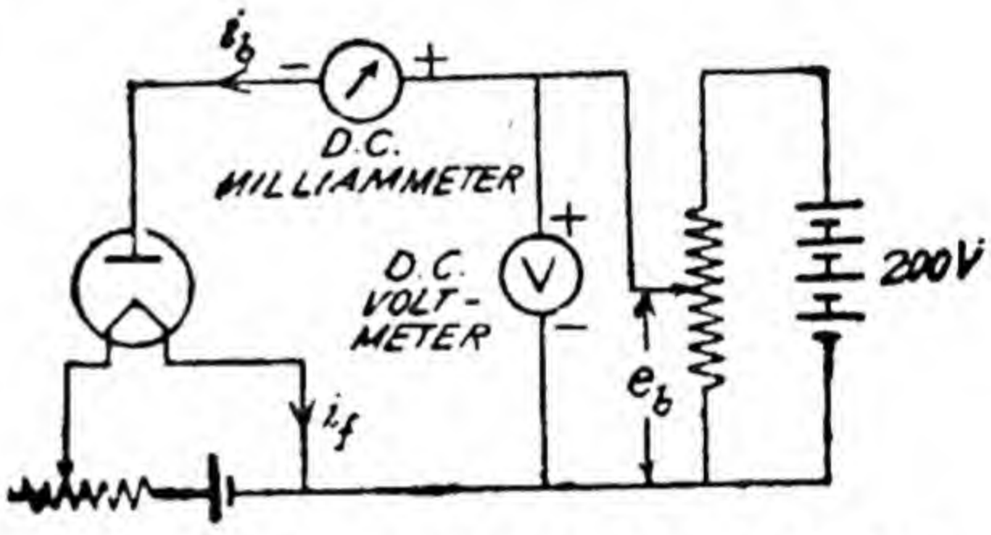


Fig. 3. Circuit of studying the characteristic of diode valve.

Suppose the cathode temperature T_1 is constant, then the plate current increases along the curve OA as the plate voltage is increased from zero. In the region OA, the number of electrons emitted from the cathode is large than the number of electrons which can be attracted by the plate voltage. The excess electrons form a cloud in the space surrounding the filament which is known as *space charge*. This negative charge retards the emission of electrons from cathode which in turn reduces the space charge. For a certain value of plate voltage, the number of emitted electrons is equal to the number of collected electrons by the plate. The diode then operates at a saturated plate current If the cathode temperature is raised to T_2, the saturated current value also increases because of the fact that at higher temperature of cathode, emission of electron will be

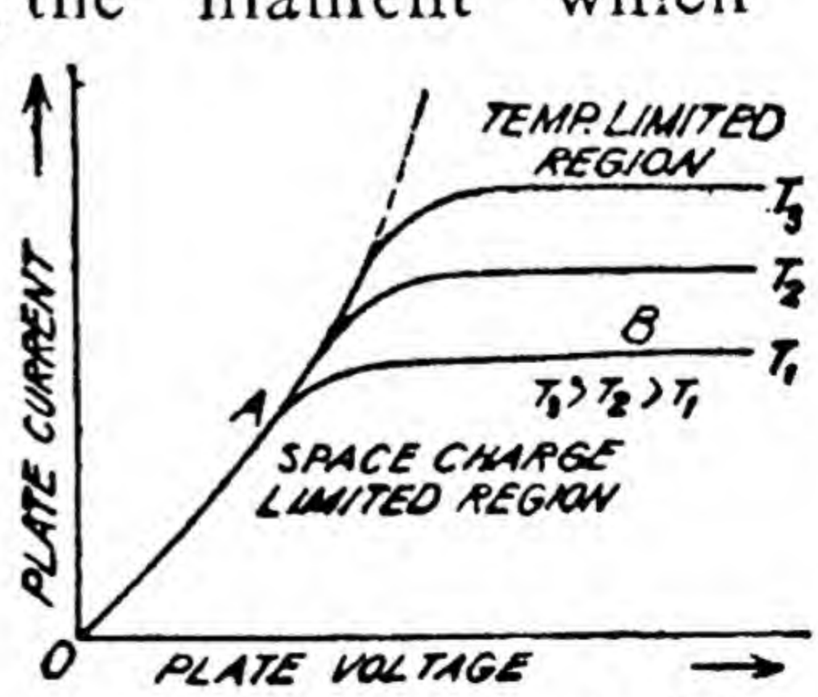


Fig. 4. Plate current-plate voltage characteristics of a diode.

greater than at temperature T_1 and then there will be more electrons to reach the plate for saturation current. When the plate potential is sufficiently high, electrons are attracted by the plate as rapidly as emitted from the cathode. The plate current is then limited by the electron emission of the cathode and so depends upon the cathode temperature rather than the plate voltage. The portion *OA*, in which the plate current is a function of plate voltage is known as *space charge limited region*, while the portion *AB*, in which plate current is dependent on plate voltage is known as *temperature limited region*.

5·1. VACUUM TRIODE :

A vacuum triode is a three electrode tube containing an emitting electrode called the 'cathode', a control electrode called the 'grid' and a current collecting electrode called the 'anode' or 'plate'. The basic structure is shown in fig. 5. The emitting electrode. may be directly or indirectly heated oxide cathode, an oxide-coated filament or a filament of tungsten or thoriated tungsten. The control grid is in the form of fine messh screen between the plate and cathode. The electrode, being nearer to cathode than the plate, will influence the electrostatic field at the cathode to a greater extent than can the flow of current from the cathode. The grid is usually operated on a slightly negative potential so that the electrons will pass between the grid wires without hitting the wires themselves.

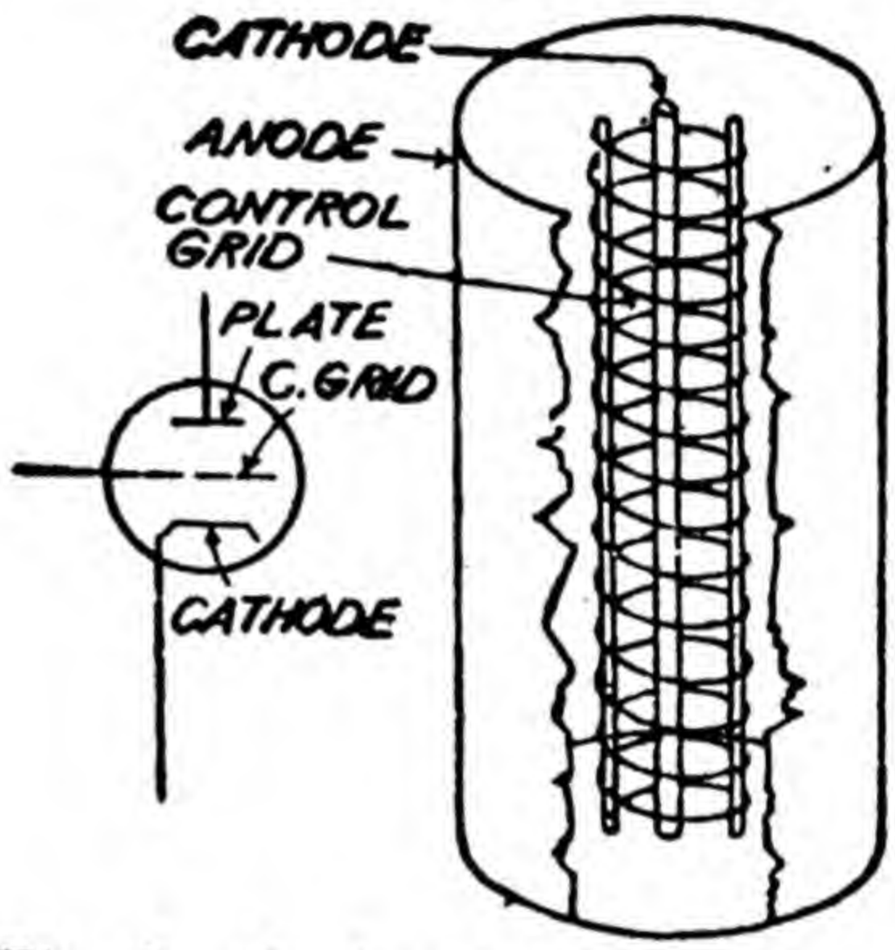


Fig 5. Basic structure and symbol of vacuum triode.

Figure 6 shows a circuit arrangement for investigating the various current voltage relations in triode. These relations can be represented by a series of characteristic curves. A single curve serves to show the variation in the plate current resulting from variation of one the curent controlling voltages, the other voltage being held cons tant

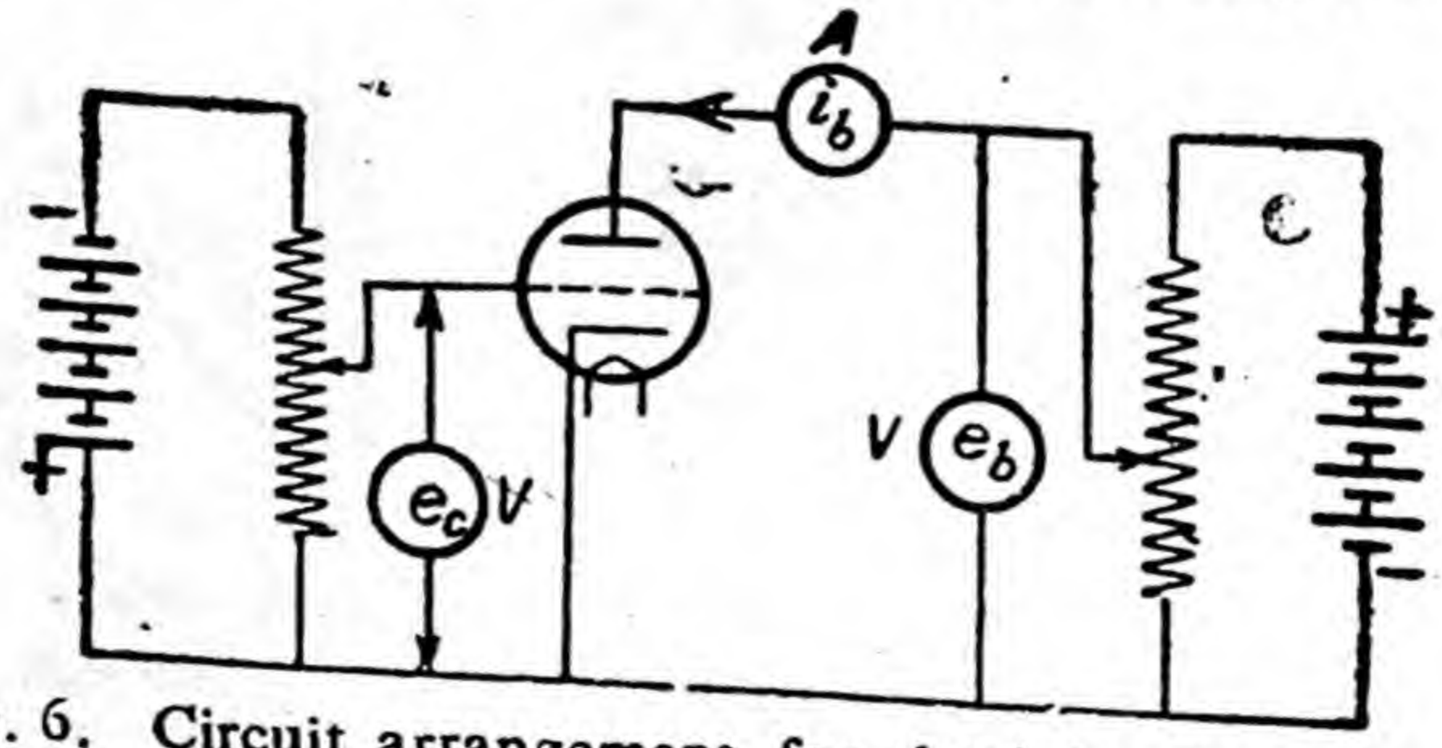


Fig. 6. Circuit arrangement for obtaining the voltage and current relation in a triode.

Figure 7 shows a family of curves for typical triode, where each curve represents a plot of plate current i_b as a function of grid voltage e_c for a fixed value of plate voltage e_b. Because the curves show the effect of a voltage variation in one circuit, the grid circuit upon the current in another circuit, the plate circuit, they are known as transfer or mutual characteristics.

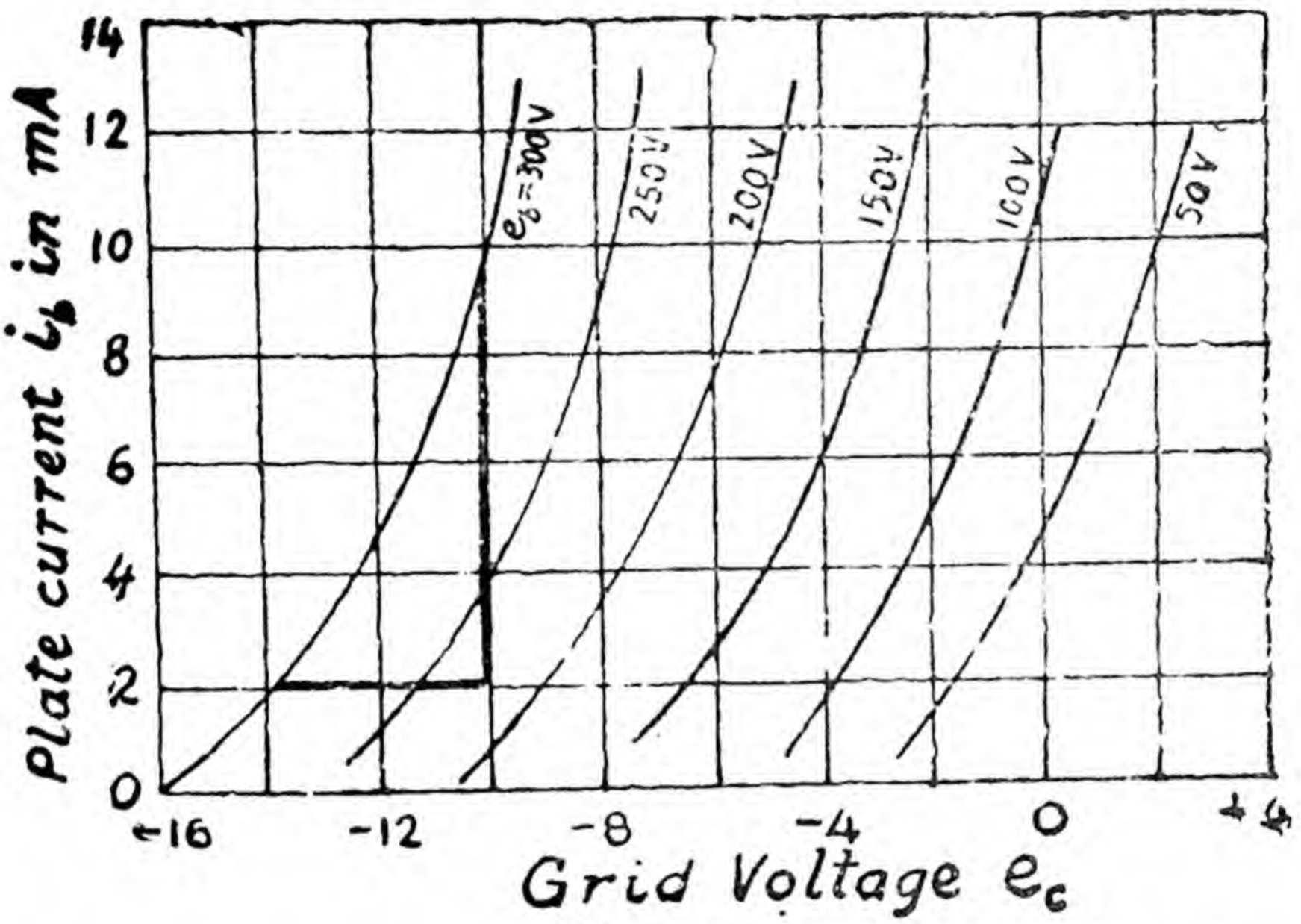


Fig. 7. Typical transfer characteristic curves of triode.

Figure 8 shows the family of other curves obtained by plot ting plate current as a function of plate voltage for various fixed values of grid voltage. The set of curves representing the current voltage relations in a triode are known as plate-characteristics.

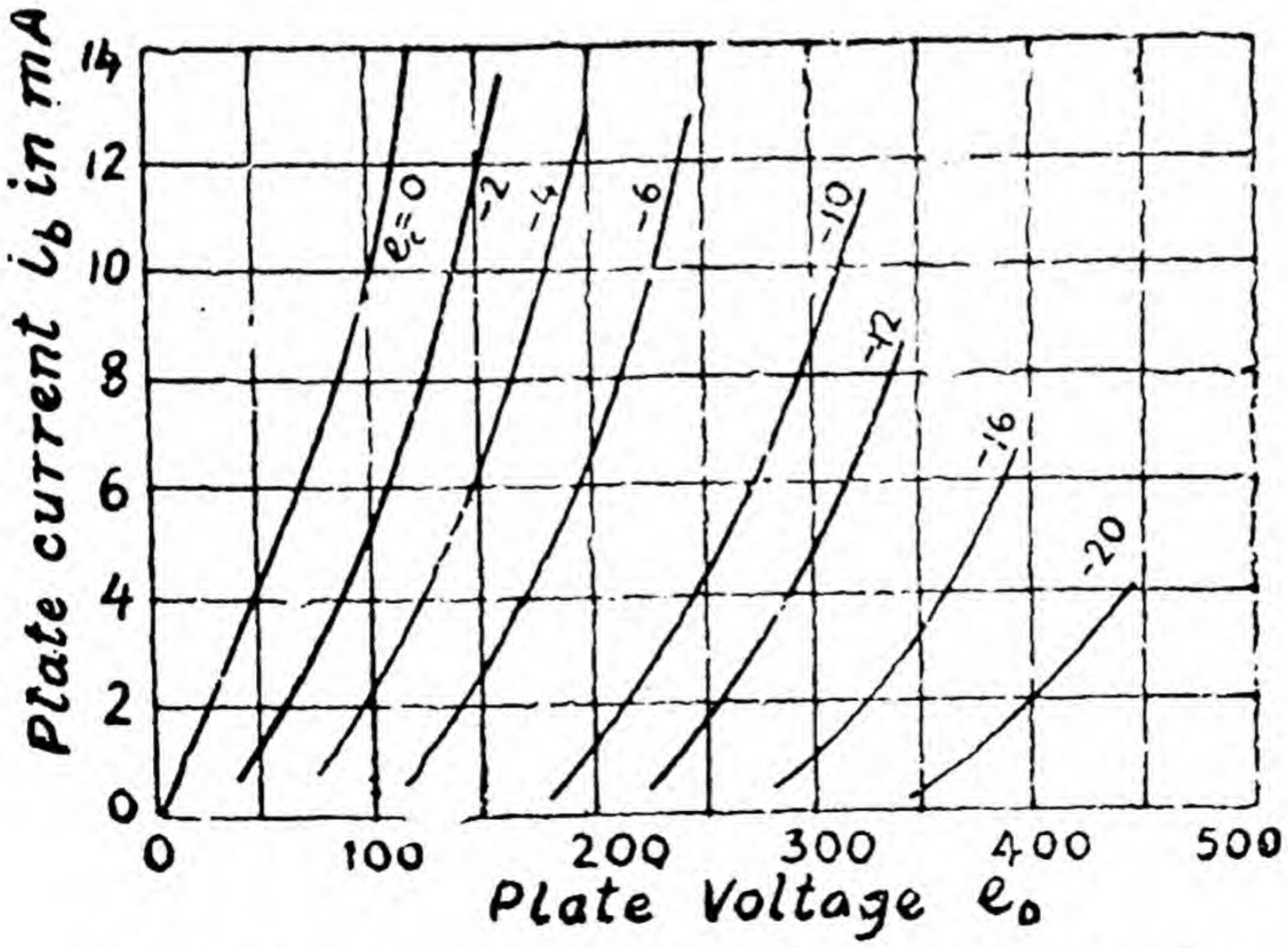


Fig. 8. Typical plate characteristic curves of a triode.

5·1-1. TUBE PARAMETERS :

Vacuum tube characteristics are often referred to as the parameters of the tube. The most commonly used parameters are :

(a) Amplification factor (μ) : This is defined as the measure of the relative ability of the plate and the grid to produce an equal change in the plate current, or

$$\mu = \frac{\text{change in plate voltage}}{\text{change in grid voltage}} \text{ for constant value of plate current}$$

$$= -\left(\frac{de_b}{de_c}\right)_{i_b}.$$

The negative sign appears because voltage changes in ratio are of opposite sign and μ is a conventionally defined as the positive number.

The amplification factor can be calculated from values obtained from the transfer characteristics of tube shown in fig. 7. We see, for example, that a change in grid voltage from —14 volts to —4 volts is associated with a change in plate voltage from 300 volts to 100 volts for a constant value of plate current (2m*A*). The amplification factor will be

$$\mu = -\frac{(300-100)}{[-14-(-4)]} = 20.$$

For most triodes amplification factor ranges between 10 and 40.

(b) Plate resistance (r_p) : It is defined as the resistance to the flow of alternating current offered by path between the cathode and plate. This is also called the dynamic plate resistance or a.c. plate resistance. Its value depends upon the grid and plate voltages applied to the tube. The expression for r_p can be written as

$$r_p = \frac{\text{change in plate voltage}}{\text{change in plate current}} \text{ for constant grid voltage}$$

$$= \left(\frac{de_b}{di_b}\right)_{e_c}.$$

The value can be estimated from fig. 7. At $e_c = -10$ volts, the change in plate current from 4 milliamperes to 10 milliamperes is associated with a change of plate voltage from 250 to 300 volts.

Therefore,

$$r_p = \frac{300-250}{(10-4)\,10^{-3}} = \frac{50}{6}.10^3 \text{ ohms.}$$

$$= 8{\cdot}33.10^3 \text{ ohms.}$$

r_p can be assumed constant over a considerable range of operation. Value of r_p varies with point of operation.

(c) Trans-conductance (g_m) : The trans-conductance of a tube may be defined as the ratio of the change in plate current to a

change in the control grid voltage when all other tube element voltages are kept constant. Mathematically,

$$g_m=\left(\frac{di_b}{de_c}\right)_{e_b},$$

where di_b=change in plate current,

de_c=change in grid voltage.

The value of g_m can be calculated from transfer characteristics (fig. 7). For example, for plate voltage $e_b=300$ volts, change in grid voltage from 14 to -10 volts is associated with a change of plate current from 2 to 10 milliamperes. Therefore,

$$g_m=\frac{8\times10^{-3}}{4}\text{mhos}=\cdot002\text{ mhos}$$

Value of g_m varies with the point of operation but variation is quite small over a considerable range of operation.

Relation in tube constants : We know that amplification factor $\mu=\frac{de_b}{de_c}$. It can be written as

$$\mu=\frac{de_b}{di_b}\cdot\frac{di_b}{de_c}$$
$$=r_p\,g_m.$$

It is obvious from figs. (7) and (8) that the slopes of mutual and plate characteristics vary with the variation of grid voltage e_c and plate voltage e_b. Hence the values of r_p (plate resistance) and g_m (Transconductance) vary accordingly; however, μ remains almost constant over a wide range of operation. Fig. (9) shows the variation of μ, g_m and r_p with plate current for a fixed plate voltage.

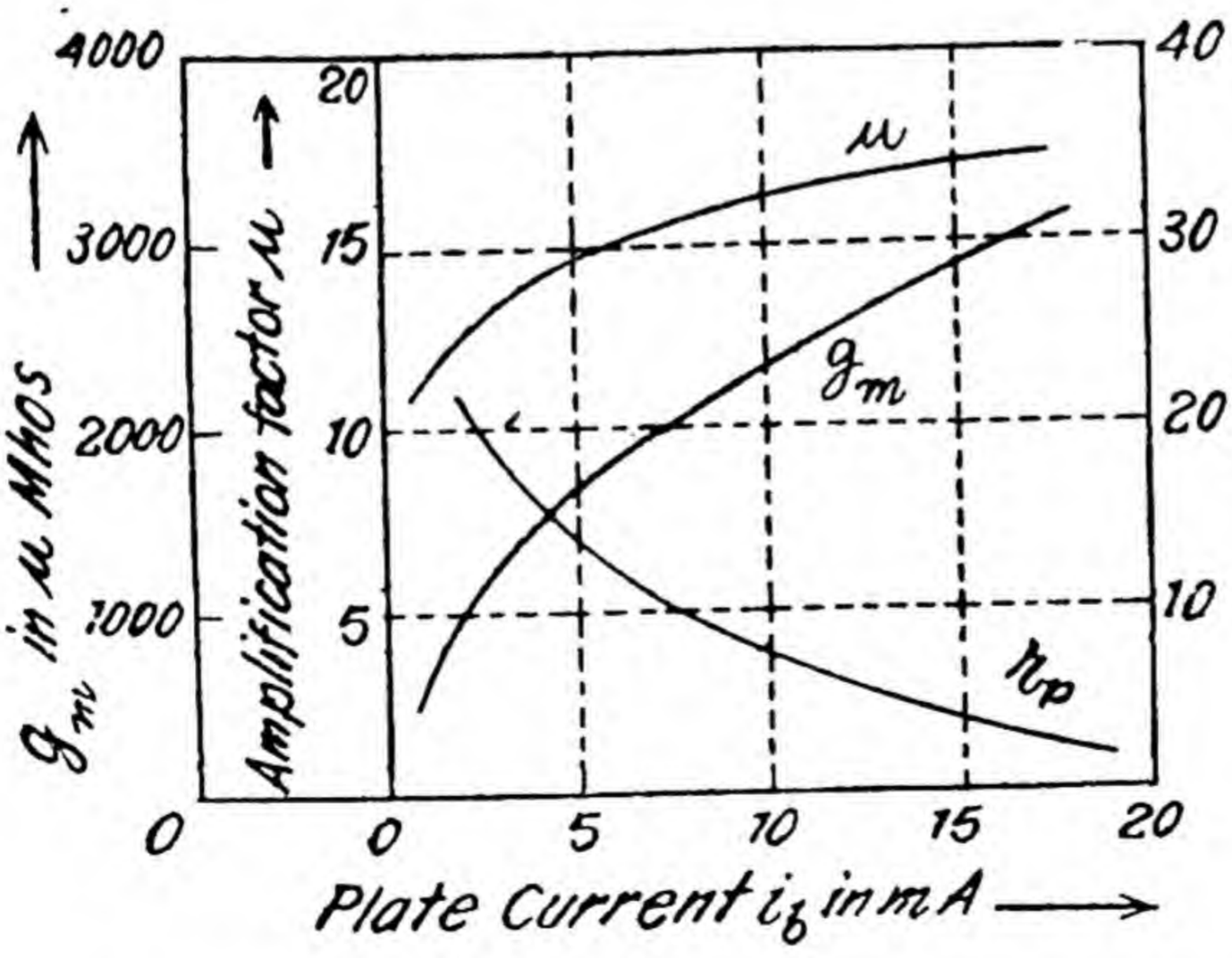


Fig. 9. Illustrating variation of μ, g_m and r_p with plate current i_b.

When the characteristic curves are linear, parallel and equi-spaced, then for equal increments in the value of parameters, these coefficients are constant. In practice this is not true. However,

within the small range, these coefficients may be taken as constant.

5·1-2. THE SHORTCOMING OF THE TRIODE :

(*i*) The value of amplification factor is relatively low. In case of voltage amplifier, the gain depends upon amplification factor μ and hence a triode should have large value of μ.

(*ii*) There exists three interelectrode capacitances between these electrodes. They are (*i*) capacity between grid and plate, C_{gp} (*ii*) the capacity between grid and cathode C_{gk} and (*iii*) the capacity between the plate and cathode, C_{pk}. The effects of these capacitances are :

(*a*) at high frequencies, the large values of C_{gp} and C_{pk} decrease the gain of amplifier and hence the triode cannot be used at high frequencies.

(*b*) when a triode of high μ is used as an amplifier then varying voltage in the plate and grid circuit will cause an a.c. to flow through C_{gp} into grid circuit. Sometimes the feed back may cause the amplifier to oscillate. Hence the triode, instead of amplifier, acts as an oscillator.

5·1-3 DYNAMIC CHARACTERISTICS :

The characteristics of a triode valve with no load in plate circuit are known as static characteristics but when a load is connected in the plate circuit, the characteristics are known as dynamic characteristics. The circuit arragement for drawing the dynamic characteristics is shown in fig. (10).

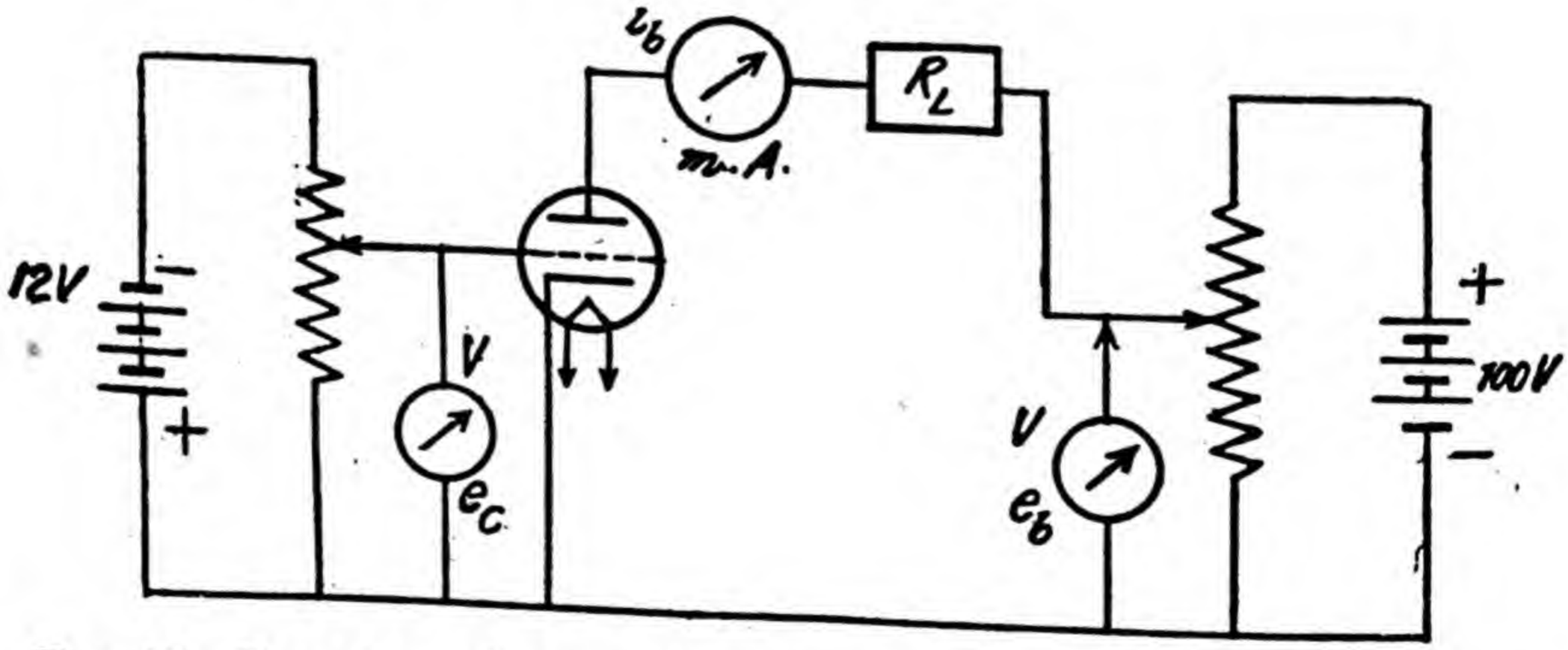


Fig. 10. Circuit arrangement for drawing the dynamic characteristics.

Fig. (11) shows the plate current-grid voltage (i_b–e_o) characteristics of a triode. Let us consider that a load R_L (20,000 Ω) is connected in plate circuit and the grid bias is such that there is no plate current. In this situation, the plate potential is the same as the plate supply (say 100 volts). This situation is shown by a point A in fig. 11.

Now suppose that the value of grid bias is such that the plate

current is 1mA. The voltage drop across the load is $20,000 \times (1 \times 10^{-3}) = 20V$. So the working conditions are now represented by a point B on 80V curve where $i_b = 1mA$ and $e_c = -6V$.

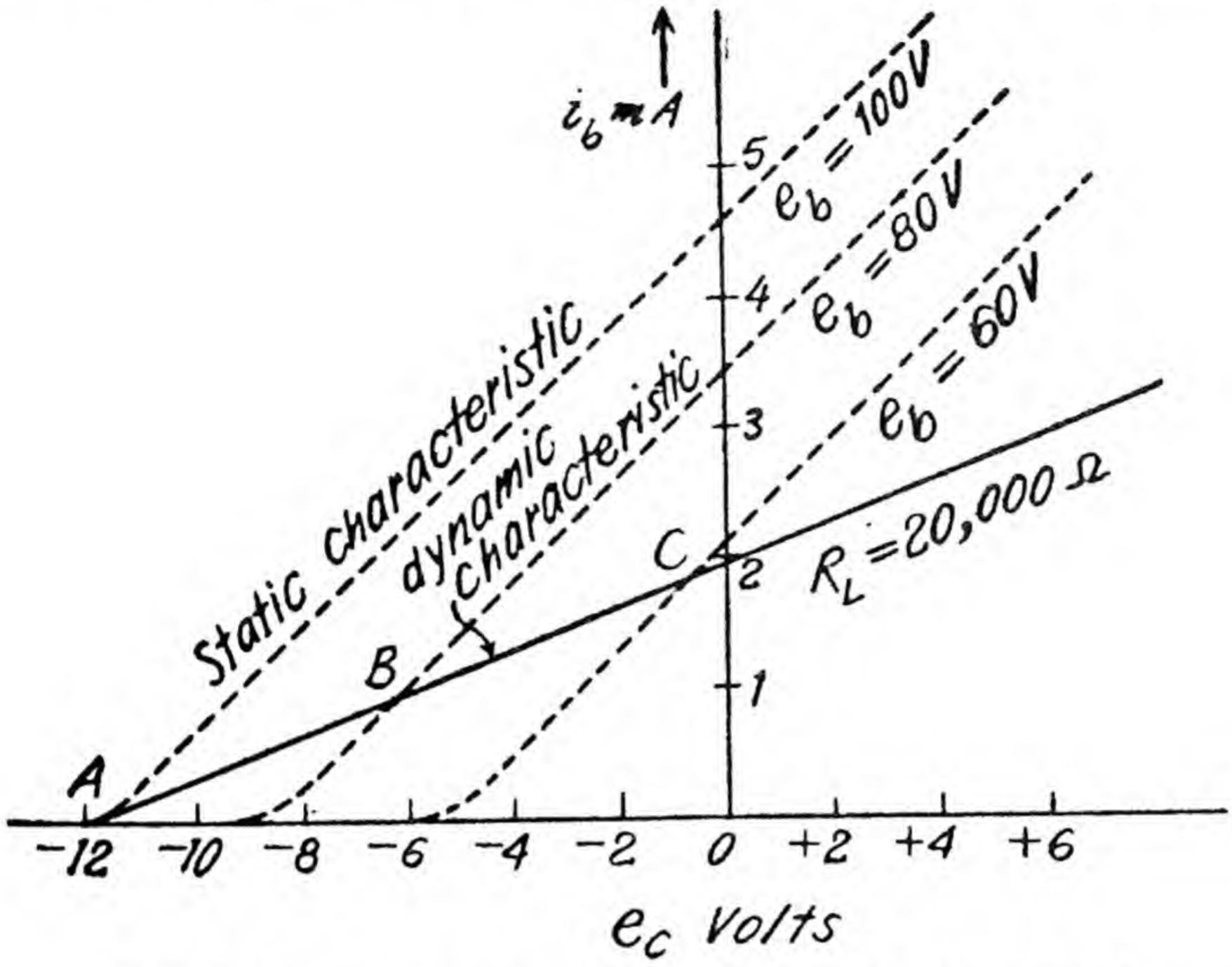


Fig. 11. Plate current grid voltage characteristics

Further suppose that the plate current is 2mA. In this case the voltage drop across load is $20,000 \times (2 \times 10^{-3}) = 40$ volts. Now the working conditions are represented by a point *C* on 60V characteristic where $i_b = 2mA$ and $e_c = -0{\cdot}6V$. The line joining the points *ABC* is the dynamic characteristic corresponding to plate load of 20,000 ohms. It is nearly straight line except at extreme and near *A* points.

5·2. VACUUM TETRODE :

Tetrode is a four-element vacuum tube. The elements are cathode, control grid, screen grid and plate. The additional grid known as screen grid, is similar in construction to the control grid except that it is usually somewhat more loosely wound. It is operated at a fixed positive potential with respect to cathode. This potential is ordinarily somewhat lower than the plate voltage.

The two main functions performed by screen grid are :

(*i*) It overcomes the space charge effect being at positive voltage, and thereby increases the plate current. Thus the screen grid reduces the effects of plate potential upon the space charge in the region surrounding the cathode, so that plate voltage variations have no appreciable influence on the plate current. Therefore in tetrode circuit the control grid voltage will produce greater influence on the output voltage variation. If we speak in terms of tube coefficients, the effect of additional grid is to increase μ and r_p.

It has no appreciable effect on g_m, the trans-conductance of the tube.

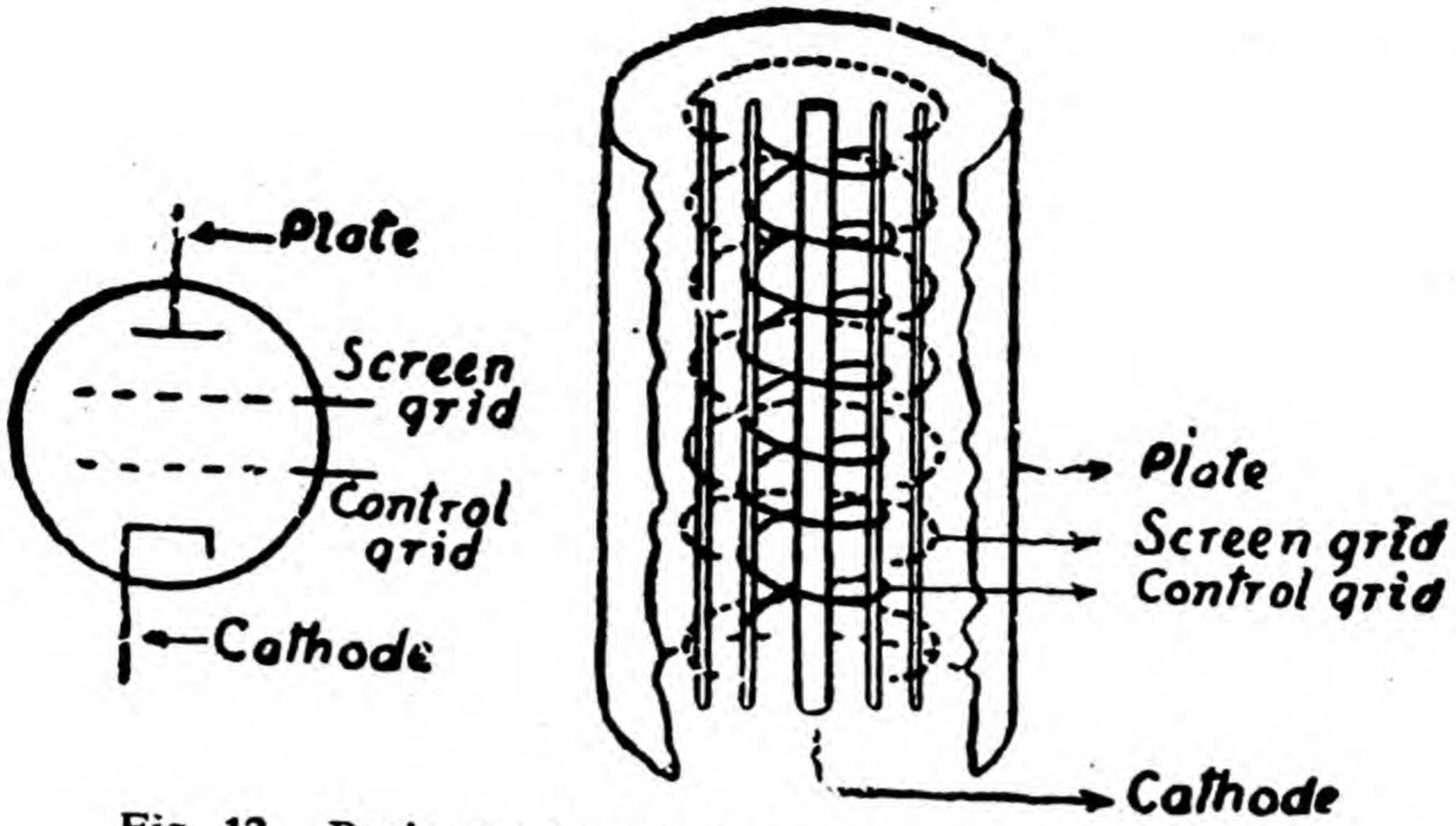


Fig. 12. Basic structure and symbol of vacuum tetrode.

(*ii*) The second function of screen grid is to improve tube performance in circuit involving high frequency voltage variations, by reducing the capacitance formed between the control grid and the plate. At radio frequencies, this inter-electrode capacitance of triode, being of the order of 10 $\mu\mu F$, becomes a source of feed back of energy from plate to grid circuit and thus causes trouble in the operation of the circuit.

The value, however, suffers from the disadvantage of *secondary emission*. When the plate is less positive than screen grid, the high velocity primary electrons from cathode strike the plate and produce secondary electrons which are attracted to the screen. Thus when the plate potential is less than screen potential, the screen current increases while the plate current decreases.

The basic circuit for drawing the characteristics of a tetrode is shown in figure 13. The screen grid potential e_s is kept at a fixed value (say 100 volts) and control grid voltage e_c is used as a parameter. Let first of all the control grid be kept at a 0 volt and the plate voltage be gradually increased in steps, the corresponding plate current i_b and screen current i_s are noted. Again the control grid potential is changed (say by 2 volts) and the observations are repeated. The curves are shown in figure 14. The space current which is the sum of plate current i_b and screen current i_s is also found. It is observed that space current is independent of plate potential.

Figure 14 shows the typical set of plate characteristics for a tetrode having a fixed screen valtage of 100 volts. In the beginning when plate voltage is less than screen voltage, and it is quite insufficient for secondary electron emission to occur from the surface of the plate, the plate current increases as the plate voltage is incre-

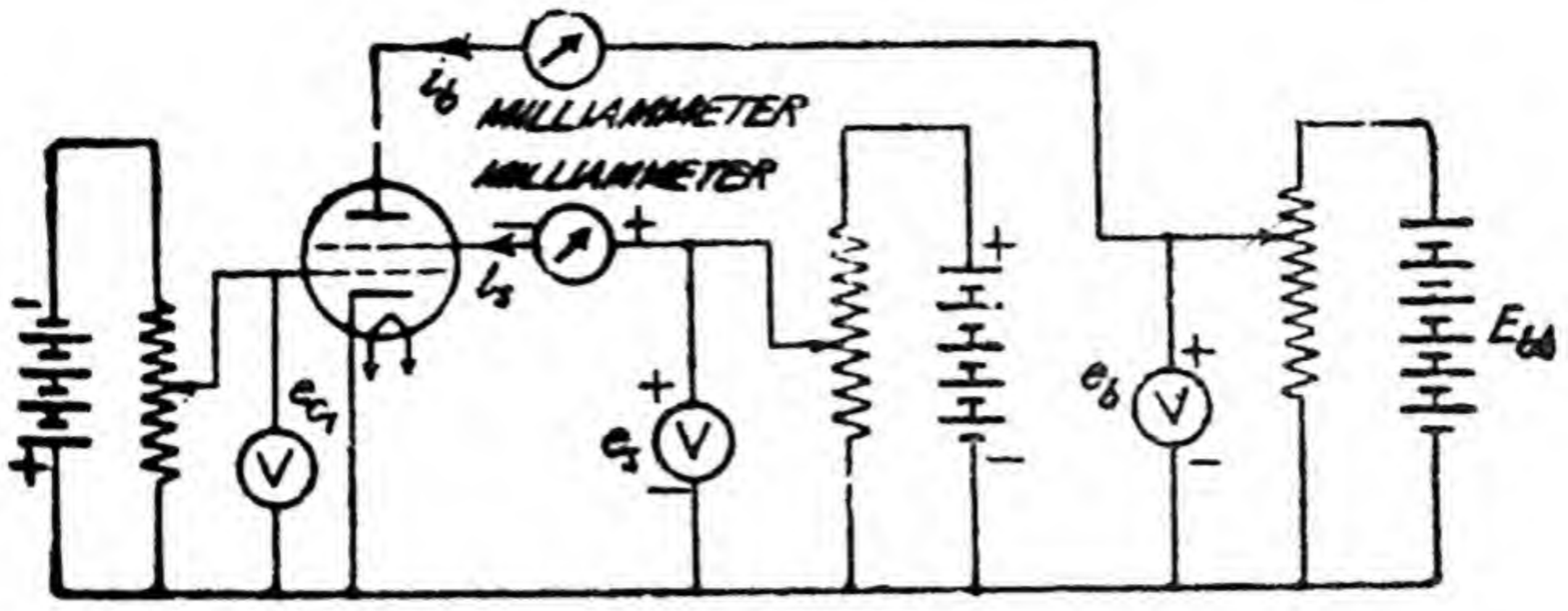


Fig. 13. Circuit connections for a tetrode.

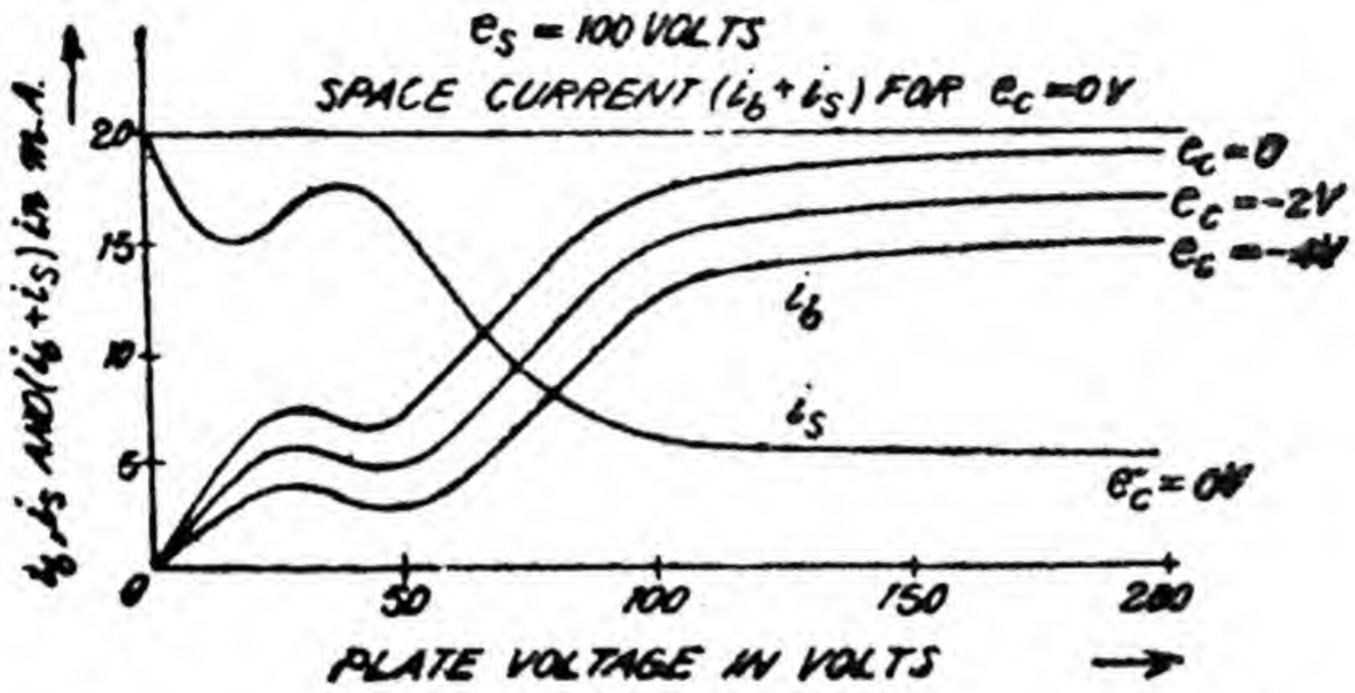


Fig. 14. i_b, i_s and $(i_b + i_s)$ as a function of plate voltage.

ased. When the plate voltage is further increased and yet plate is less positive than the screen grid, the curves show reduction in plate current. This effect is due to secondary electrons emitted from the plate by the impact of electrons coming to it with sufficient velocity. The phenomenon is called dynatron effect. Although this is undesirable when the tube is used as an amplifier yet it may be utilised effectively in case of oscillator. As the screen voltage is higher than plate voltage, these secondary electrons will be attracted by the screen grid and thus increases screen current at the cost of plate current. But when plate voltage becomes higher than screen voltage, these electrons are returned back to the plate and thus plate current increases.

5·3. PENTODE TUBE :

The dip in the plate current curve, due to secondary emission limits the voltage-variation range over which a tetrode can be operated. To avoid the dip in the plate current curve, another electrode, called the suppressor grid, is inserted between the screen grid and plate which is kept at the cathode potential The secondary electrons because of the cathode potential of suppressor grid are returned back to the plate and thus dip is eliminated as shown in figure 15 (*b*), representing the charateristic curves for vacuum pentode.

Thus pentode is a vacuum tube having cathode, control grid, screen grid, suppressor grid and plate as shown in fig 15 (*a*).

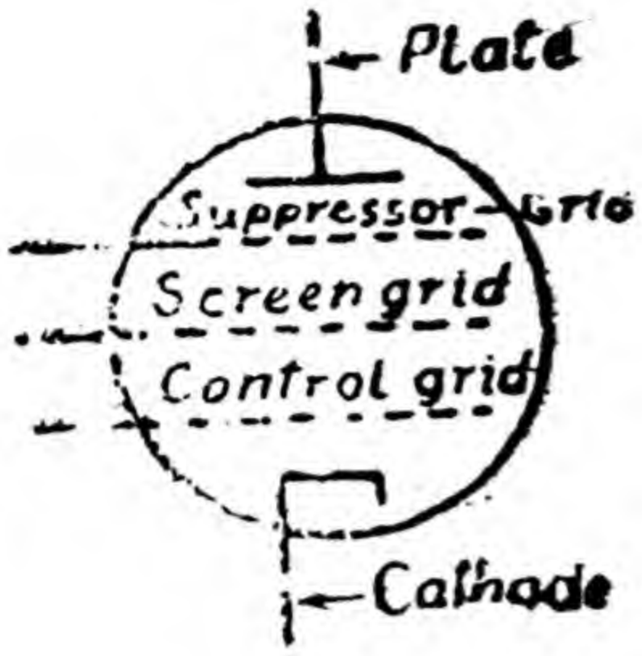


Fig. 15 (*a*) Symbolic representation of pentode.

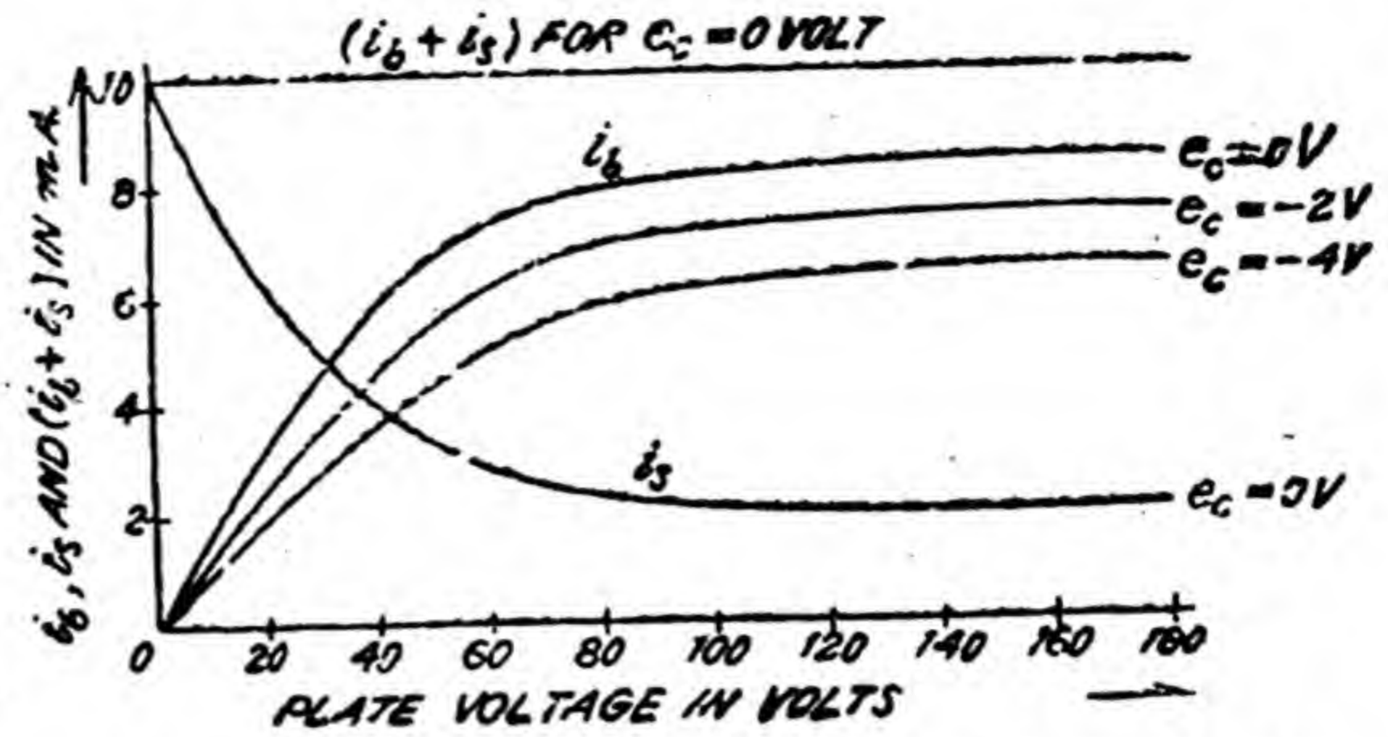


Fig. 15 (*b*) Plate characteristic of a pentode.

In normal operation the cathode is operated near ground potential, the control grid at a small negative potential, the screen grid at a relatively large positive potential, the supperssor grid at the cathode potential and the plate at the screen potential or a more positive potential.

Circuit for obtaining the static characteristics of pentode is shown in fig. 16. Starting with zero voltage on the control grid and screen at 100 volts, values of plate and screen currents are noted as the plate voltage is increased from zero to maximum working value. The test is repeated for various negative values of control grid potential. The family of curves is obtained on plotting plate current against plate potential for various negative control grid potentials.

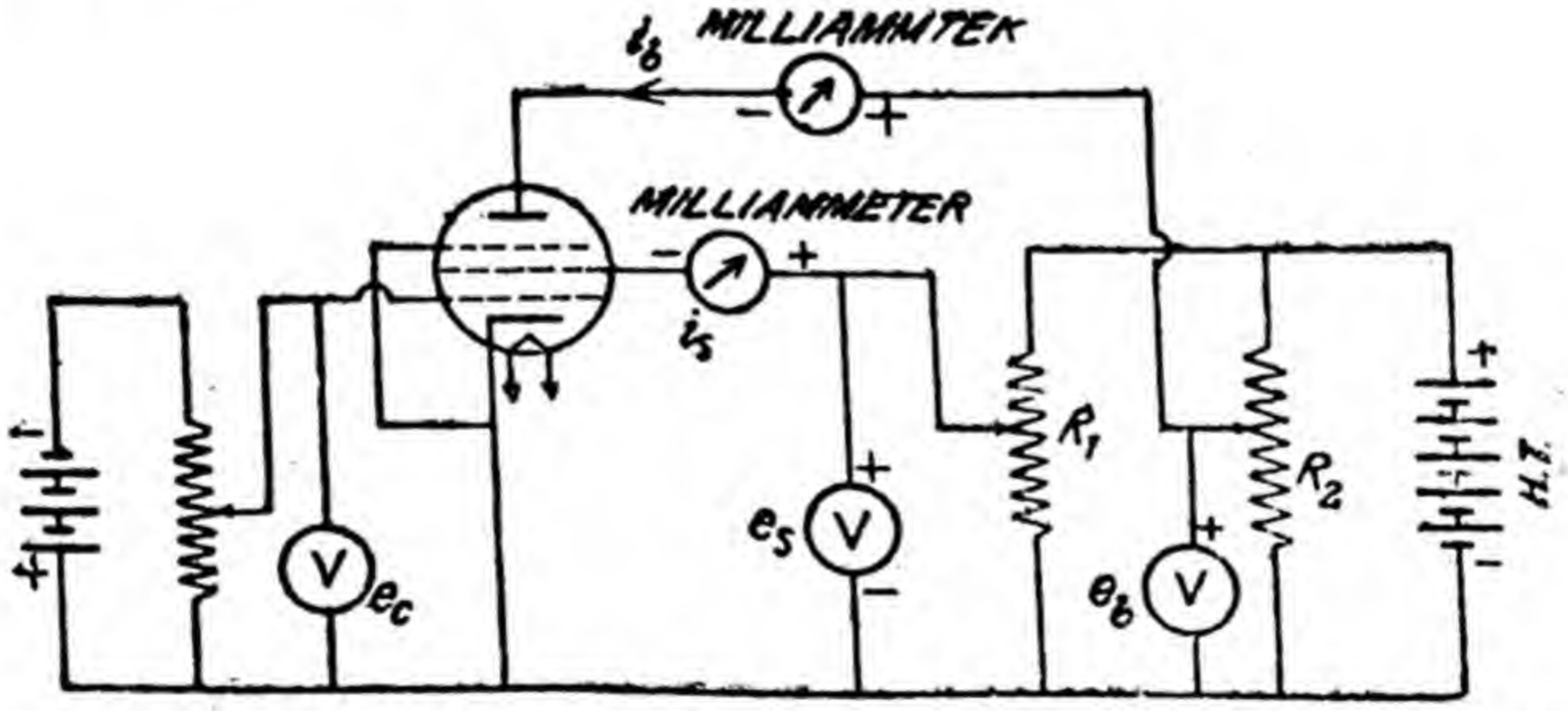


Fig. 16. Circuit for obtaining the static characteristics of pentode.

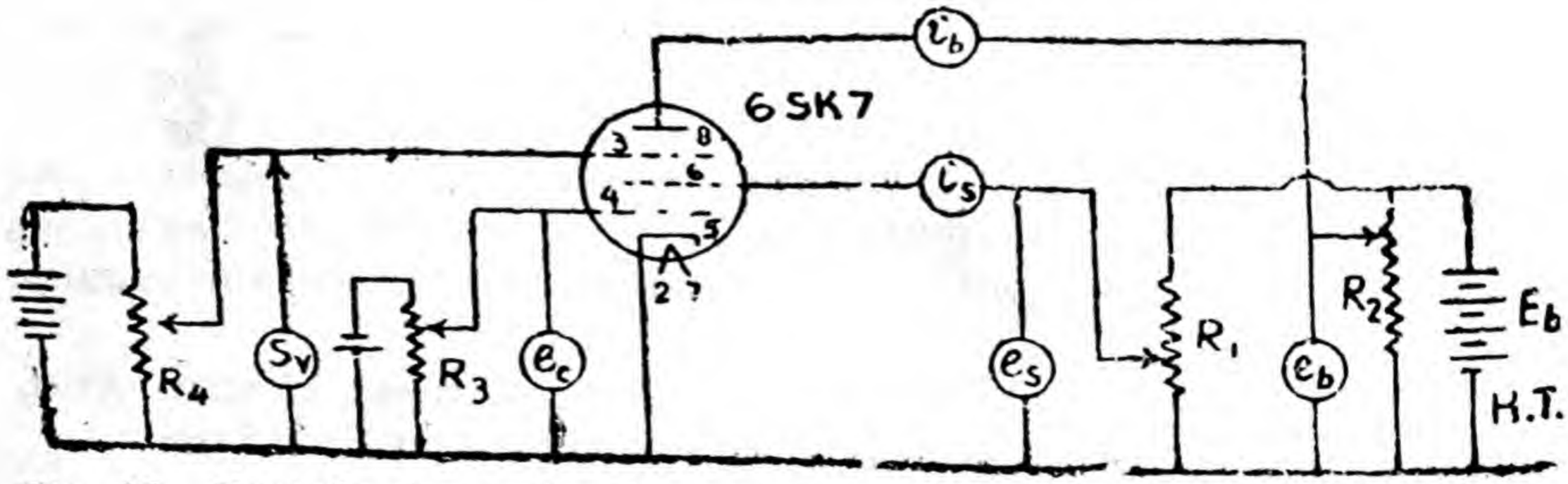


Fig. 17. Pentode circuit for observing the effect of suppressor grid potential.

To see the effect of suppressor grid voltage, it is connected to some voltage source fig. 17. Various voltages to the suppressor grid are applied. The plate and screen currents thus obtained are plotted as a function of plate potential as shown in figure 18.

The tube parameters for pentode tube are defined in the same manner as for triode or tetrode. Due to the high screening action in pentode, the amplification factor of the pentode is extremely high, of the order of 1000 or more, the plate resistance is also high. It can also be verified from the curves of fig 15*b*. that they have very small gradient so plate resistance and amplification factor are high. The value of g_m is not markedly different from those of tetrodes or triodes.

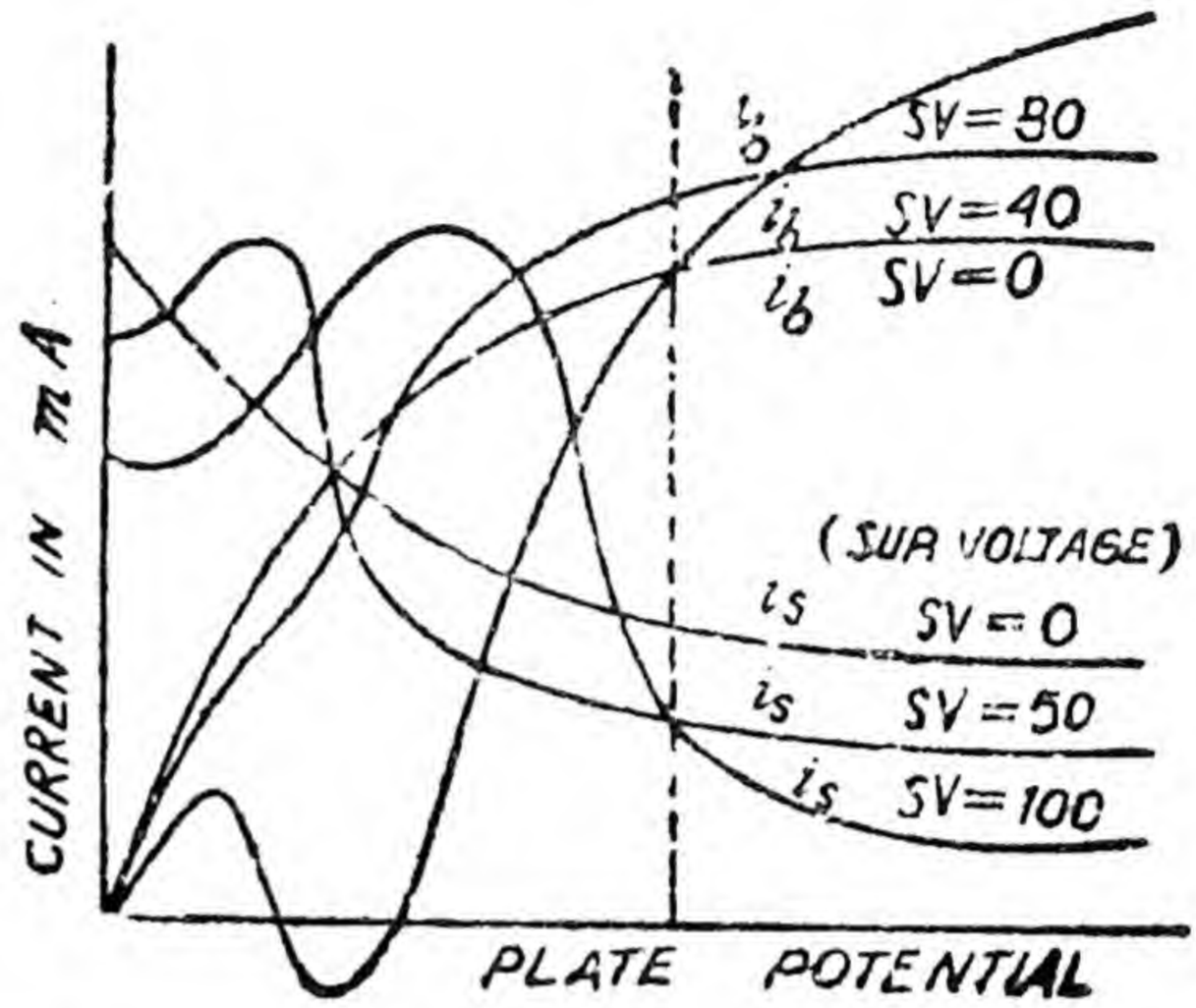


Fig. 18. Static characteristics of pentode showing the effect of suppressor grid voltage

Secondary Emission Ratio : The number of secondary electrons produced at an electrode is proportional to the number of arriving primary electrons and is affected by the electrode voltage and its surface condition. Therefore secondary emission coefficient N, for certain anode potential and surface condition, is defined as the ratio of the number of secondary electrons, n_s, emitted per second from anode to the number of primary electrons, n_p, arriving at the anode

or
$$N = \frac{n_s}{n_p}.$$

For the determination of secondary emission ratio, suppressor grid of tube 6SK7 is used to collect the secondary electrons emitted from the anode. The anode potential is fixed at a certain value (say 100 volts) and screen grid is given about half the potential of the anode. The control grid is kept at cathode potential. Suppressor grid potential is varied from —40 to 100 volts, *i.e.* upto plate potential.

The anode current and suppressor grid currents are plotted against suppressor grid voltage. The curves obtained are shown in figure 19. The straight part of suppressor current curve *AB*, *CD*, *EF* are extrapolated as shown in figure 19. *KL* represents the primary current to the suppressor grid and *KP* the total current in the anode circuit. *ML* represents the secondary emission current.

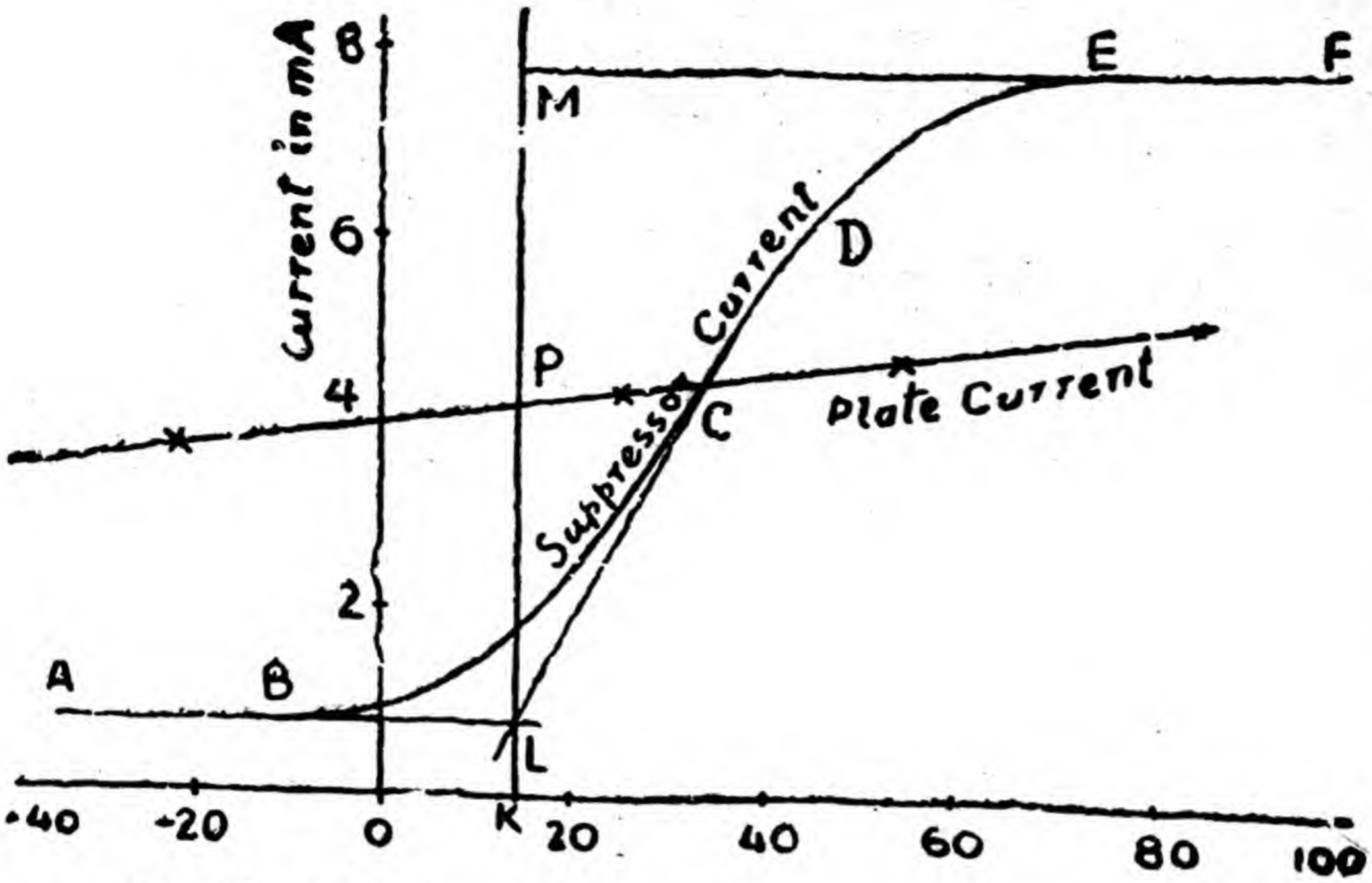


Fig. 19. Measurement of secondary emsision coefficient.

Since the currents will be of the same ratio as the number of electrons, secondary emission coefficient N is given by

$$N = \frac{ML}{LP}.$$

If the electrode surface has been specially treated to resist secondary emission, N *i e.* the ratio of secondary to primary electrons is much less than unity, being of the order of 0·2 to 0·1.

Uses of Pentode.

(1) As plate resistance and amplification factor are high, so a pentode can be used as voltage amplifier.

(2) Pentode can be used as mixer or frequency converter.

(3) Specially designed pentode (*i.e.*, keeping spacing between control grid and cathode small and pitch of the control grid wires small) can be used as a wideband amplifier or video amplifier in television and radar.

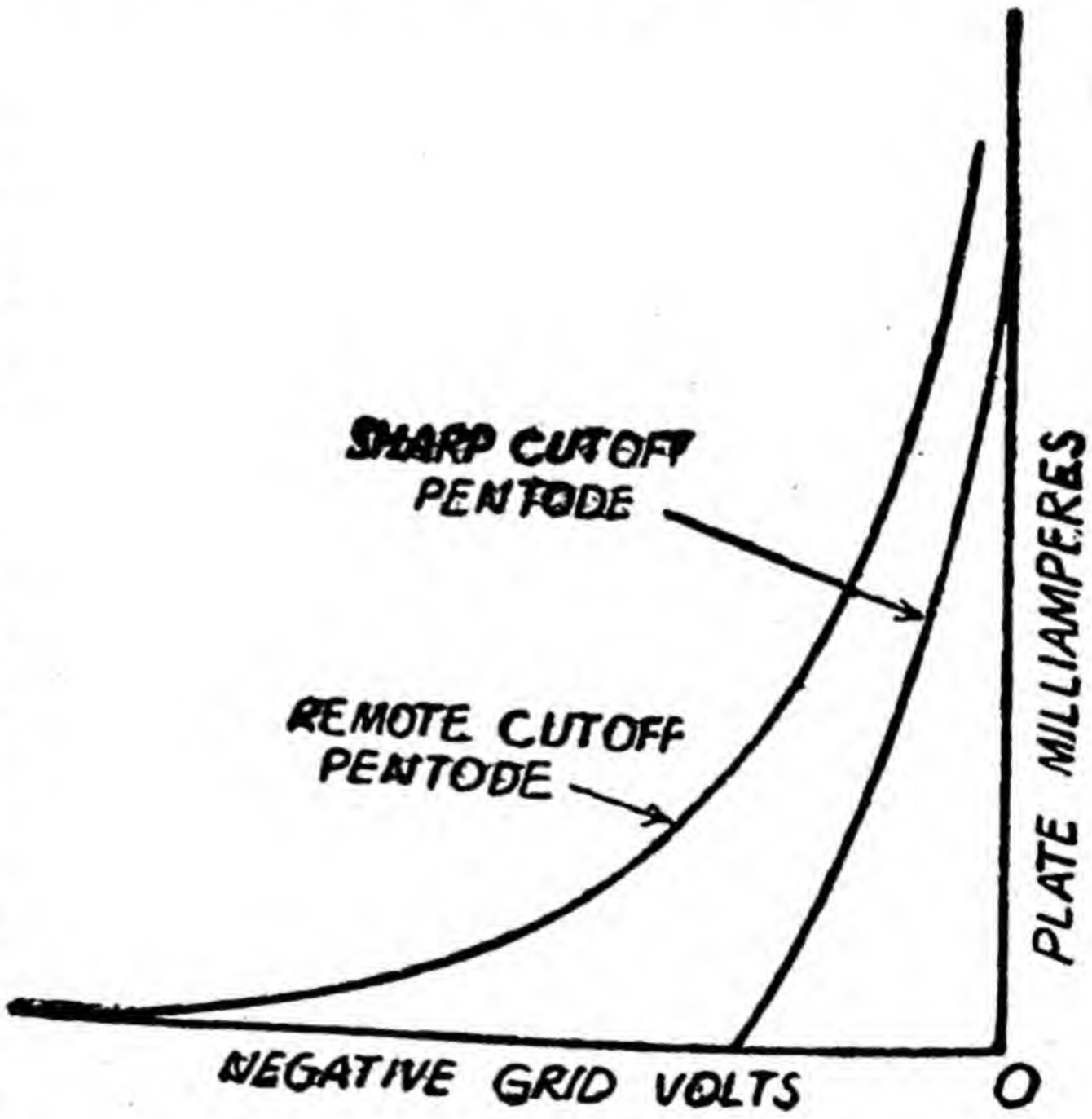


Fig. 20. Variable μ pentode characteristic as compared to sharp cut-off pentode characteristic.

(4) This can be used as suppressor grid modulator. In this

case the modulating voltage is applied at suppressor grid while the carrier is applied to the control grid. The suppressor grid is used for controlling the plate current.

5·3-1. VARIABLE PENTODES (Remote cut-off pentodes) : To accommodate large-amplitude signals without distortion which arise due to plate current cut-off, remote cut-off or the gradual cut-off characteristic of the tube, as shown in figure 20, is required. Since the amplification factor largely depends on the spacing of the control grid wires, a non-uniform control grid structure (varying pitch between grid wires) will result into the construction of variable-mu tube. Control wires are kept closer near the centre while high at the ends. When the control grid bias is increased, plate current cut off will reach at the control grid end regions due to high μ, while the central portion of the control grid (low μ) will still permit the flow of plate current A further increase in bias will also reduce the flow of electrons through the cental portion and consequently the amplification factor decreases considerably. Thus in order to obtain complete plate current cut-off, the bias should be very high.

In various circuits, it is desired to control the gain of the amplifying stage by the signal strength. This variable μ characteristic is helpful in providing such automatic gain controls.

5·4. BEAM POWER TUBE :

We have studied that a suppressor grid is used to reduce the effects of secondary emission in pentode valve. Beam power tube is a special type of tetrode which has general characteristics of pentode. In this tube the secondary emission is reduced by utili-

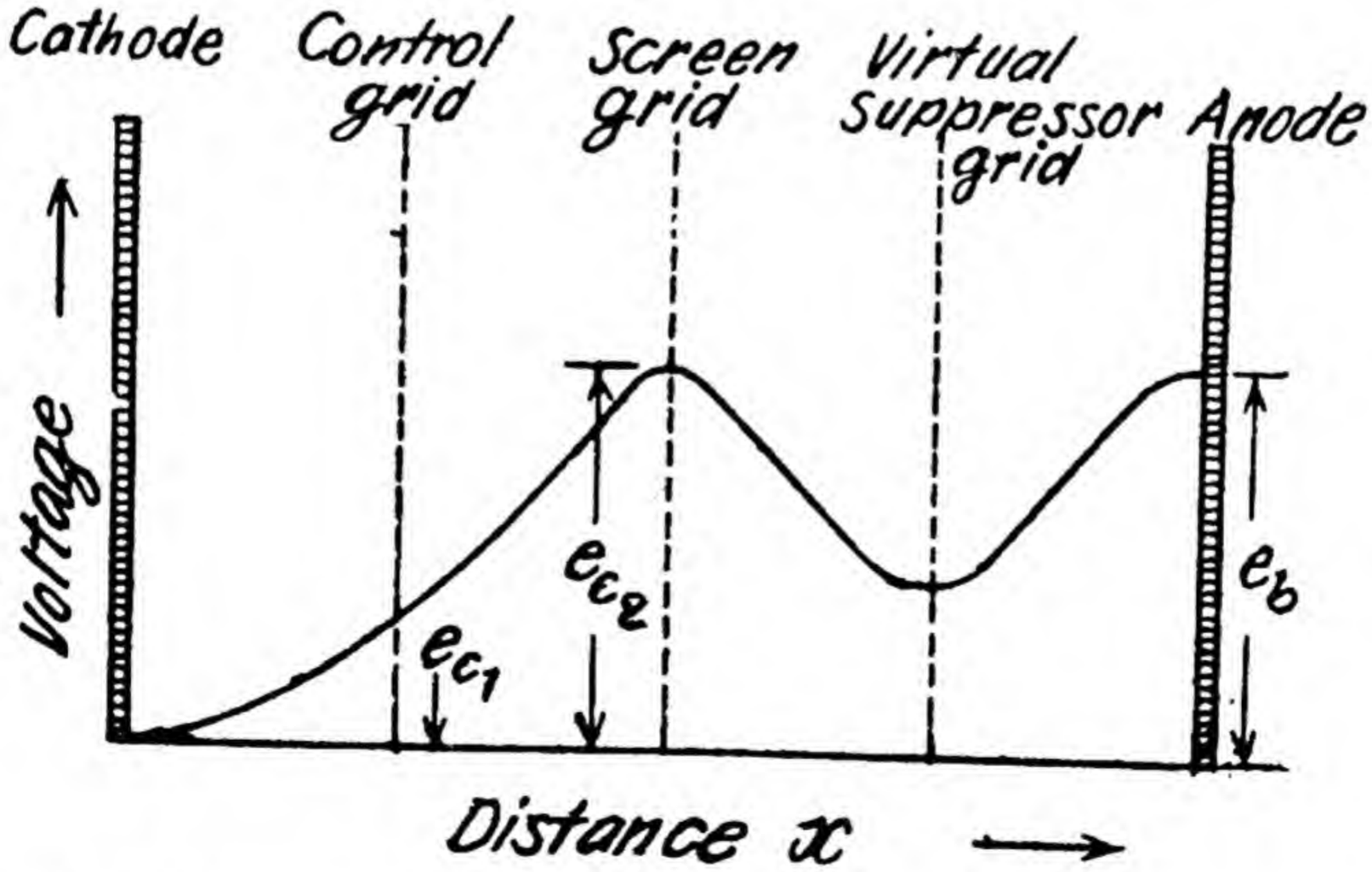


Fig. 21. Potential distribution curve of a beam power tube.

sing the space charge between screen grid and the plate The electrodes are spaced and designed in such a way that the space

charge at some point in the region between screen grid and plate reaches a very high negative value, thus reducing the potential at that point far below the plate potential. This low voltage surface acts as a suppressor grid and eliminates the effects of secondary emission. The potential distribution curve is shown in fig. 21.

Beam power tube is designed so that the electrons move from cathode to plate in dense sheets. The effect is achieved by making the control grid and screen grid of the same pitch and aligning the grid wires. The electrode structure of the tube is shown in fig. 22.

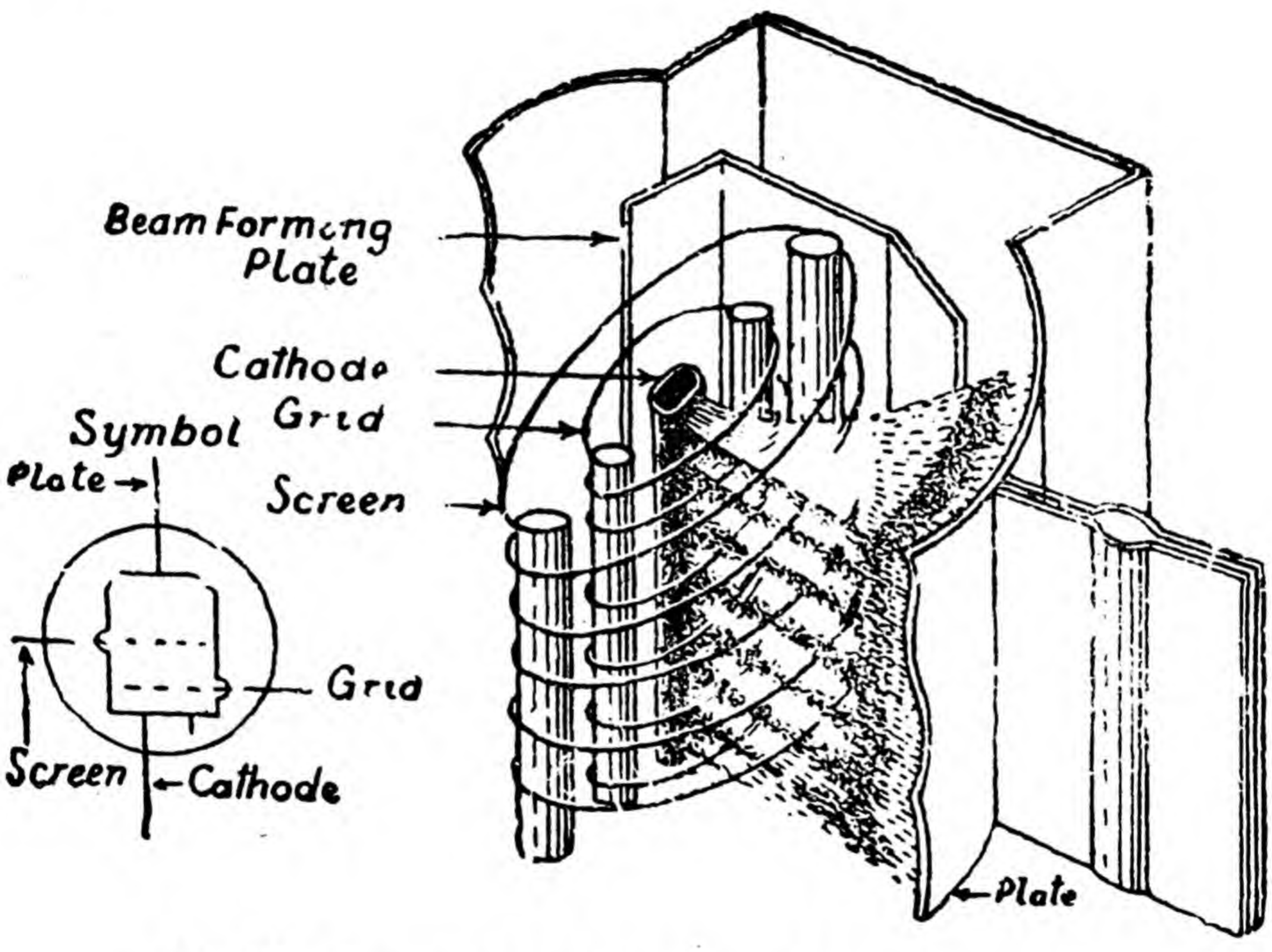


Fig. 22. Cut away view of the electrode arrangement in a beam power tube.

A negative potential is established in the region between the screen and the plate by securing a high electron density in this space. This dense negative charge sheet serves the same purpose as the suppressor grid in the conventional pentode. The potential between the screen grid and plate is depressed enough so that secondary electron flow from plate to grid is suppressed. Since there is no physical suppressor grid, the phenomenon of electron dispersion around suppressor grid wires is eliminated and hence the rounding of the knee of plate characteristic is avoided.

The plate characteristics of a beam power tube are shown in fig. 23. It is seen that the curves are essentially free from secondary

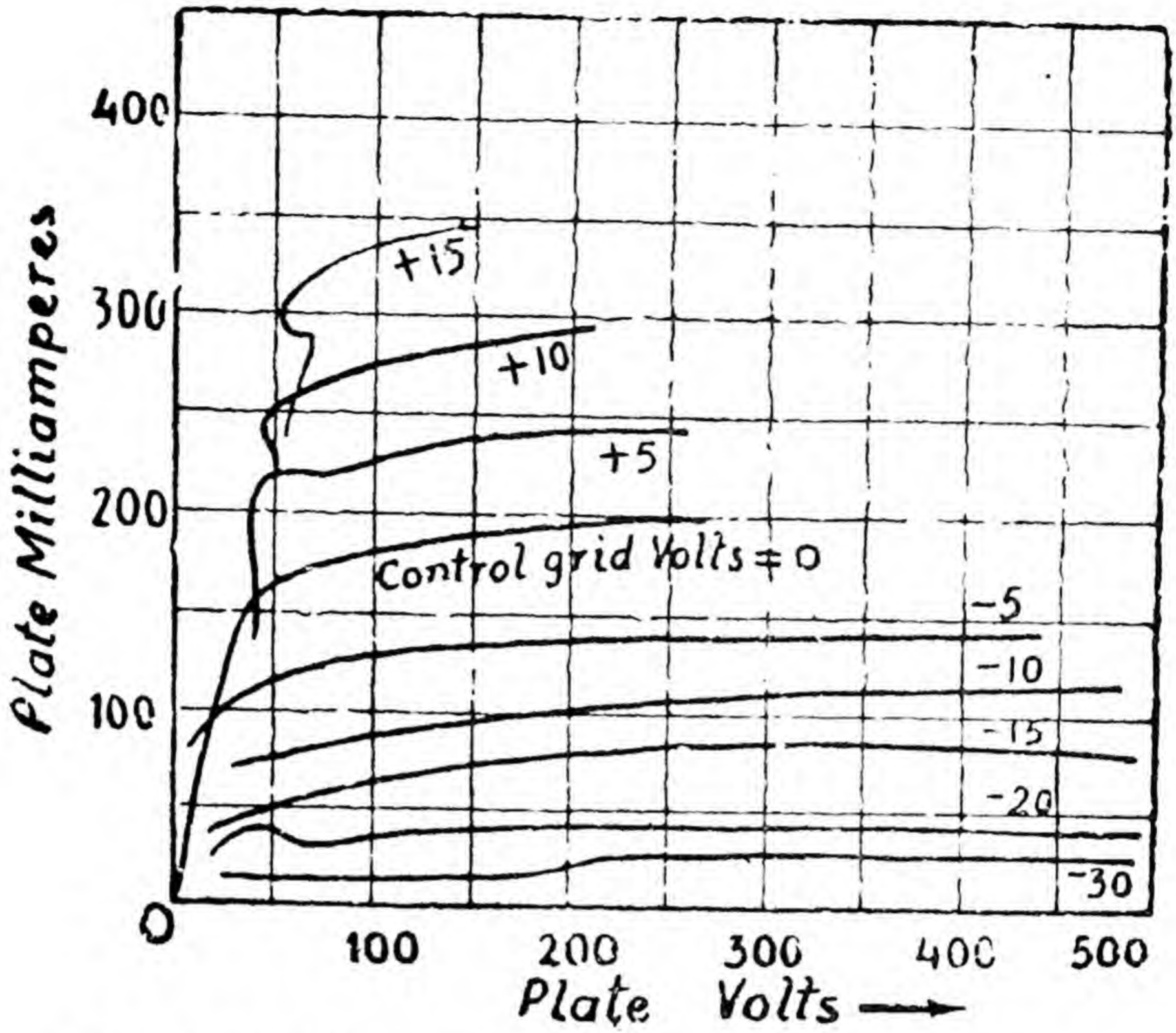


Fig. 23. Plate current-plate voltage characteristics of a beam power tube.

emission dips. Another difference from pentode characteristics is that the plate current rises more rapidly with plate voltage at low plate potentials. This is due to the relatively low screen current of this tube type. Beam power tubes are therefore highly efficient and are usable over a wide operating range.

5·5. CO-EFFICIENTS OF SCREEN GRID, BEAM AND PENTODE TUBES :

The constants required to define completely all the characteristics of the tube consisting four or more electrodes are large in number and can be divided into three types : the dynamic resistance of various electrodes (screen and plate). the mu-factor and the trans-conductance.

1. Dynamic electrode resistance : The resistance offered by the electrode to a small increase in applied voltage is called dynamic electrode resistance. For pentode, tetrode and beam power tube, it can be written as

Plate resistance $r_p = \left(\frac{de_b}{di_b}\right)$ e_{c_1} and e_{c_2} constant.

Screen resistance $r_s = \left(\frac{de_{c_2}}{di_{c_2}}\right)$ e_b and e_{c_1} constant,

where e_b, e_{c_1} and e_{c_2} represent the plate, control grid and screen grid potentials respectively i_b and i_{c_2} represent the plate and screen currents.

Plate resistance of pentode, beam and tetrode under usual operating conditions is very high (of the order of mega-ohms while the screen resistance is moderate (about four or five times the plate resistance of the same tube used as a triode).

2. Mu-factor : In pentodes or similar tubes, the relative effectiveness of control grid and screen voltage on the plate current is represented as μ factor. To express mathematically, μ factor of electrode 1 relative to electrode 2 with respect to current in electrode 3 in a given tube will be

$$\mu = \left(\frac{de_2}{de_1}\right)_{I_3 \text{ constant}}$$

where I_3 represents the currents of electrode 3. Generally in pentode, beam and tetrode tube, two μ-factors are of importance, one concerning the control of plate current and the other regarding the control of total space current by the control grid and screen voltage. The former one is called *amplification factor* and latter is termed as *cut-off amplification factor* because it determines the relative screen and control-grid potentials for which plate current cut-off is reached.

3. Transconductance : In a general way, the transconductance is defined as the change of current to electrode 2 caused by the increment in voltage applied to electrode 1, *i.e.*

$$g_m = \frac{dI_2}{de_1}.$$

The important transconductance, even in pentodes, tetrodes is from control grid to plate which is almost the same as that of a triode and has nearly the same numerical value

5·6. DESIGN CONSIDERATION OF POWER AND HIGH FREQUENCY TUBES :

At very high frequencies, the important effect—known as transit time effect (see oscillators)—limits the design for negative grid operation at such frequencies. The characteristics for the design are :

(1) to provide high transconductance in proportion to electrode capacities and to minimise transit time effect, the spacing between electrodes must be quite small ;

(2) to minimise power losses in the leads and lead inductance, short leads of relative large diameters should be applied,

(3) suitable electrode and lead arrangement, so as to facilitate operation with resonant transmission line,

(4) for the tubes to be used at high voltage, the structural arrangement should be such as to provide sufficient amount of insulation, and

(5) the tube should be designed such that there can be no stray electrons for which the ends of the tube structure should be closed.

The tubes used in power amplifier must bear the correct capability of dissipating plate and grid losses in proportion to the desired power for which cathode emission and the voltage which can be applied to anode with safety must be in proportion to the tube power output

In power tubes, during operation, the grid current flows which causes power dissipation at the grid of the tube. Consequently, the grid wire material chosen for power tubes should be tungsten, tantalum and molybdenum.

5·7. CATHODE RAY TUBE OSCILLOSCOPE :

Now a-days cathode ray oscilloscope is used for a number of purposes such as measurement of current, voltage, observation of waveforms of alternating voltages, recreation of television images, as indicator in radar for usual presentation of target data such as distance, height etc.

It consists of the following main constituents :

1. Cathode ray tube.
2. Power supplies.
3. Time base circuit.
4 Deflection voltage amplifiers.

We shall discuss these constituents In cathode ray tube, the deflection of electron is governed by electric and magnetic fields and hence we would first like to discuss the motion of electron in electric and magnetic fields.

5·7-1. MOTION OF AN ELECTRON IN PARALLEL PLATE DIODE :

(a) Zero initial velocity :

Let figure 24 show the simplest diode configuration having parallel plates at a distance d apart. Let a potential difference V be maintained between the two plates such that the direction of electric field is in the negative x-direction. As the electron is negatively charged, it tends to flow from cathode to plate. Our problem is to consider the motion of an electron under the following boundary conditions :

velocity of the electron

$$\begin{cases} v=\dfrac{dx}{dt}=0 \\ \text{and } x=0 \text{ when } t=0. \end{cases}$$

We know that the potential V and the electric intensity E are related by the relation

$$E=-\frac{dV}{dx}$$

or $$dV=-E\,dx. \qquad \ldots(1)$$

Again the difference of potential is defined as the amount of work done in moving a unit

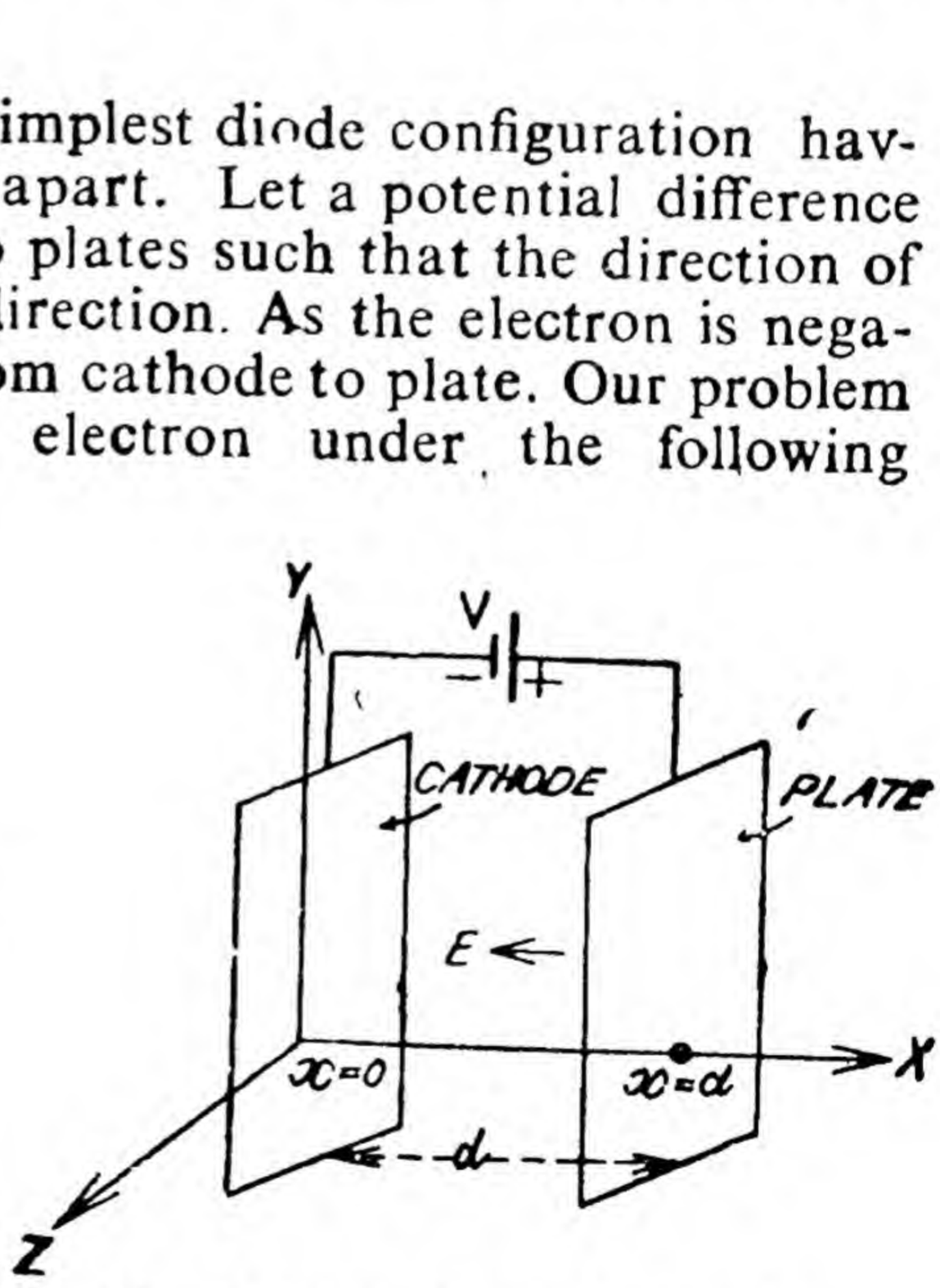


Fig 24. Motion of an electron in parallel plate diode.

positive charge from one plate to another *i.e.*,

$$\int dV = V = -\int_0^d E\,dx = -Ed$$

or
$$E = -\frac{V}{d}. \qquad \ldots(2)$$

The force exerted on the electron having charge Q_e ($-1{\cdot}6\times10^{-19}$ coulomb) in electric field E will be

$$F = Q_e.E = -Q_e\,\frac{V}{d} \qquad \ldots(3)$$

But according to Newton's second law of motion

Force = mass × acceleration

$$F = m \times \frac{d^2x}{dt^2}, \qquad \ldots(4)$$

where m is the mass of the electron.

From equations (3) and (4), we have

$$m\,\frac{d^2x}{dt^2} = -Q_e\,\frac{V}{d}$$

or
$$\frac{d^2x}{dt^2} = -\frac{Q_e}{m}\,\frac{V}{d}. \qquad \ldots(5)$$

Also
$$\frac{d^2y}{dt^2} = \frac{d^2z}{dt^2} = 0. \qquad \ldots(6)$$

Integrating equation (5), we get

$$v = \frac{dx}{dt} = -\frac{Q_e.V}{m.d}.t + C_1, \qquad \ldots(7)$$

where C_1 is a constant of integration and can be obtained by applying the boundary condition *i.e.*, $v=0$ when $t=0$. The value of C_1 comes out to be zero. Hence

$$v = \frac{dx}{dt} = -\frac{Q_e}{m}\,\frac{V}{d}.t. \qquad \ldots(8)$$

Integrating once again, we have

$$\int dx = \int -\frac{Q_eV}{md}.t\;dt$$

$$x = -\tfrac{1}{2}\,\frac{Q_eVt^2}{m.d} + C_2. \qquad \ldots(9)$$

Again $C_2=0$ as $x=0$ when $t=0$.

Thus the distance traversed by the electron in time t is given by

$$x = -\tfrac{1}{2}\,\frac{Q_e.V.t^2}{m.d} \qquad \ldots(10)$$

The time t_p taken by the electron to move from cathode to plate can be obtained by putting $x=d$ and $t=t_p$ in equation (10),

i.e., $$d=-\tfrac{1}{2}\frac{Q_e V}{m}\frac{t_p^2}{d}$$

or $$t_p=\left(\frac{2m}{-Q_e.V}\right)^{1/2} d. \qquad \ldots(11)$$

Similarly, the velocity v_p with which the electrons strikes the plate can be obtained by putting $t=t_p$ and $v=v_p$ in equation (8),

i.e. $$v_p=-\frac{Q_e}{m}\frac{V}{d}.\ t_p$$

Substituting the value of t_p from equation (11), we get

$$v_p=\left(-\frac{Q_e V}{m\ d}\right)\left(\frac{2m}{-Q_e\ V}\right)^{1/2} d.$$

$\because$ $$v_p=\left(\frac{-2Q_e V}{m}\right)^{1/2} \qquad \ldots(12)$$

The kinetic energy with which the electron strikes the plate is

$$\tfrac{1}{2}mv_p^2=\tfrac{1}{2}\ m\times\left(\frac{-2Q_e V}{m}\right)=-Q_e V. \qquad \ldots(13)$$

For an electron $Q_e=-1{\cdot}6\times10^{-19}$ coulomb, and

$$m=9{\cdot}11\times10^{-31}\ \text{K. gm.}$$

$$\therefore\ v_p=5{\cdot}93\times10^5.V^{1/2}\ \text{meter/sec.}$$

(v) With a definite initial velocity in the direction of anode :

Let v_i be the initial velocity of the electron in the direction from cathode to plate. The boundary conditions are

x, y, z, $\frac{dy}{dt}$ and $\frac{dz}{dt}$ are zero at time $t=0$

$\frac{dx}{dt}=v_i$ at time $t=0$

Now from equation (5), we have

$$\frac{d^2x}{dt^2}=-\frac{Q_e V}{md}.$$

Integrating this equation, we have

$$\frac{dx}{dt}=-\frac{Q_e}{m}.\frac{V}{d}.\ t+C_3 \qquad \ldots(14)$$

where C_3 is the constant of integration.

Now $\frac{dx}{dt}=v_i$ when $t=0$.

Hence $C_3=v_i$.

Substituting this value in equation (14), we have

$$v=\frac{dx}{dt}=-\frac{Q_e}{m}\frac{V}{d}.t+v_i. \qquad \ldots(15)$$

Again integrating, we have

$$x=-\frac{Q_e}{m}.\frac{V}{d}.\frac{t^2}{2}+v_i\ t+C_i.$$

But $x=0$ when $t=0$, hence $C_4=0$.

$$\therefore \quad x=-\frac{Q_e}{m}\cdot\frac{V}{d}\cdot\frac{t^2}{2}+v_i t. \qquad \ldots(16)$$

v_p may be determined by applying the principle of conservation of energy, *i.e.*,

$$\tfrac{1}{2}\,mv_p^2=-Q_eV+\tfrac{1}{2}\,mv_i^2$$

or

$$v_p=\left(-\frac{2Q_eV}{m}+v_i^2\right)^{1/2} \qquad \ldots(17)$$

Average velocity from cathode to anode is given by

$$\bar{v}=\frac{v_p+v_i}{2}.$$

$\therefore$ time of flight from cathode to anode

$$t_p=\frac{d}{(v_p+v_i)/2} \qquad \ldots(18)$$

Example 1. *In a diode, the anode is at* 120 *volts with respect to the cathode and spacing between the anode and cathode is* 5 *mm. Assuming an electron to be emitted with zero initial velocity, calculate the velocity and distance travelled after* $0{\cdot}4\times10^{-9}$ *second. Also calculate the above quantities if the initial velocity of emission in the direction of anode is* $1{\cdot}5\times10^{6}$ *metres/sec.*

(*a*) Zero initial velocity :

The velocity v of the electron after a time t sec is given by

$$v=-\frac{Q_e}{m}\frac{V}{d}.t.$$

Now for an electron $Q_e=-1{\cdot}6\times10^{-19}$ coulomb,

$$m=9{\cdot}11\times10^{-31}\text{ k. gm.}$$

According to the given problem :

$V=120$ volts, $d=5$ m.m. $=5\times10^{-3}$ meters, $t=0\,4\times10^{-9}$ se

$$\therefore \quad v=\frac{1{\cdot}6\times10^{-19}\times120\times0{\cdot}4\times10^{-9}}{9{\cdot}11\times10^{-31}\times5\times10^{-3}}$$

$$=1{\cdot}69\times10^{6}\text{ meters/sec.}$$

The distance x is given by

$$x=-\frac{Q_e}{m}\frac{V}{d}\cdot\frac{t^2}{2}\cdot$$

$$=\frac{1{\cdot}6\times10^{-19}\times120\times(0{\cdot}4\times10^{-9})^2}{9{\cdot}11\times10^{-31}\times5\times10^{-3}\times2}$$

$$=0{\cdot}338\times10^{-3}\text{ meters.}$$

(*b*) With initial velocity $v_i=1{\cdot}5\times10^{6}$ meters/sec.

$$v=-\frac{Q_e}{m}\frac{V}{d}.t+v_i$$

$$= \frac{1{\cdot}6\times10^{-19}\times120\times(0{\cdot}4\times10^{-9})}{9{\cdot}11\times10^{-31}\times5\times10^{-3}} + (1{\cdot}5\times10^{6})$$

$$= \mathbf{3{\cdot}19\times10^{2}\ \text{meters/sec.}}$$

and
$$x = -\frac{Q_e}{m}\frac{V}{d}\frac{t^2}{2} + v_i t$$

$$= \frac{1{\cdot}6\times10^{-19}\times120\times(0{\cdot}4\times10^{-9})^2}{9{\cdot}11\times10^{-31}\times5\times10^{-3}\times2} - $$

$$+ (1{\cdot}5\times10^{6}\times0{\cdot}4\times10^{-9})$$

$$= \mathbf{0{\cdot}938\times10^{-3}\ \text{meters.}}$$

Example 2. *An electron starts at rest from the negative plate of a plane parallel capacitor, potential across which is* 1,000 *volts. The distance between the plate is* 1·5 *cm. How far has the electron travelled when it acquires a speed of* 10^7 *m/sec. If at this instant the potential on the plates is reversed in sign and changed to* 1,500 *volts, find the voltage against which the electron would move.*

We know that

$$v = \sqrt{\left(\frac{2\,(-Q_e)\,V}{m_e}\right)} \qquad \ldots(1)$$

The electric field strength

$$E = \frac{V}{d} = \frac{1000\text{ volt}}{1{\cdot}5\times10^{-2}\text{ metre}}$$

Let the distance travelled by the electron when it has acquired a speed of 10^7 *m*/sec. is x. Then

$$v = \sqrt{\cdot\left(\frac{2\,(-Q_e)}{m_e}E.x\right)}$$

or
$$0^7 = \left(2\times1{\cdot}76\times10^{11}\times\frac{1000}{1{\cdot}5\times10^{-2}}.x\right)^{1/2}$$

Solving for x, we have

$$x = \mathbf{0{\cdot}42\ cm.}$$

When the potential is reversed, the electron is acted upon by a retarding force opposite to positive direction of x. Now

$$E_x = \frac{V}{d}.x = \frac{1000}{1{\cdot}5\times10^{-2}}\times\frac{0{\cdot}42}{100} = 280\text{ volts}$$

Where E_x refers to the electric field strength at a distance x from the plate. When the potential is reversed, the electron moves against a p.d. of (1500 − 280)

$$= \mathbf{1220\ volts.}$$

Example 3. *An electron enters through a hole from one of the plates of a parallel plate capacitor at right angles to the plate where its motion is retarded by an electric field as shown in figure* 24. *Calculate the retarding fields to bring it to rest in a distance of* 4 *cm. if* (*a*) *the electron is travelling with initial velocity of* 6×10^7 *m./sec.*

(*b*) *the initial velocity of the electron corresponds to a kinetic energy of* 100 *eV*.

When the electron enters the plate it is acted upon by a retarding force in the opposite direction of x. The equation of motion is

$$m\frac{d^2x}{dt^2}=-Q_e E$$

or
$$v.\frac{dv}{dx}=-\frac{Q_eE}{m}$$

Integrating, we have

$$\frac{v^2}{2}=-\frac{Q_eE}{m}x+C.$$

Assuming that the electron enters with a velocity u, the initial conditions at $t=0$ are $x=0$ and $v=u$.

$$\therefore \quad C=\frac{u^2}{2}$$

The velocity of the electron at a distance x from the left hand plate is given by

$$\frac{v^2}{2}=-\frac{Q_eE}{m}x+\frac{u^2}{2}$$

or
$$v^2=-\frac{2Q_e}{m}Ex+u^2$$

(a) When $x=4$ cm., $v=0$ and $u=6\times10^7$ m/sec.

$$\therefore \quad u^2=\frac{2Q_eEx}{m}$$

or
$$E=\frac{mu^2}{2Q_ex}=\frac{36\times10^{14}}{2\times176\times10^{11}\times0{\cdot}04}$$
$$=256\times10^2 \text{ volt/metre.}$$

(b) The initial velocity of the electron corresponding to 100 *eV* is given by

$$\tfrac{1}{2}mu^2=100\times Q_e \text{ joules}$$
$$u=\sqrt{[2\times100\times(Q_e/m)]}$$
$$=\sqrt{(2\times100\times1{\cdot}76\times10^{11})}$$
$$=5{\cdot}94\times10^6 \ m/\text{sec.}$$
$$\therefore \quad E=\frac{mu^2}{2Q_ex}=\frac{(5{\cdot}94\times10^6)^2}{2\times1{\cdot}76\times10^{11}\times0{\cdot}04}$$
$$=250 \text{ KV/metre.}$$

Example 4. *In a high vacuum parallel plane plate diode the spacing between the anode and cathode is* 8 *mm. A square wave alternating voltage of amplitude* 1 *volt and frequency* 5×10^7 *cycles/second is applied between the electrodes. An electron starts from rest at the cathode surface at the instant when anode to cathode a.c.*

voltage is zero and is going in the positive direction. Find (i) the velocity and displacement of electron from cathode at the end of first, second etc. half cycles and (ii) the velocity and time of travel of electron on reaching the anode.

The square wave of amplitude 1 volt is shown in fig. (25).

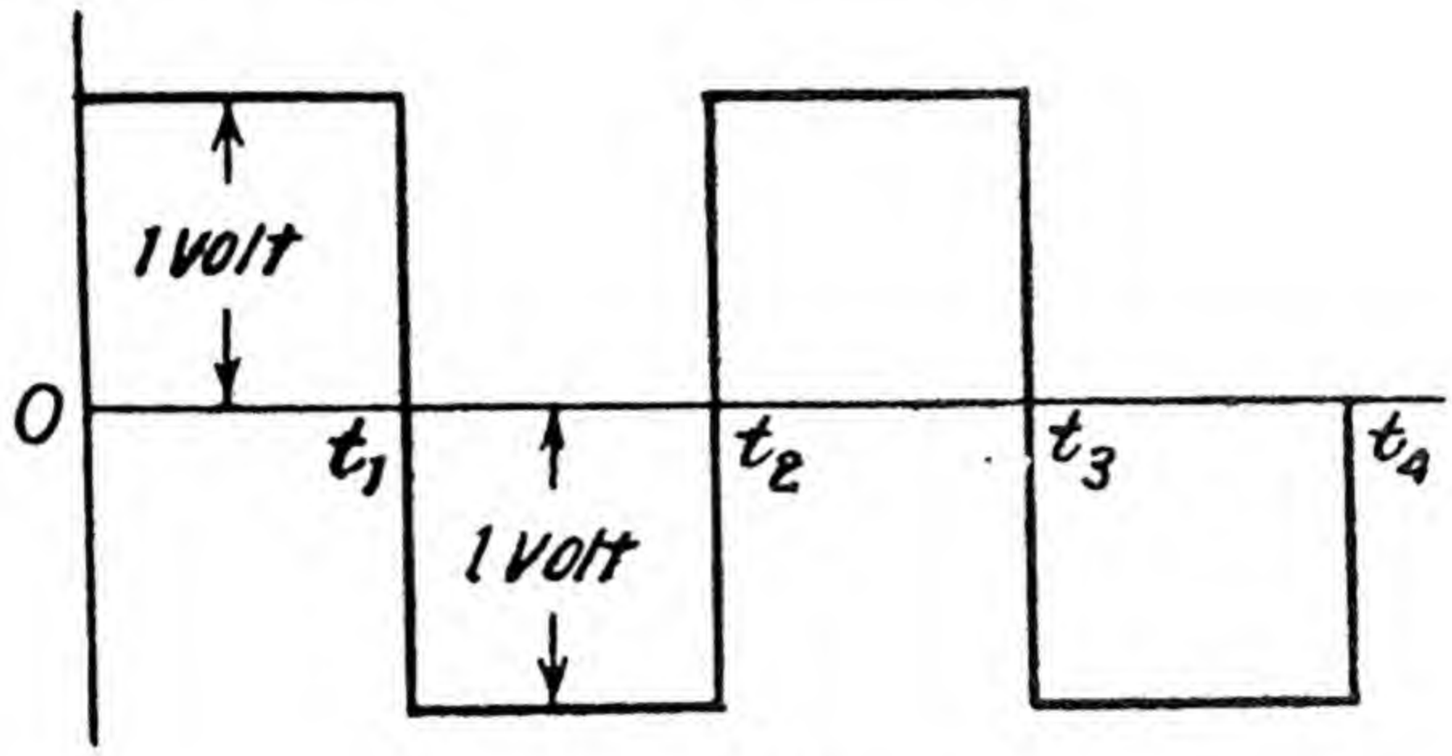


Fig. 25

Half of the periodic time

$$= \frac{1}{2 \times 5 \times 10^7} = 10^{-8} \text{ second.}$$

At the end of first half cycle *i.e.*, at t_1

velocity $$v_1 = -\frac{Q_e}{m}\frac{V}{d}.t$$

$$= +\frac{1{\cdot}6 \times 10^{-19}}{9{\cdot}1 \times 10^{-31}} \times \frac{1}{8 \times 10^{-3}} \times 10^{-8} \text{ metres/sec.}$$

$$= 0{\cdot}22 \times 10^6 \text{ metres/second.}$$

Distance travelled during first half cycle

$$x_1 = -\tfrac{1}{2}\frac{Q_e V}{m\, d}.t^2$$

$$= -\tfrac{1}{2}\frac{(-1{\cdot}6 \times 10^{-19})}{9{\cdot}1 \times 10^{-31}} \times \frac{1}{(8 \times 10^{-3})} \times (10^{-8})^2 \text{ metres}$$

$$= 1{\cdot}1 \text{ mm.}$$

At time t_2. velocity $v_2 = 0$ and distance travelled $x_2 = 2x_1 = 2{\cdot}2$ m.m. because there is retardation equal to the acceleration during first positive half cycle.

Similarly,

at time t_3, $v_3 = v_1 = 0{\cdot}22 \times 10^6$ metres/sec. and $x_3 = 3x_1 = 3{\cdot}3$ mm.

at time t_7, $v_7 = v_1 = 0{\cdot}22 \times 10^6$ metres/sec. and $x_7 = 7x_1 = 7{\cdot}7$ mm.

The remaining distance $= (8 - 7{\cdot}7) = 0{\cdot}3$ mm $= 3 \times 10^{-4}$ metre.

For this distance, we have

$$\text{distance} = -\frac{Q_e}{m}\frac{V}{d} \cdot \frac{t^2}{2} + v_i\, t \quad \text{where } v_i = v_7$$

$$\therefore \quad 3\times10^{-4}=-\left(\frac{-1{\cdot}6\times10^{-19}}{9{\cdot}1\times10^{-31}}\right)\cdot\frac{1}{8\times10^{-3}}\cdot\frac{t^2}{2}$$
$$+0{\cdot}22\times10^{6}\ t$$

Solving this equation for t, we get

$$t=1{\cdot}852\times10^{8}\text{ second}\quad\text{or}\quad 0{\cdot}1474\times10^{-8}\text{ sec.}$$

The first value is inadmissible, hence $t=0{\cdot}1474\times10^{-8}$ sec.
Now total time $=\mathbf{7{\cdot}1474\times10^{-8}}$ **second.**
Velocity on reaching the anode

$$=-\frac{(-1{\cdot}6\times10^{-19}}{9{\cdot}1\times10^{-31}}\times\frac{(-1)}{3\times10^{-8}}\times0{\cdot}1474\times10^{-8}$$
$$+(0{\cdot}22\times10^{6})\text{ metres/sec.}$$

$$=\mathbf{1{\cdot}876\times10^{5}\ metres/sec.}$$

Example 5. *In a high vaccum parallel plate diode, the distance between anode and cathode is* 1 *cm. Between anode and cathode is applied a square wave alternating voltage of amplitude* 1 *volt and frequency* 100 *MHz. An electron starts from rest at the cathode at the instant anode to cathode a.c. voltage is zero and is going in the positive direction. Find the position and velocity of this electron at the end of* (*i*) *positive half cycle and* (*ii*) *subsequent negative half cycle. Find also above quantities if a* 3 *volts d.c. voltage is connected between anode and cathode* (*positive at the anode*) *in series with the square wave a.c. voltage.*

(1) With square wave A.C. voltage only.

(i) Since the frequency of a.c. voltage $=100\ MHz$, the half time period of the wave is 5×10^{-9} second.

During positive half cycle, the displacement x of the electron is given by

$$x=-\tfrac{1}{2}\frac{Q_e}{m}\cdot\frac{V}{d}\cdot t^2$$
$$=-\tfrac{1}{2}\frac{(-1{\cdot}6\times10^{-19})}{(9{\cdot}1\times10^{-31})}\times\frac{1}{(10^{-2})}\times(5\times10^{-9})^2$$
$$=\mathbf{0{\cdot}2199\ mm}$$

velocity after first half cycle is given by

$$v=-\frac{Q_e}{m}\cdot\frac{V}{d}\cdot t=-\frac{(-1{\cdot}6\times10^{-19})}{(9{\cdot}1\times10^{-31})}\times\frac{1}{(10^{-2})}$$
$$\times(5\times10^{-9})$$
$$=\mathbf{8{\cdot}795\times10^{4}\ m/s.}$$

(ii) During the negative half cycle, the retardation is equal and opposite. The distance travelled during this negative half cycle is the same as travelled during the positive half cycle. Hence during the end of second half cycle, the total distance covered is $2\times0{\cdot}2199\text{ mm}=0{\cdot}4398$ mm and velocity becomes zero.

(2) With superimposed D.C. voltage

(i) During the first positive half cycle, the net anode to cathode voltage is 4 volts while during negative half cycle, the voltage is 2 volt.

During positive half cycle, the displacement is given by

$$x_1 = -\tfrac{1}{2}\,\frac{Q_e}{m}\cdot\frac{V}{d}\cdot t^2$$

$$= -\tfrac{1}{2}\,\frac{(-1{\cdot}6\times10^{-19})}{(9{\cdot}1\times10^{-31})}\cdot\frac{4}{(10^{-2})}\cdot(5\times10^{-9})^2$$

$$= \mathbf{0{\cdot}8796\ mm.}$$

Velocity at the end of positive half cycle

$$v_1 = -\frac{Q_e}{m}\cdot\frac{V}{d}\cdot t = -\frac{(-1{\cdot}6\times10^{-19})}{(9{\cdot}1\times10^{-31})}\times\frac{4}{10^{-2}}\times(5\times10^{-9})$$

$$= \mathbf{35{\cdot}180\times10^4\ m/s}$$

During the negative half cycle, the initial velocity

$$v_1 = 35{\cdot}180\times10^4\ \text{m/s}$$

The initial displacement $x_1 = 0{\cdot}8796$ mm.

∴ Velocity at the end of negative cycle

$$v_2 = -\frac{Q_e}{m}\cdot\frac{V}{d}\cdot t + v_1$$

$$= (1{\cdot}76\times10^{11})\ \frac{2}{(10^{-2})}\times(5\times10^{-9}) + 35{\cdot}180\times10^4\ \text{m/s}$$

$$= \mathbf{52{\cdot}77\times10^4\ m/s}$$

Distance travelled at the end of negative half cycle

$$x_2 = -\frac{Q_e}{m}\times\frac{V}{d}\cdot\frac{t^2}{2} + v_1 t + x_1$$

$$= (1{\cdot}76\times10^{11})\times\frac{2}{(10^{-2})}\cdot\frac{(5\times10^{-9})^2}{2} + (35{\cdot}180\times10^4)\ (5\times10^{-9}) + 0{\cdot}8796\ \text{mm.}$$

$$= 2{\cdot}1967\ \text{mm} + 0{\cdot}8796\ \text{mm.}$$

$$= \mathbf{3{\cdot}0763\ mm.}$$

5·7-2. MOTION OF AN ELECTRON IN A CONSTANT MAGNETIC FIELD :

Let the uniform magnetic field be applied on the right hand side of the line OO' such that the flux density B is perpendicular to the plane of the paper in which the electron moves with a constant velocity v.

As a moving charge constitutes a current, the moving electron constitutes a minute electric current of magnitude $(-vQ_e)$ and is deflected by uniform magnetic field. The direction of deflection is given by Flaming's left hand

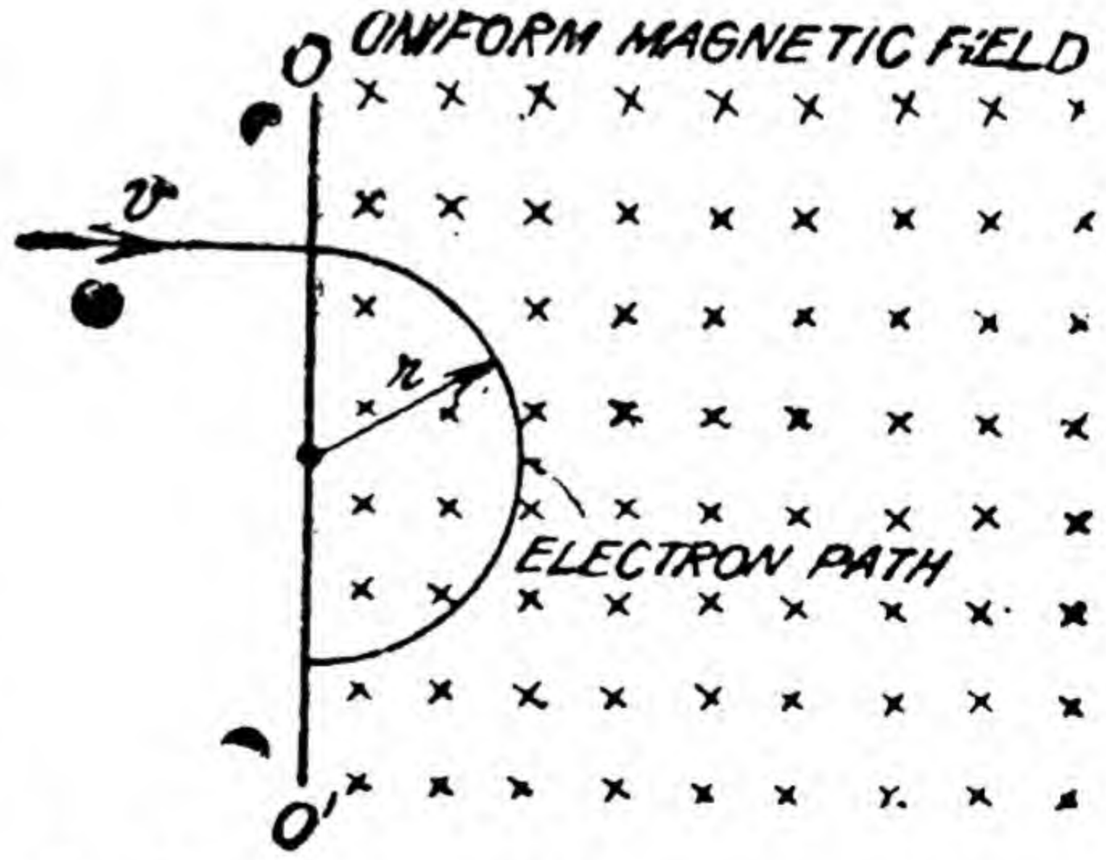


Fig. 26. Path of electron in uniform magnetic field.

rule, which is perpendicular to the plane containing the direction of the magnetic field. As the electron is deflected in a direction perpendicular to its instantaneous direction of motion, it describes a circular path.

Since the force on the electron, due to the magnetic field is perpendicular to the direction of motion of the electron, no work is done by this force on the electron. Due to this fact the kinetic energy and speed of the charged particle do not change.

The force acting on the charged particle is given by

$$F = -BQ_e v \qquad \ldots(1)$$

where v is the velocity of the particle.

The acceleration due to this force is v^2/r, where r is the radius of curvature of the path. Further

$$F = \frac{mv^2}{r} \qquad \ldots(2)$$

Comparing equations (1) and (2), we have

$$\frac{mv^2}{r} = -BQ_e v.$$

or

$$r = -\frac{mv}{BQ_e}. \qquad \ldots(3)$$

As the velocity of the electron remains constant, Q_e and m are also constant, the radius r is constant provided B is constant. Thus the path traced by the charged particle is a circle.

Now

$$\text{angular velocity} = \frac{\text{linear velocity}}{\text{radius of curvature of path}}$$

$$\therefore \quad \omega = \frac{v}{r} = -\frac{Q_e B}{m} \text{ radians/sec.} \qquad \ldots(4)$$

The time period T for a complete revolution is

$$T = \frac{1}{f} = \frac{2\pi}{2\pi f} = \frac{2\pi}{\omega}$$

$$= \left(\frac{2\pi m}{-Q_e B}\right) \text{ second.} \qquad \ldots(5)$$

We know that for an electron

$$m = 9{\cdot}11 \times 10^{-31} \text{ k. gm.}$$

$$Q_e = -1{\cdot}6 \times 10^{-19} \text{ coulomb.}$$

$$\therefore \quad T = \frac{2 \times 3{\cdot}14 \times 9{\cdot}11 \times 10^{-31}}{1{\cdot}6 \times 10^{-19} \times B} = \frac{35{\cdot}7 \times 10^{-12}}{B} \text{ second,}$$

where B is the magnetic flux density expressed in Weber's per square meter.

Example 1. *An electron moving with a velocity* 10^8 *meters per sec. enters a magnetic field at an angle of* $20°$ *to the direction of the field. Calculate*

(a) the flux density so that the helical path radius be 2 meters,
(b) the time required to execute one revolution of the helical path, and
(c) the pitch of the spiral.

(a) The force acting on a charged particle is given by

$$F=B\,(-Q_e)\,v=m_e\,a$$

$$\therefore \quad a=\frac{v^2\sin^2\theta}{r}$$

$$\therefore \quad m_e\,a=\frac{m_e\,v^2\sin^2\theta}{v}=B\,(-Q_e)\,v\sin\theta$$

or
$$B=\frac{m_e\,v\sin\theta}{r\,(-Q_e)}=\frac{9\times10^{-31}\times10^8\sin 20}{2\times1\,6\times10^{-19}}$$
$$=\mathbf{9{\cdot}6\times10^{-5}\ wb./sq.\ m.}$$

(b) Velocity component perpendicular to the magnetic field

$$v_y=v\sin 20^\circ=10^8\times0{\cdot}342=3{\cdot}42\times10^7\ \text{m/sec.}$$

velocity component parallel to magnetic field

$$v_y=v\cos 20^\circ=10^8\times0{\cdot}9397=9{\cdot}347\times10^7\ \text{m/sec.}$$

$$T=\frac{2\pi m_e}{B\,(-Q_e)}=\frac{2\pi}{9{\cdot}6\times10^{-5}}\times\frac{1}{1{\cdot}76\times10^{11}}=\mathbf{3{\cdot}67\times10^{-7}\ sec.}$$

(c) Pitch of the spiral $=v_y t=v\cos 20\times t$
$$=9{\cdot}347\times10^7\times3{\cdot}67\times10^{-7}$$
$$=\mathbf{34{\cdot}51\ metres.}$$

Example 2. *An electron is accelerated in an electrostatic field and then projected at right angles to the magnetic field of 50 Wb./sq. m. as shown in figure 20. Calculate*

(i) the accelerating voltage so that the electron revolves in a path of 3·2 cm. radius, and

(ii) the angular frequency of revolution.

(i) We know that the velocity of the electron after emerging out of the electrostatic field is

$$v=\sqrt{\left(\frac{2\,(-Q_e)\,V}{m_e}\right)}$$
$$=\sqrt{(2V.1{\cdot}7\times10^{11})}=0{\cdot}593\times10^6\ \sqrt{V}.m/\text{sec.}$$

The radius of the path is given by

$$r=\frac{m_e\,v}{B(-Q_e)}$$

$$\therefore \quad v=\frac{B\,(-Q_e)\,r}{m_e}$$

or $\quad 0{\cdot}593\times10^6\ \sqrt{V}=50\times10^{-3}\times1{\cdot}76\times10^{11}\times0{\cdot}002$

$$V=\left[\frac{50\times10^{-3}\times1{\cdot}76\times10^{11}\times0{\cdot}002}{0{\cdot}593\times10^6}\right]^2$$
$$=[29{\cdot}7]^2=\mathbf{880\ volts.}$$

(ii) $$\omega=\frac{(-Q_e)\,B}{m_e}=1{\cdot}76\times10^{11}\times50\times10^{-3}$$

$$=88\times10^{8}\text{ radians/sec.}$$

5·7-3. MOTION OF AN ELECTRON IN ELECTRIC AND MAGNETIC FIELDS AT RIGHT ANGLE TO EACH OTHER.

We shall now consider the motion of an electron in electric and magnetic fields at right angle to each other. Let the electron be situated initially at origin and at rest so that the components of velocity in three directions are zero. Thus the boundary conditions are

(i) $x=y=z=0$ when $t=0$

(ii) $v_x=v_y=v_z=0$ when $t=0$.

If the electron starts from rest in such a crossed electric and magnetic fields, first of all it will be accelerated in the direction of electric fields and then, with increase in velocity, gets deflected in a direction perpendicular to the plane containing the electric and magnetic field vectors. Let at any subsequent time v_x, v_y and v_z be the components of velocity in X, Y and Z directions respectively. Considering the electric field along Y axis (potential increasing with +ve Y) and magnetic field along Z axis, the equations of motion are

$$m\frac{dv_y}{dt}=Q_e E-BQ_e v_x \qquad \ldots(1)$$

$$m\frac{dv_x}{dt}=BQ_e v_y \qquad \ldots(2)$$

$$m\frac{dv_z}{dt}=0, \qquad \ldots(3)$$

because :

(a) due to the existence of B in Z direction and the velocity of electron in X direction, a force of magnitude $-BQ_e v_x$ will act in Y direction,

(b) due to the existence of B in Z direction and the velocity of electron v_y in the Y direction, a force of magnitude $BQ_e v_y$ will act in X-direction, and

(c) since the Z component of velocity v_z of the electron and magnetic field B are parallel to each other, there will be no force in Z direction.

From equation (1), we have

$$\frac{dv_y}{dt}=\frac{Q_e E}{m}-\frac{BQ_e}{m}v_x=a-\omega v_x, \qquad \ldots(4)$$

where $$a=\frac{Q_e E}{m} \text{ and } \omega=\frac{BQ_e}{m},$$

and equation (2) becomes

$$\frac{dv_x}{dt}=\frac{BQ_e v_y}{m}=\omega v_y. \qquad \ldots(5)$$

Differentiating equation (4) and substituting the value of $\frac{dv_x}{dt}$ from equation (5), we get

$$\frac{d^2v_y}{dt^2}=-\omega\ \frac{dv_x}{dt}=-\omega^2 v_y$$

or
$$\frac{d^2v_y}{dt^2}+\omega^2 v_y=0. \qquad \ldots(6)$$

The solution of equation (6) is

$$v_y=A_1 \cos \omega t+A_2 \sin \omega t, \qquad \ldots(7)$$

where A_1 and A_2 are constants and can be determined by applying the boundary conditions.

When $t=0$, $v_y=0$, then

$$0=A_1$$

$$\therefore \quad v_y=A_2 \sin \omega t \qquad \ldots(8)$$

Differentiating equation (8), we get

$$\frac{dv_y}{dt}=A_2\omega \cos \omega t. \qquad \ldots(9)$$

From equation (4), when $t=0$, $v_x=0$ *i.e.* $\frac{dv_y}{dt}=a$. Substituting in equation (9), we get

$$\therefore \quad a=A_2\omega \quad \text{or} \quad A_2=\frac{a}{\omega}$$

$$v_y=\frac{a}{\omega} \sin \omega t. \qquad \ldots(10)$$

Integrating equation (10), we have

$$\int \frac{dy}{dt}=\frac{a}{\omega}\int \sin \omega t\ dt$$

$$y=-\frac{a}{\omega^2} \cos \omega t+D$$

When $t=0$. $y=0$, hence $D=\frac{a}{\omega^2}$.

$$\therefore \quad y=\frac{a}{\omega^2}\ (1-\cos \omega t).. \qquad \ldots(11)$$

The corresponding value of x, as obtained by integrating equation (5), is

$$x=\frac{a}{\omega^2}\ (\omega t-\sin \omega t). \qquad \ldots(12)$$

The path of the electron is given by equations (11) and (12) and is shown in fig. 27.

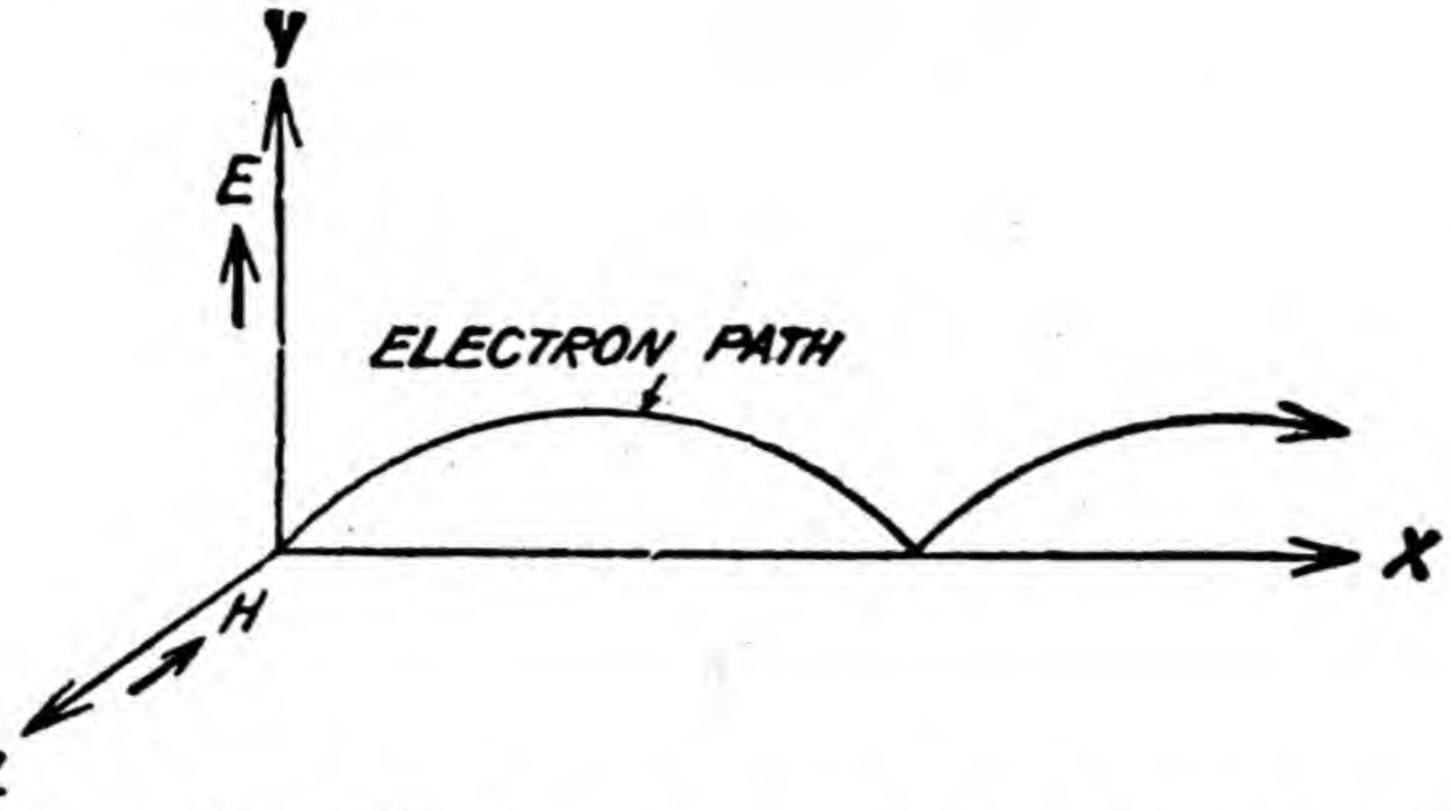
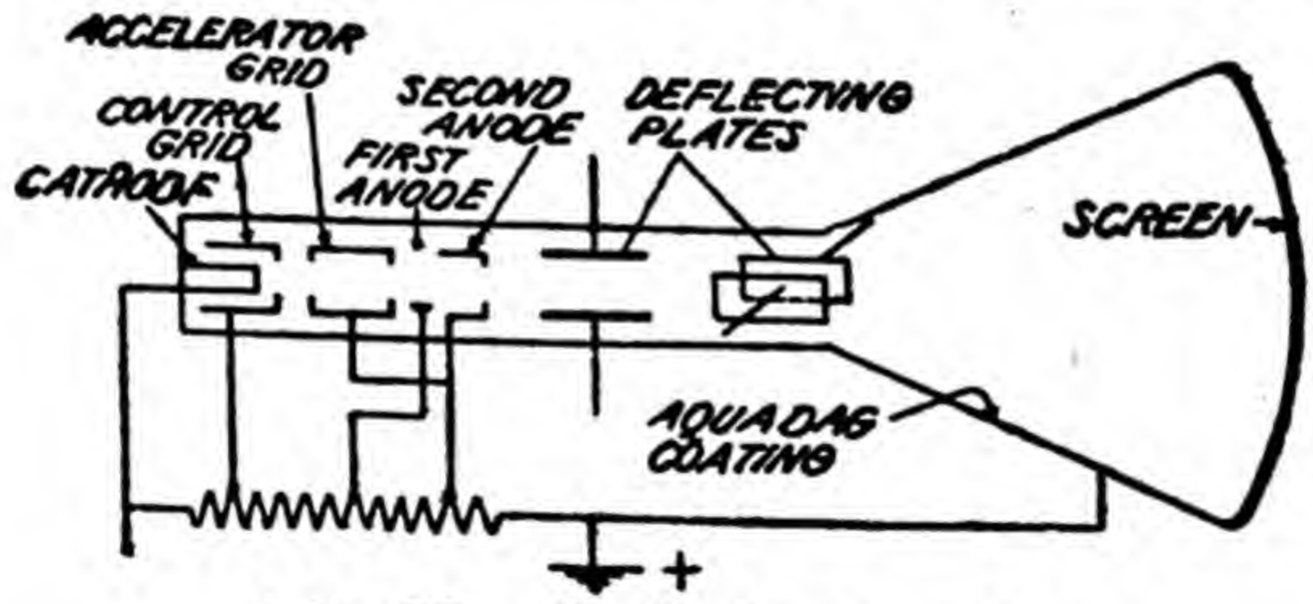


Fig. 27. Path of the electron in combined electric and magnetic fields at right angle to each other.

5·7-4. CATHODE-RAY TUBE :

This is a special type of tube among vacuum tubes which makes use of the geometrical form rather than the intensity of its electron stream and converts the energy of its electron stream into a visual indication.

In physical size the ordinary cathode ray tube ranges from 3 to 5 inches in diameter. The tubes operate-with a beam accelerating potential between 800 to 10,000 volts. The tube is evacuated to a pressure of about 10^{-5} mm. or 10^{-7} mm. of mercury. The electrode arrangement in a cathode ray tube is shown in fig. (28).



ig. 28. Cathode ray tube.

Cathode ray tube consists of three basic components :

1. *the electron gun*, which produces, accelerates and focusses the emitted electrons into a narrow beam,

2. *a deflecting system*, which deflects the electron beam either electrically or magnetically inaccordance with the phenomenon (voltage or current waveform) to be displayed,

3. *a fluorescent screen*, upon which the beam of electrons impinges to produce a spot of visible light.

The electron gun : The electron gun consists of a cathode, an accelerating electrode or grid, a focussing anode and a final accelerating anode. The electrodes are in the form of cylinders surrounding the cathode. The filament emits electrons and the control grid is used to concentrate the electrons emitted by filament because they have a tendency to spread out due to the force of

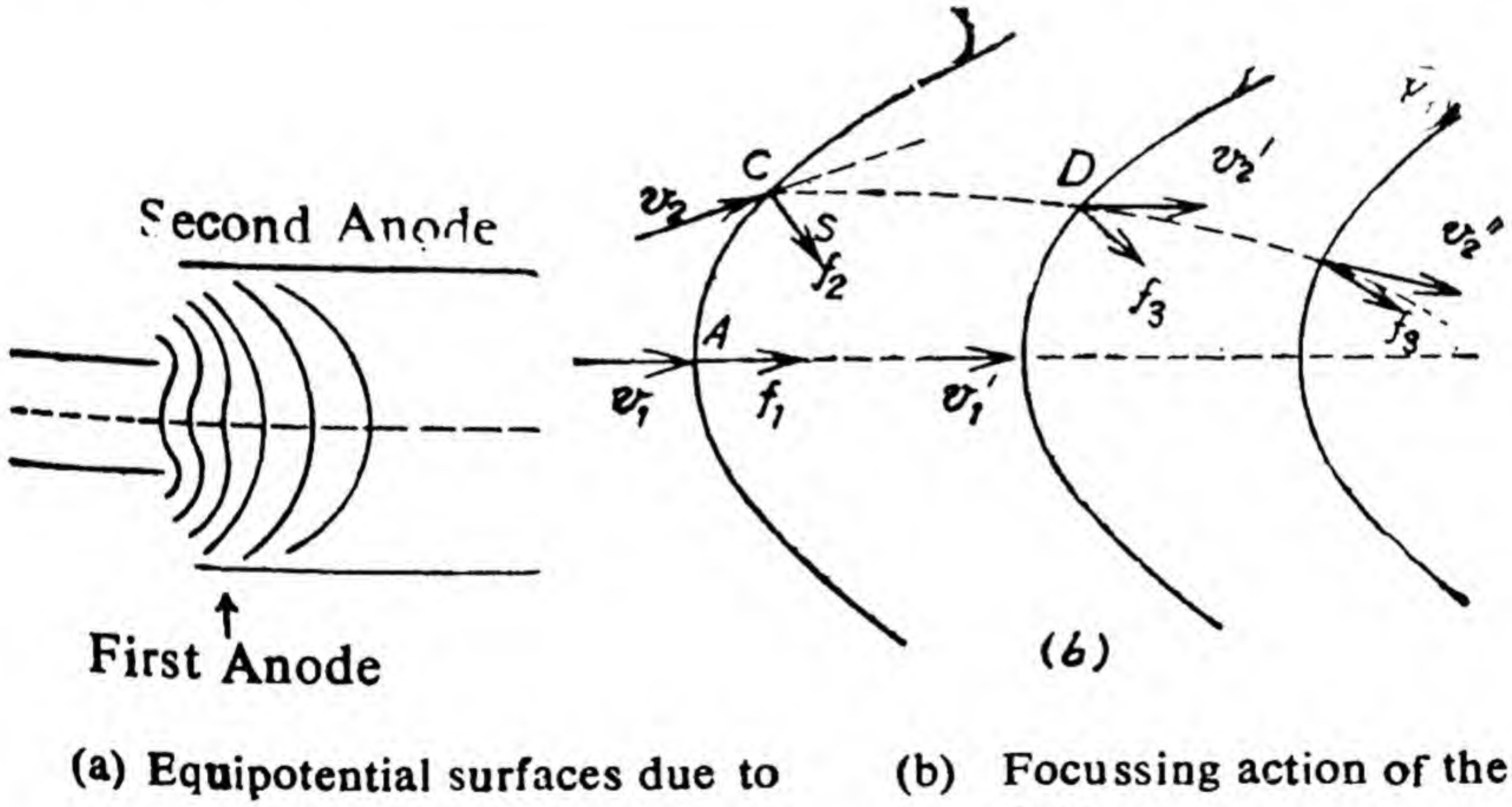
mutual repulsion betwen one another. It is charged negatively with respect to the cathode by the power supply. The applied potential can be varied as to control the rate of flow of electrons and thus, the intensity of beam is controlled. Electrons passing through it are repelled and consequently, are bunched together The first anode is made positive by the power supply which attracts the negatively charged electrons emitted by the cathode. The highly positive anode number two then focusses them into a thin dense beam, *i.e.* these positively charged electrodes serve not only to accelerate the electrons, but also to form them into a narrow them.

Focussing : The convergent beam from accelerating electrode has a tendency to spread because of the mutual repulsion between the electrons. Hence some focussing device is required to bring the beam to a sharp focus at the screen. Two methods of focussing most commonly used are :

(*a*) Electromagnetic focussing.

(*b*) Magnetostatic focussing.

(a) *Electrostatic focussing* : The anodes 1 and 2 (fig. 29*a*), provide an electron lens system for focussing the beam into a spot on the fluorescent screen. Anode 2 is kept at a higher positive



(a) Equipotential surfaces due to two cylindrical electrodes.

(b) Focussing action of the electric field.

Fig. 29.

potential than anode 1 ; moreover final anode 2 also has a central aperture so to permit only a well defined beam to pass through it. The electrostatic field between them is shown by its equipotential surfaces.

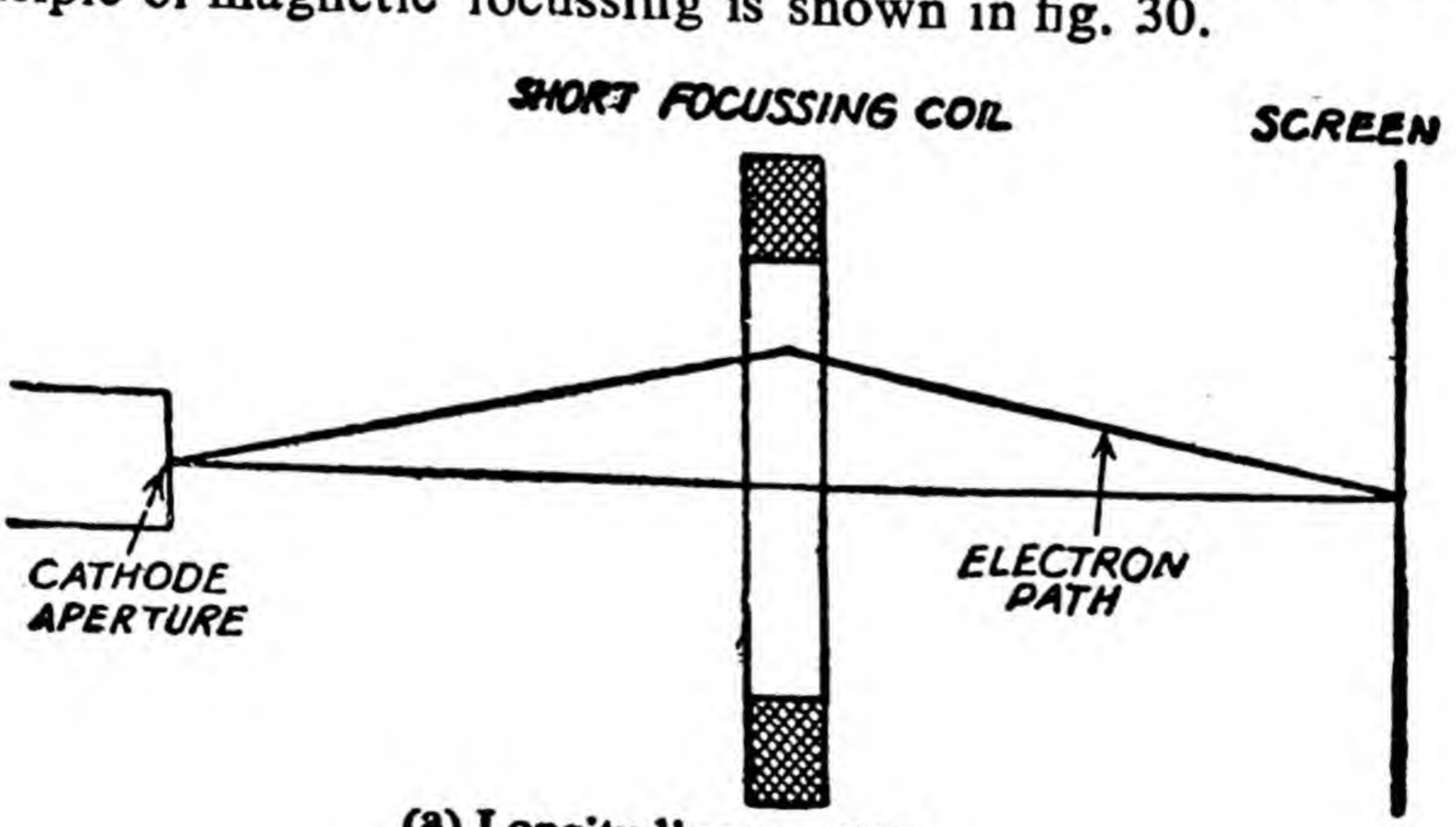
Let the electron strike the equipotential surface normally at a point *A* with velocity v_1 (Fig. 29 *b*). We know that a force is exerted by the electric field on the electron which is given by $-Q_e$ grad V.

Grad V is acting normally to the surface at A hence the direction of the electron motion remains along the X-axis but it is accelerated.

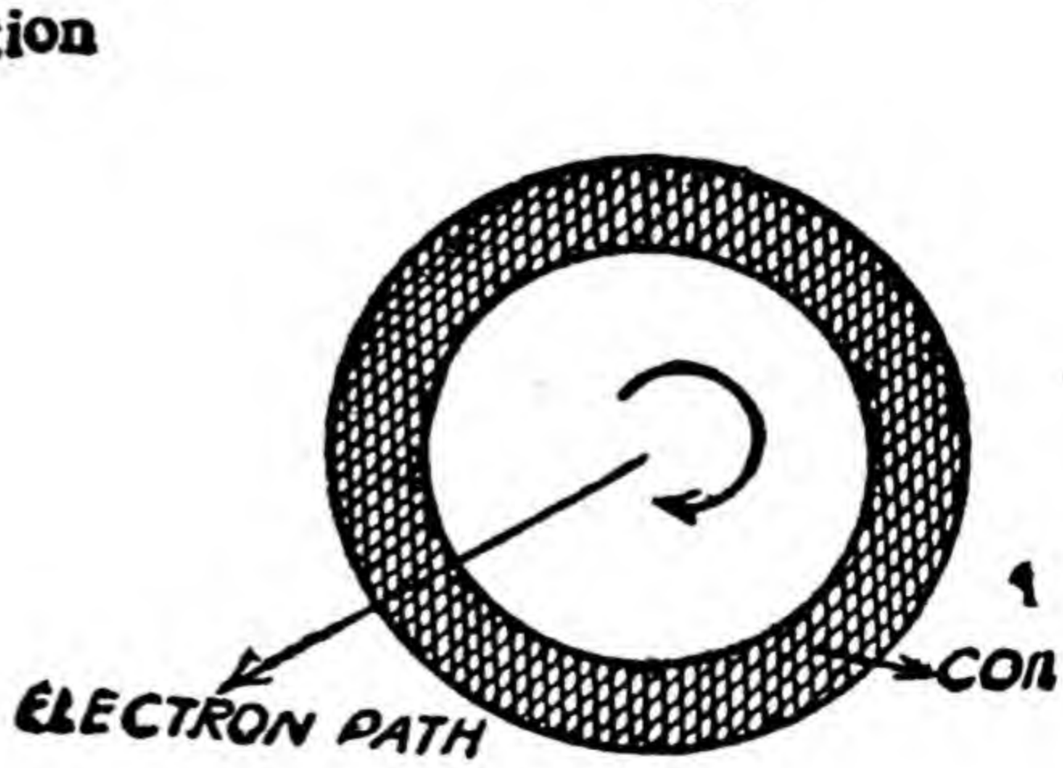
Now let the electron strike the equipotential surface at a point C with velocity v_2. A force $f_2 = -Q_e$ grad V_1 acts on the electron in a direction CS normal to the equipotential surface V_1. Due to this force, the direction of motion of the electron as well as velocity are changed. The velocity increases and the direction will be along CD. Thus the path is some-where in between the direction of the initial velocity and direction of the force. In this way the beam is bent towards the axis of the cathode ray tube as shown in (fig 29*b*).

The focussing action of the electron lens depends upon (i) ratio of the diameters of the cylindrical electrodes, and (ii) ratio of the electrode potentials with respect to cathode.

(*b*) *Magnetostatic focussing.* The electron beam may also be focussed by a magnetic field achieved either by permanent magnet or by passing current through a coil of wire, the axis of which coincides with the beam axis. This method is known as magnetostatic focussing. On account of the advantages of the magnetostatic focussing over electrostatic focussing, it is used in television, where a sharp focus over the whole screen is highly desirable. The principle of magnetic focussing is shown in fig. 30.



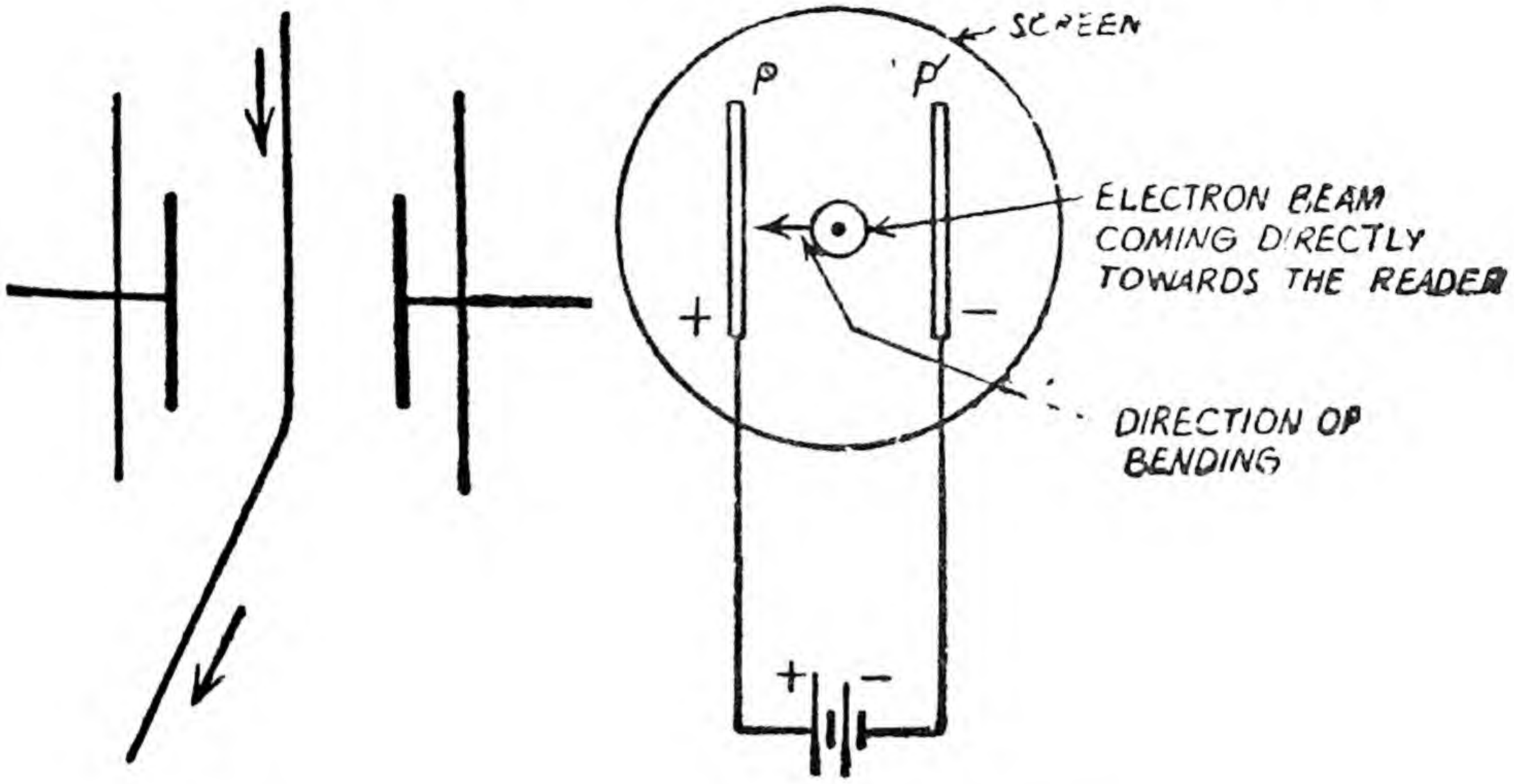
(a) Longitudinal section



(b) End view

Fig. 30. Focussing action of a short focussing coil.

Deflection system : – Next to the focussing system, there are two pairs of deflection plates. One pair of parallel plates is kept vertical, while the other pair of plates is kept horizontal. The pair



(a) Top view (b) End view

Fig. 31. Horizontal deflecting plates.

producing the horizontal deflection of the spot is called the horizontal or *X*-plates while, the other pair causing a vertical displacement of the spot is called the vertical or *Y*-plates. The deflecting voltages are applied between the two plates of the pair concerned. Figure 31 shows a pair of metal plates that makes possible control of the electron stream in a horizontal direction. The plate *P* is made positive and *P'* negative, the electron passing through them will be deflected (curved) towards the positive plate *P*. The resulting light spot will, therefore, move to the side screen.

Fluorescent screen : To make the screen fluorescent, it is coated with calcium tungstate, zinc silicate or some other material. When electrons of high velocity strike it, the energy absorbed by the treated surface fluoresces or emits light so that a round spot glows on the screen, the colour depending on the nature of the screen material.

5·7-5. ELECTROSTATIC DEFLECTION SENSITIVITY :

It is defined as the amount of deflection (in mm or inch) of the electron spot produced when a voltage of one volt d.c. is applied between the corresponding deflection plates. This is different for *X* and *Y* plates.

Fig. 32 shows the configuration of electronic deflection systems.

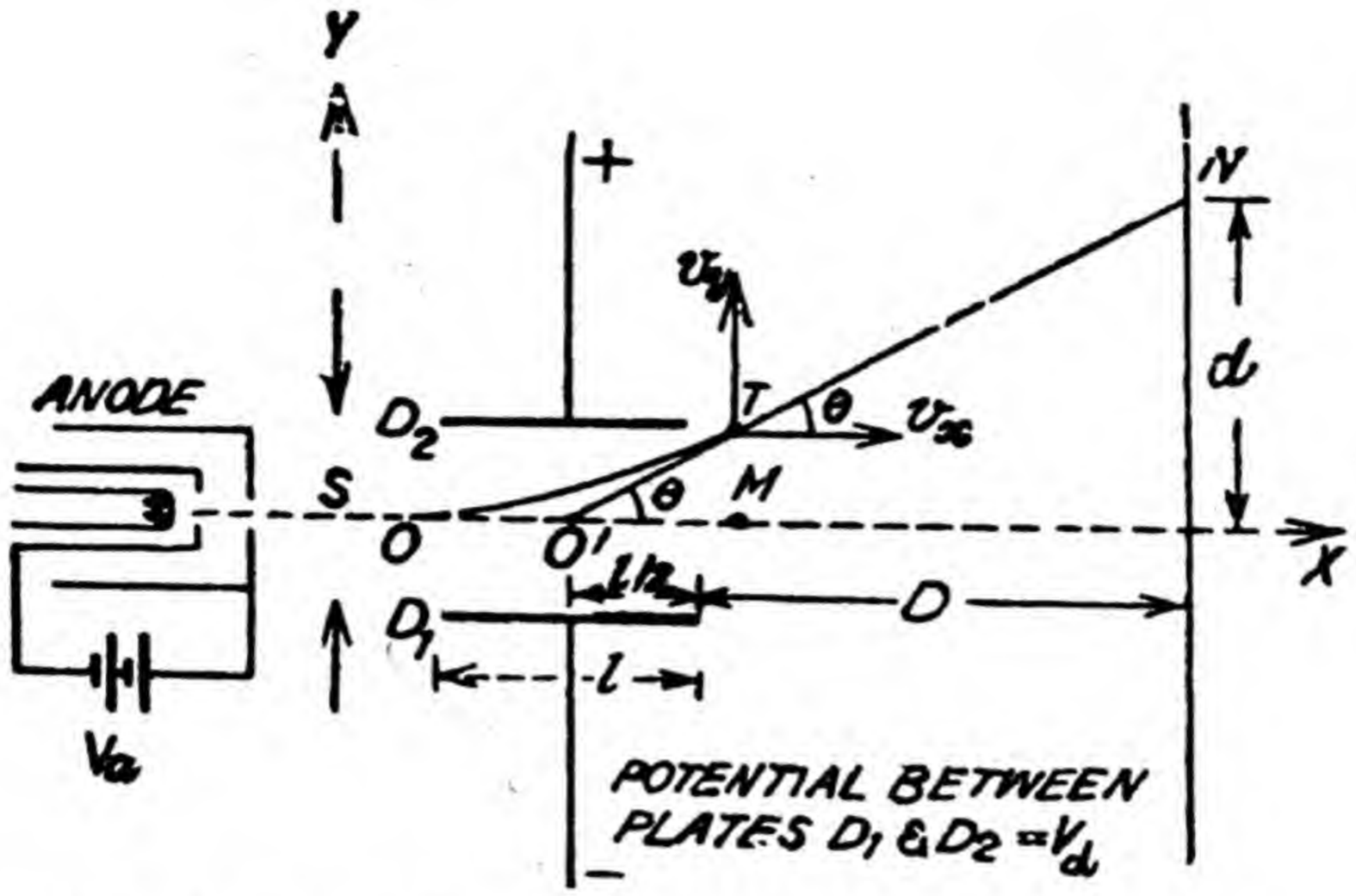


Fig. 32. Electrostatic deflection in cathode ray tube.

Let the length of two similar plates $=l$,
spacing between two plates $=S$,
distance of the plates from the screen $=D$,
voltage applied between the plates $=V_d$.
velocity of an electron when entering the field of deflection plates $=v$.

Now
$$v=\sqrt{\left(\frac{-2Q_eV_a}{m}\right)} \qquad \ldots(1)$$

where $Q_e=-1{\cdot}6\times10^{-19}$ is the charge on an electron in coulombs, m the mass of the electron and V_a accelerating potential applied to second anode. Here it is assumed that the emission velocity of the electron from cathode is zero. Substituting for m and Q_e, we get

$$v=5{\cdot}94\times10^5\sqrt{(V_a)}\ \text{m./sec.}$$

The force of attraction F, experienced by the electron in the electrostatic field towards the positive plate is given by

Force = mass × acceleration $=Q_eE$

$$\therefore \quad m\frac{d^2y}{dt^2}=-Q_e\frac{V_d}{S} \qquad \text{where } E=\frac{V_d}{S}$$

or
$$\frac{d^2y}{dt^2}=-Q_e\frac{V_d}{mS}, \qquad \ldots(2)$$

where m is mass of the electron $=9{\cdot}11\times10^{-31}$ kg. and $Q_e=-1{\cdot}6\times10^{-19}$ coulomb. With the help of equation (2) the velocity and displacement of the electron can be calculated provided the initial boundary conditions are known. It is assumed that the electron reaches at O whose coordinates are $x=0$ and $y=0$ at $t=0$.

At the point O, $v_y=0$ and $v_x=\sqrt{\left(\frac{-2O_eV_a}{m}\right)}$

Integrating equation (2), we get

$$\frac{dy}{dt}=v_y=\frac{-Q_eV_d}{m.S}t+C,$$

where C is the constant of integration. When $t=0$, $v_y=0$ hence $C=0$. Therefore

$$\therefore \quad \frac{dy}{dt}=\frac{-Q_eV_d}{m.S}.t \qquad \ldots(3)$$

Integrating again equation (3), we have

$$y=\frac{-Q_eV_d}{2m.S}.t^2+D$$

But $D=0$ as $y=0$ when $t=0$ so that

$$y=\frac{-Q_eV_d}{2m.S}.t^2. \qquad \ldots(4)$$

Equation (4) gives the displacement of the electron along Y axis at any instant.

The distance travelled by the electron in time t along X axis is

$$x=\text{velocity}\times\text{time}=\left(\frac{-2Q_eV_a}{m}\right)^{1/2}.t. \qquad \ldots(5)$$

Eliminating t from equations (4) and (5), we have

$$y=\frac{-Q_eV_d}{2m\,S}\cdot\frac{x.^2m}{2\,(-Q_e).V_a}=\frac{V_d.x^2}{4.S.V_a} \qquad \ldots(6)$$

This is the equation of parabola and hence the electron moves in a parabolic path between the plates. After the point T, it follows a straight path and strikes the screen at a point N, the tangent NT makes an angle θ with the X-axis. (fig. 32).

From figure, $$\tan\theta=\frac{d}{(D+l/2)} \qquad \ldots(7)$$

We know that $$\tan\theta=\left.\frac{dy}{dx}\right|_{x=l}$$

From equation (6) $$\tan\theta=\left.\frac{dy}{dx}\right|_{x=l}=\left.\frac{V_d.x}{2.S.V_a}\right|_{x=l}$$

or $$\tan\theta=\frac{V_d\,l}{2.S.V_a} \qquad \ldots(8)$$

From equations (7) and (8), we have

$$\frac{d}{(D+l/2)}=\frac{V_dl}{2.S.V_a}.$$

$$\therefore \text{ Deflection sensitivity}=\frac{d}{V_d}=\frac{l}{2.S.V_a}\,(D+l/2). \qquad \ldots(9)$$

Thus deflection sensitivity is

(*i*) directly proportional to the length of the deflection plates l,

(*ii*) inversely proportional to the spacing S between the plates,

(*iii*) directly proportional to $(\frac{1}{2}l+D)$,

(*iv*) inversely proportional to the anode voltage of the final stage

5·7-6. MAGNETIC DEFELECTION SENSITIVITY :

The magnetic field deflects the beam perpendicular to its direction of motion. The electron follows a circular path in the field (fig. 33).

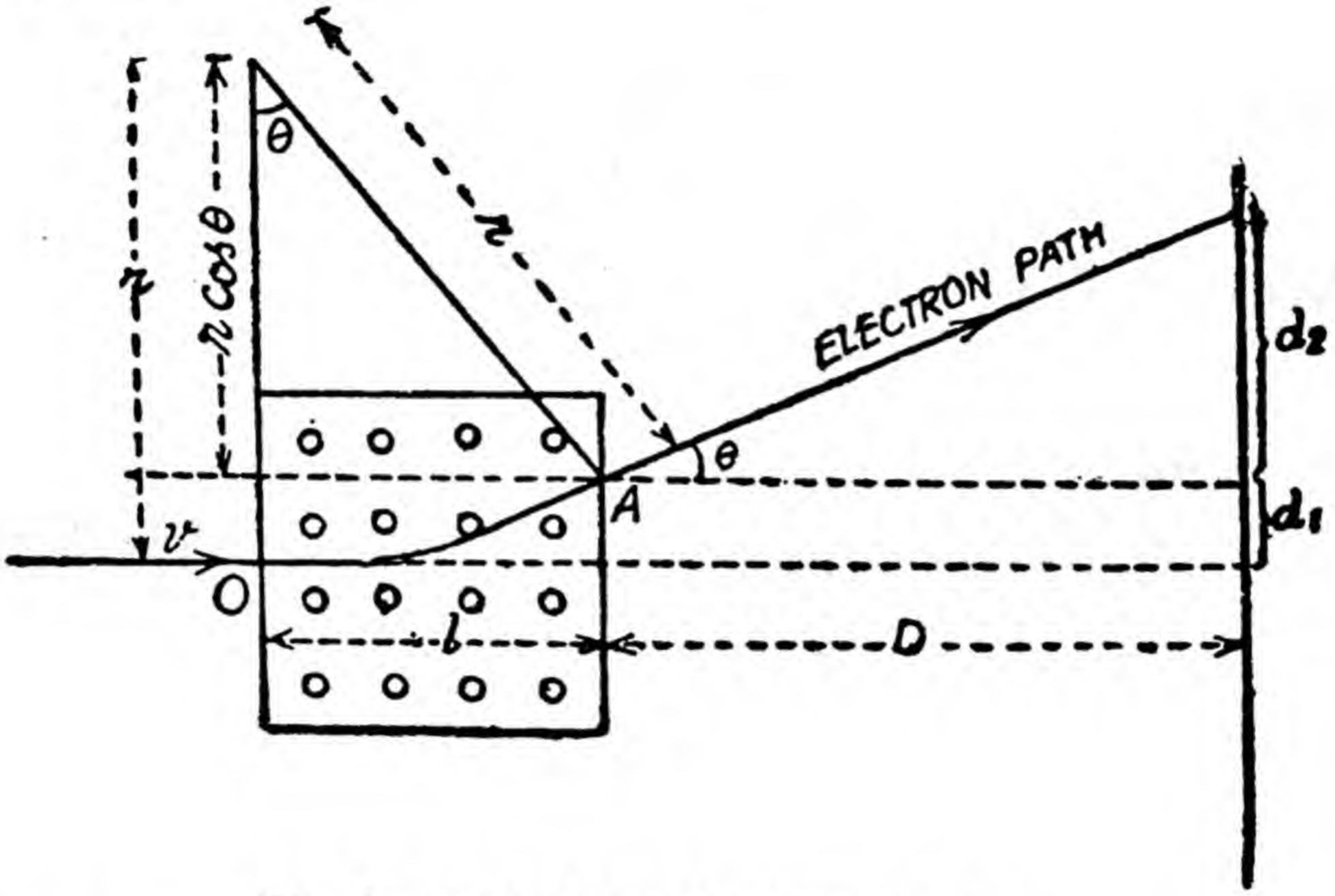


Fig. 33. Magnetic deflection system.

Let the radius of the circle followed by the electron in the field = r metres.

velocity of the electron = v metres/sec.

mass of the electron = m k.gm.,

the centrifugal force on electron = $\frac{mv^2}{r}$ ...(1)

The force F exerted on the electron by the magnetic field is given by

$$F = -B\,Q_e.v \sin\theta, \quad \text{.. (2)}$$

where B = magnetic flux density in webers/sq. metre.

θ = angle which the electron path makes with the direction of magnetic flux.

Q_e = charge on the electron.

If $\theta = 90°$, *i.e.* $\sin\theta = 1$, then

$$F = -B\,Q_e\,v. \quad \text{...(3)}$$

Equating the centrifugal force to that exerted by magnetic field

$$-B.Q_e.v = \frac{mv^2}{r}.$$

$$r = -\frac{mv}{BQ_e} \quad \text{.. (4)}$$

Now the electron follows a straight line path as shown in figure 33.

Total displacement $d=d_1+d_2$. ...(5)

The angular change in the direction of the electron after passing through the magnetic field is given by θ ; then

$$\sin\theta=\frac{l}{r},$$

$$\begin{aligned} d_1 &= r-r\cos\theta \\ &= r-r\sqrt{(1-\sin^2\theta)} \\ &= r-r\sqrt{\left(1-\frac{l^2}{r^2}\right)}. \end{aligned} \quad \text{..(6)}$$

Also $$\tan\theta=\frac{d_2}{D},$$

$$d_2=D\tan\theta=D\,\frac{\sin\theta}{\cos\theta}$$

$$=\frac{D.\frac{l}{r}}{\sqrt{(1-\sin^2\theta)}}$$

$$=\frac{Dl}{r\sqrt{\left(1-\frac{l^2}{r^2}\right)}}$$

$$=\frac{Dl}{\sqrt{(r^2-l^2)}}.$$

Now $$d=r-r\sqrt{\left(1-\frac{l^2}{r^2}\right)}+\frac{Dl}{\sqrt{(r^2-l^2)}}.$$

If $l << r$, then l and $\frac{l^2}{r^2}$ can be neglected.

$$\therefore \quad d=\frac{Dl}{r}, \qquad \text{...(7)}$$

Now $$\frac{Dl}{r}=-\frac{DlB.Q_e}{mv}=\frac{-DlBQ_e}{m\sqrt{\left(-\frac{2V_a}{m}.Q_e\right)}}$$

$$\therefore \quad d=\frac{-DlB}{\sqrt{\left(-2V_a\,\frac{m}{Q_e}\right)}}=-\frac{DlB\sqrt{\left(\frac{-Q_e}{r}\right)}}{\sqrt{(2V_a)}}, \qquad \text{...(8)}$$

The deflection per unit field intensity

$$\frac{d}{B}=\frac{Dl}{\sqrt{(2V_a)}}\sqrt{(-Q_e/m)}\ \text{meter/weber}$$

is called the magnetic deflection sensitivity

Thus deflection sensitivity is

(*i*) directly proportional to flux density B, length l and distance D between screen and field,

(*ii*) inversely proportional to the square root of the final anode voltage V_a and products of electron mass and its charge.

The important uses of cathode-ray tube are in radar sets and ionospheric height measurements. In laboratory, these are also used for the comparison of frequencies and plot of *B-H* curves of iron.

5·7-7. CATHODE RAY TUBE CONNECTIONS :

Fig. 34. shows the connections for a cathode ray tube.

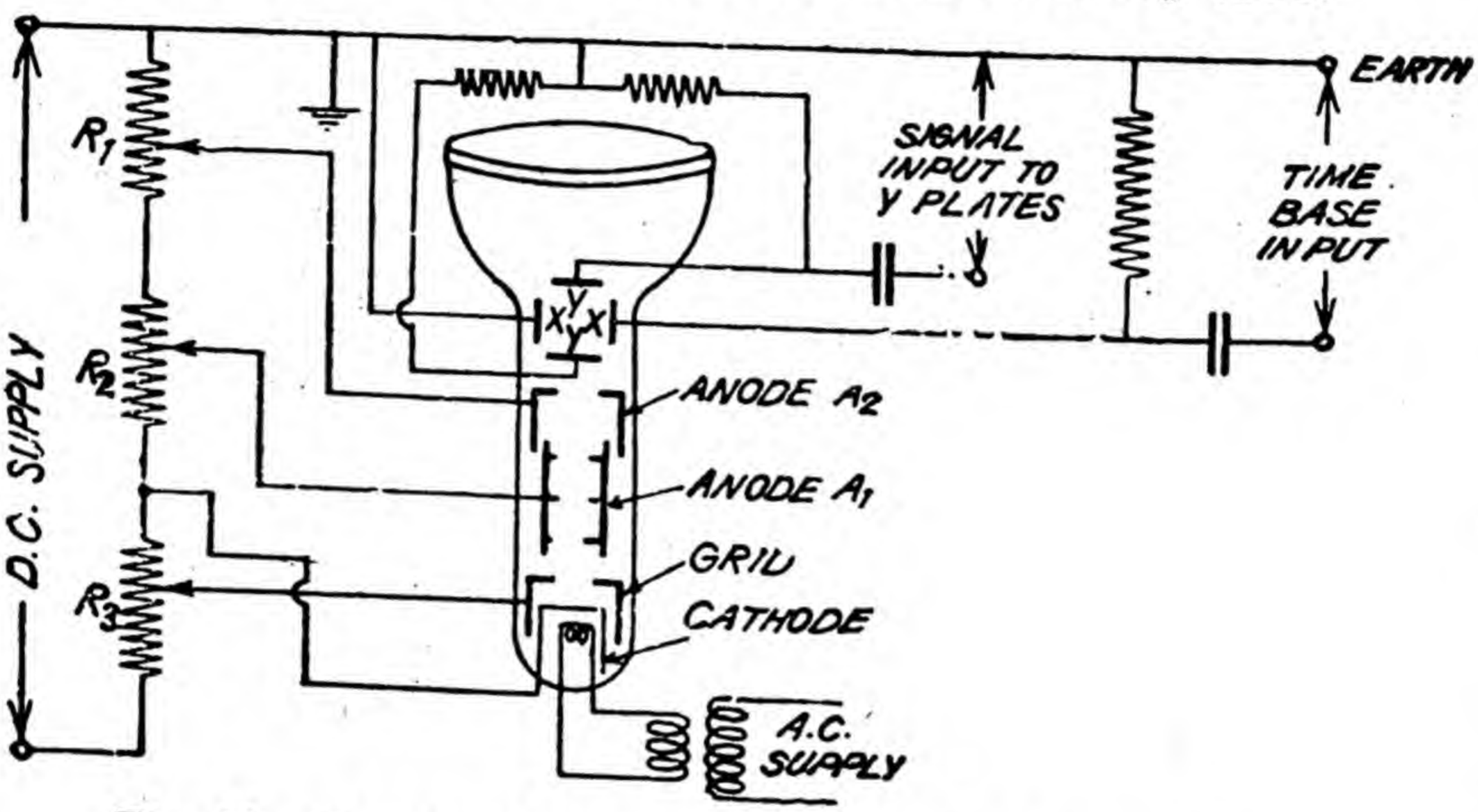


Fig. 34. Showing the connections of cathode ray tube.

The high tension supply for accelerating anode A_1, focussing electrode A_2 and grid is taken from vacuum tube rectifier. With the help of the resistances R_1, R_2 and R_3 used as voltage divider, proper voltages are applied. The positive potential applied to the first anode can be controlled by the potentiometer R_2, so that the ratio of second to first anode potential is varied This varies the distribution of lines of force in the electrostatic lens between the anodes, and the focussing of the beam takes place. By varying the voltage applied to A_2, the size of the spot and intensity may be varied.

In order to make measurements with C.R.O one of the pairs of the plates is connected to a circuit known as time base. It provides potential of variable frequency to be applied to say (X, X) plates ; the spot on screen moves according to the frequency and voltage of time base. The time base e.m.f. increases at a uniform rate to a positive potential and then decreases suddenly to its initial value during one cycle as shown in fig. 35.

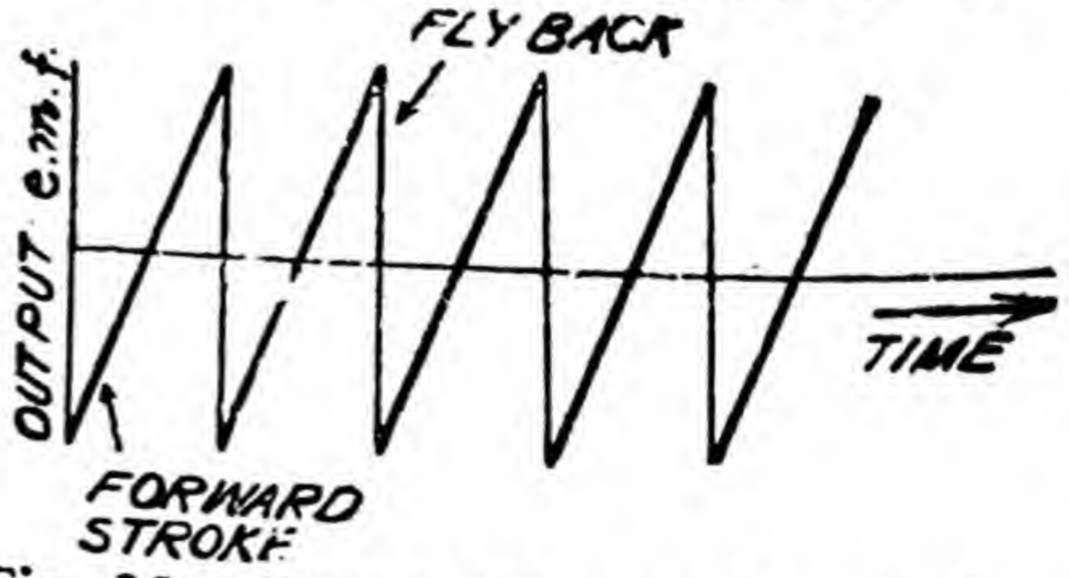


Fig. 35. Saw-tooth waveform output from a time base generator.

In successive cycles, this linear increase and rapid decrease are repeated. Frequency used in time base in a linear time base in which the spot moves in X direction from left hand side to right hand side with a definite but variable speed and flies back rapidly after reaching the extremity of screen.

Time base : There are number of time base circuits :

(1) Neon time base circuit.
(2) Thyratron time base circuit.
(3) Hard valve time base circuit.

(1) *Neon time base circuit* :

The simplest time base circuit, shown in fig. 36, is made by placing a condenser C across the terminals of a neon lamp N, charging the condenser through a resistance R from D.C. mains and discharging it through the neon lamp.

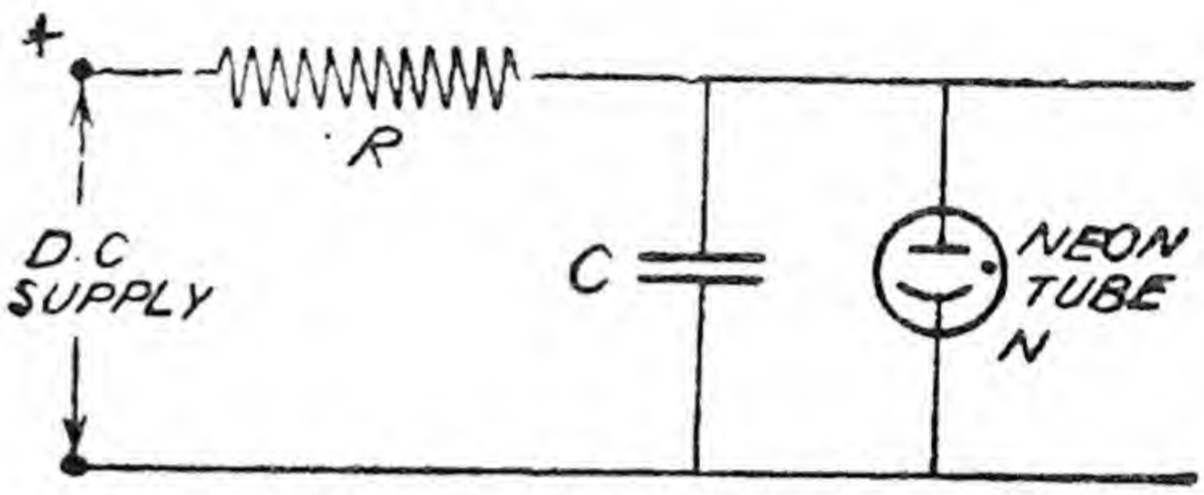


Fig. 36. Neon time base circuit.

When high tension supply is switched on, the condenser charges up through the resistance until the voltage across it reaches the striking value (150 volts) of the neon lamp. Now neon lamp glows which short circuits the condenser and hence it discharges through the lamp. The process is repeated again and again. The neon lamp flashing at regular intervals depends upon the product of the values of resistance, and capacity, and when both are large, the flashing rate will be low. The voltage across the condenser is applied to the X plates. When the condenser charges up, the spot on the screen moves slowly in the horizontal direction until the condenser gets charged and then due to discharge through the neon lamp, the spot is brought back instantaneously to the striking point. This sweep of the spot is repeated at a flashing rate of the neon

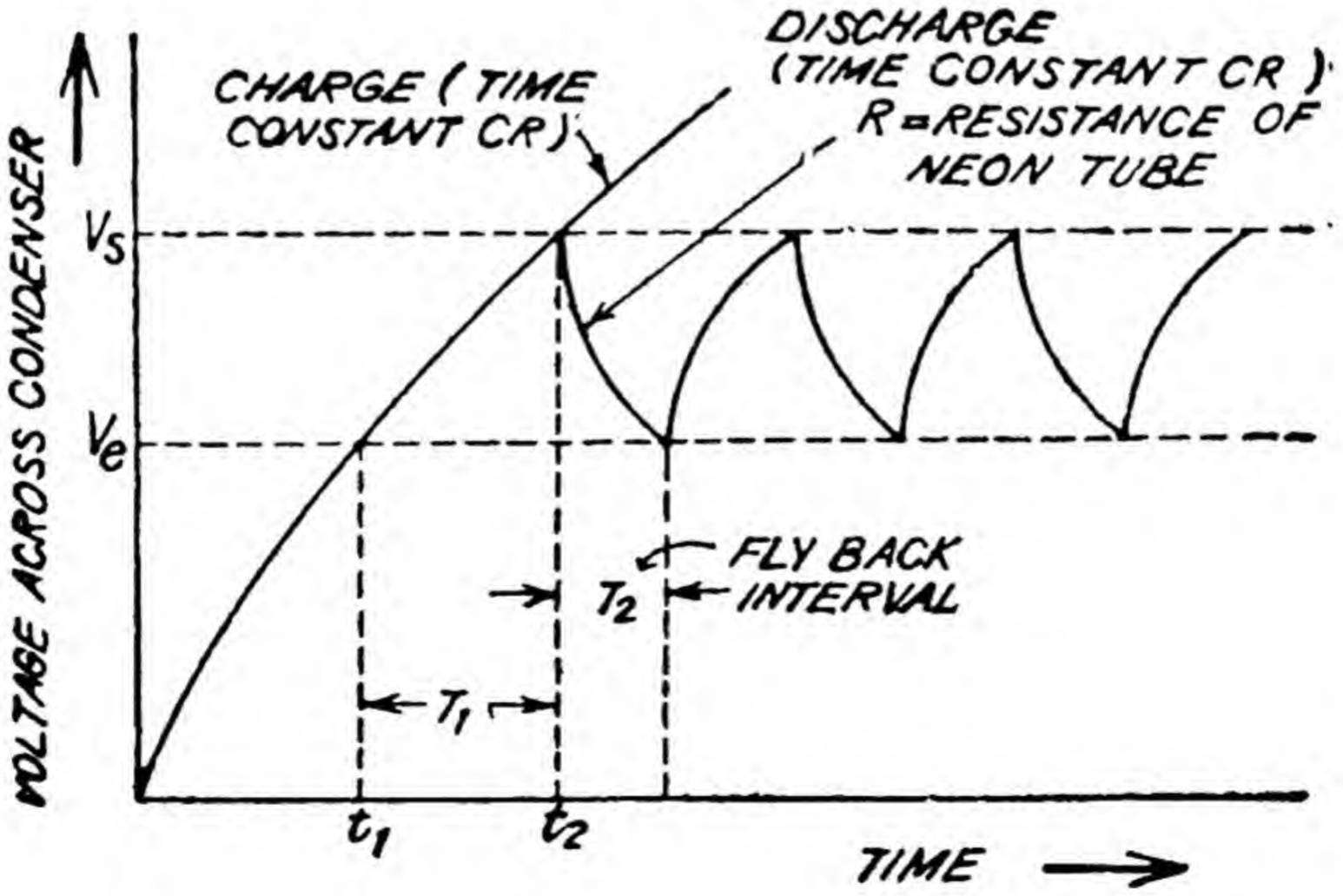


Fig. 37. Generation of saw tooth voltage using neon tube.

lamp. The frequency of the time base can be calculated as follows : Figure 37, shows the charging and discharging of a condenser in time base circuit. The periodic time T is sum of the time of charge T_1 and time of discharge or flyback period T_2. As T_2 is very small in comparison to T_1 it can be neglected in calculation of periodic time.

There is a definite striking voltage for neon to conduct. If the voltage is reduced, continuously, then at some stage conduction stops as ionisation ceases. This voltage is referred to as extinguishing voltage. These voltages are given as

$$V_s = E_{d.c.}(1 - e^{-t_2/CR}) \qquad \ldots(1)$$

$$V_e = E_{d.c.}(1 - e^{-t_1/CR}) \qquad \ldots(2)$$

where $E_{d.c.}$ is the voltage applied across condenser.

From equation (1)

$$e^{-t_2/CR} = \frac{E_{d.c.} - V_s}{E_{d.c.}}$$

From equation (2)

$$e^{-t_1/CR} = \frac{E_{d.c.} - V_e}{E_{d.c.}}$$

$$\therefore \quad e^{(t_2 - t_1)/CR} = \frac{E_{d.c.} - V_e}{E_{d.c.} - V_s}$$

or $$(t_2 - t_1)/CR = \log_e \frac{E_{d.c.} - V_e}{E_{d.c.} - V_s}$$

$$T_1/CR = \log_e \frac{E_{d.c.} - V_e}{E_{d.c.} - V_s}$$

Since $T_2 << T_1$, we can have $T \approx T_1$

$$T = CR \log_e \frac{E_{d.c.} - V_e}{E_{d.c.} - V_s}$$

and $$f = \frac{1}{CR \log_e \frac{E_{d.c.} - V_e}{E_{d.c.} - V_s}}$$

It is seen that sweep voltage curve is not linear, so that time base circuit can not be used in commercial C. R. O.

(2) *Thyratron controlled time base* :

A thyratron controlled time base circuit is shown in figure 38.

The current in a pentode is almost independent of plate potential above a particular voltage known as knee voltage. It is controlled by screen grid potential, which can be used as a fine frequency control. The capacitor C will charge up linearly *via* the pentode valve V_1 which is operated on the part of its anode characteristic above the knee voltage. As the pentode charges the condenser at a constant rate, the potential of the condenser and hence that of the plate of thyratron increases also at a constant rate until the critical

potential is reached when the thyratron allows a discharge and then itself goes out of action since the discharging of the condenser has caused the plate potential to drop practically to zero. The charging

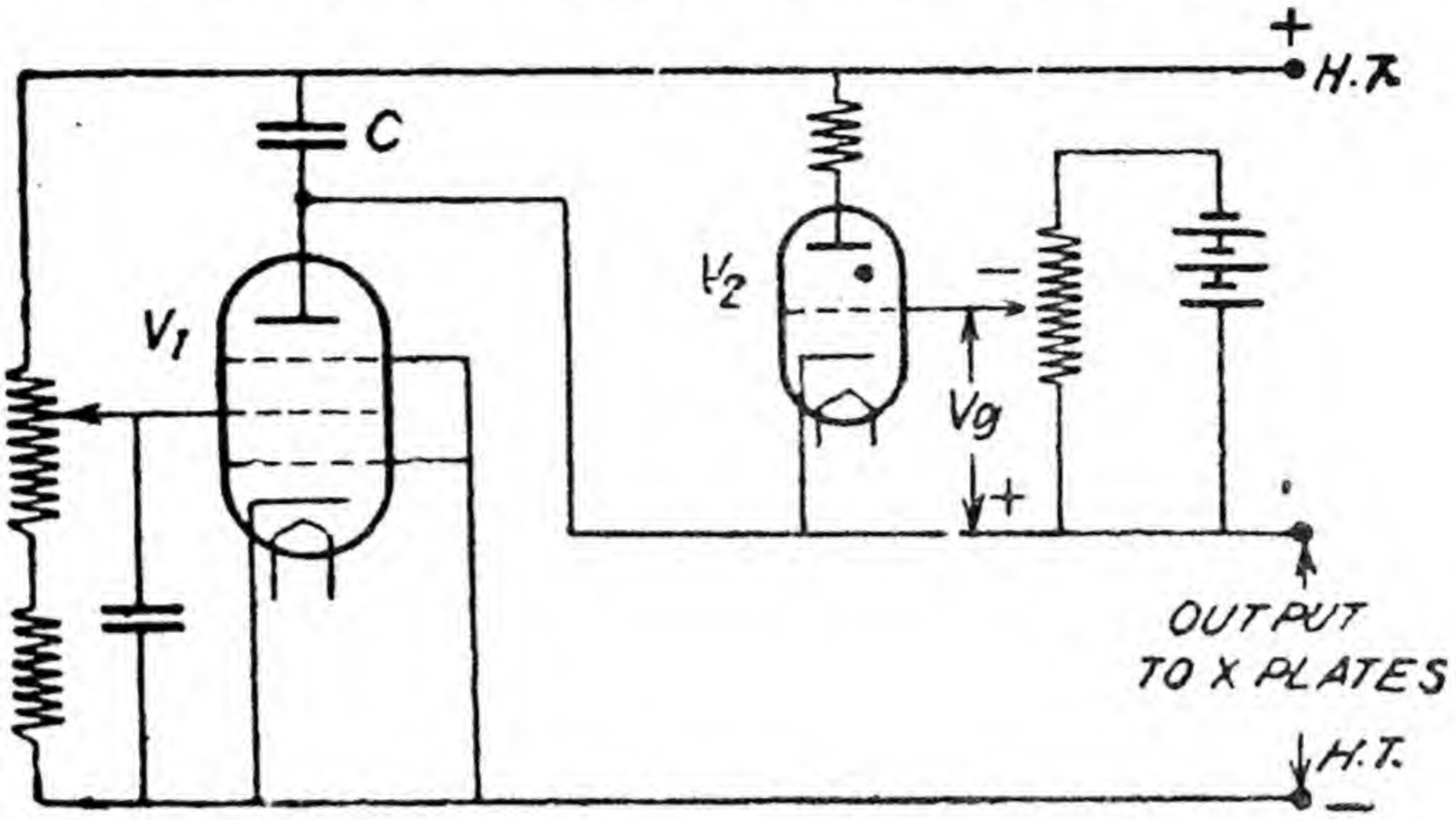


Fig. 38. A thyratron controlled time base.

action then begins again. In this way a saw-tooth waveform p.d. is set up across of which the period is decided by the capacitance C and the screen potential of the pentode valve. The striking potential of thyratron also depends on the potential of thyratron grid. A safety resistance is usually put in the plate circuit of the thyratron in the order to prevent it from passing too much current, which would damage the cathode by excessive ionic bombardment.

(3) *Puckle's hard valve (vacuum tube) time base :*

Fig. 39 shows a simple time base circuit designed by Puckle. This circuit utilises the constant current property of the pentode valve and gives highly linear sweep voltage.

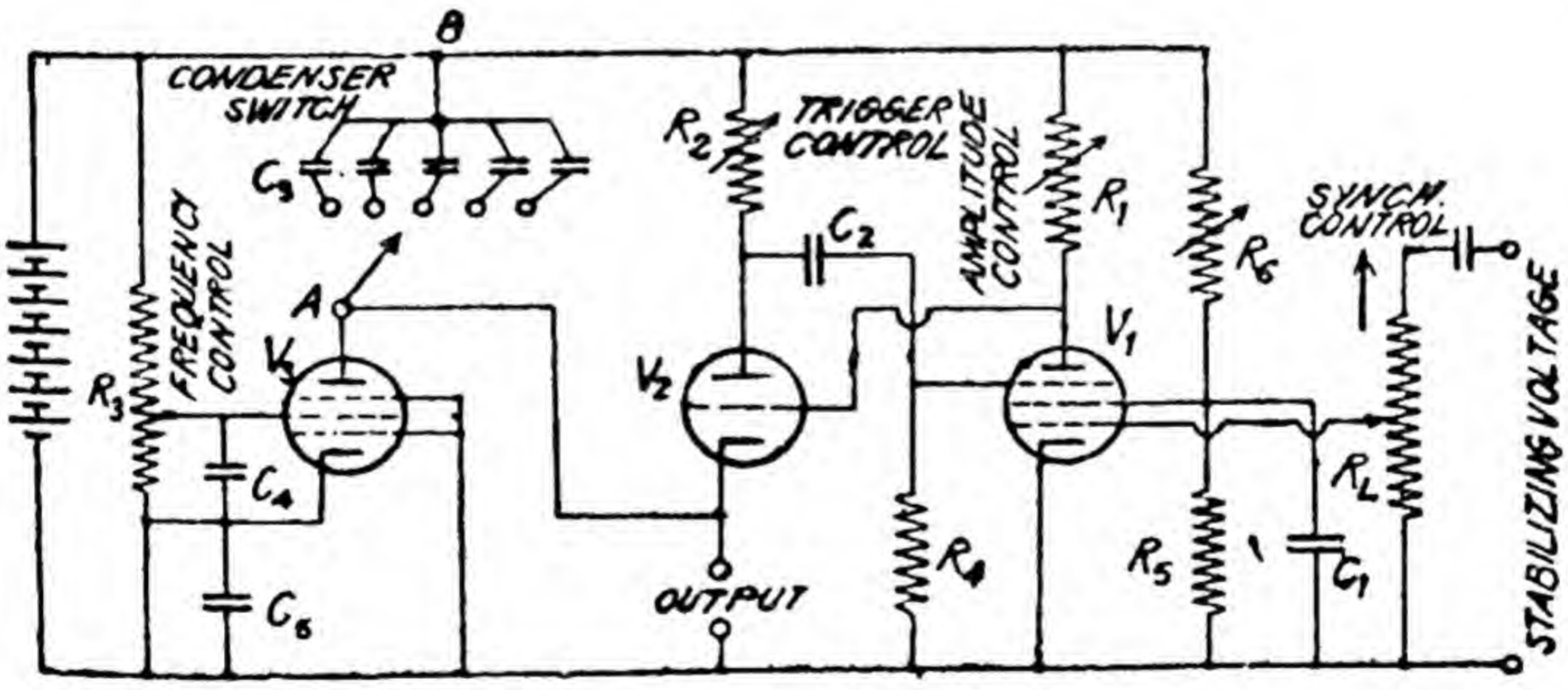


Fig. 39. Simplified circuit of Puckle's hard valve time base.

The condensor C_3 charges through the valve V_3 acting as a resistance when conduting. The voltage across C_3 increases linearly with time as the current through pentode V_3 remains constant. Consider that pentode valve V_1 is also conducting and it passes anode current through the resistance R_1. It is also assumed that at the begining the voltage drop across R_1 is much larger than the voltage drop across C_3, then valve V_2 is cut off as the point A is directly connected to the cathode of valve V_2.

When the condenser C_3 charges, the potential of the point A falls with respect to the point B and as the plate of valve V_3 is directly connected to the cathode of valve V_2, this fall in potential is conveyed to the cathode of V_2. Due to this fall in cathode potential of V_2 a current flows in it This current produces a potential drop across R_2 which lowers the potentials of plate of valve V_2 and suppressor grid of valve V_1. This results in the fall of plate current of V_1 and the voltage drop across R_1 is reduced. Now the grid of valve V_2 rises in potential. This increases the plate current of valve V_2 and the cycle is repeated.

Thus the grid of V_2 continuously rises in potential while the cathode falls. This cumulative action is very rapid and results in a heavy current through valve V_2 discharging the condenser C_3 through V_2. This discharge continues until the initial condition is restored. The charge and discharge curves of condenser C_3, so obtained, give a highly linear time base voltage.

The different controls in time base circuit are :

(1) *Frequency control or velocity control* : This controls the grid voltage of valve V_3 and thus determines the plate current. In this way this determines the charging rate of condenser C_3 and hence the frequency of time base circut.

(2) *Trigger control* : The condenser C_3 discharges through R_2 and variation of R_2 permits the adjustment of flyback. This is known as trigger control.

(3) *Amplitude control* : Variation of resistance R_1 determines the voltage drop across R_1 and hence the voltage upto which the condenser C_3 may charge before flyback starts. Thus this determines the amplitudes of time base voltage.

(4) *Condenser switch* : A number of condensers are connected at a point B and any one of them may be connected to a point A. Each condenser corresponds to a definite range of frequency of sweep voltage.

(5) *Synchronization control* : The sweep voltage frequency may be synchronized with any other alternating voltage by applying this voltage to the control grid of valve V_1. Synchronization may be done (*i*) internally with the frequency of the power supply or (*ii*) externally with some other a.c. voltage.

USES OF CATHODE RAY OSCILLOSCOPE :

Cathode ray oscilloscope can be used in a number of ways. Few of them are given below :

(1) *In television* : In television receiver cathode ray oscilloscope is used for recreation of images.

(2) *In Radar* : In radar, cathode ray oscilloscope is used for giving the visual indication *i.e.* to detect and to give the position of the target such as aircraft.

(3) *Measurement of direct or alternating voltage* :

The deflection on a cathode ray oscilloscope screen is directly proportional to the volta ge applied to the deflecting plates. First of all the screen is directly calibrated in terms of the known voltages *i.e.* the deflection sensitivity is calculated. The direct voltage to be measured is applied between a pair of deflecting plates. The deflection produced multiplied by the deflection sensitivity gives the value of direct voltage.

The alternating voltage of a sinusoidal waveform to be measured is applied between a pair of deflection plates and the length of the straight line produced on the screen is measured. Knowing the deflection sensitivity, peak value of the applied voltage can be calculated.

(4) *Measurement of direct or alternating current* :

We know that cathode ray oscilloscope is a voltage indicating device and hence current to be measured is first passed through a known resistance and then the voltage developed is passed to vertical deflecting plates. The horizontal plates are short circuited. The voltage is noted from the deflection on the screen. Current may then be calculated.

(5) *Study of waveforms* : In order to study the waveform of alternating voltage, sinusoidal or otherwise, it is applied to the *Y* deflection plates and a time base voltage of suitable frequency and amplitude is applied to *X* plate. The waveform is then obtained on the screen. Now the frequency of time base is adjusted to some exact submultiple of frequency of the applied a.c. voltage such that a steady pattern is obtained. The number of complete a. c. voltage waves will be equal to the ratio of the frequency of a. c. voltage to the frequency of time base voltage.

(6) *Measurement of phase angle.*

When one sinusoidal voltage is applied to the *X* plates and the other to the *Y* plates, a Lissajous figure pattern is obtained on the screen. This pattern is a measure of the phase difference between the two voltages. Generally an ellipse is obtained when

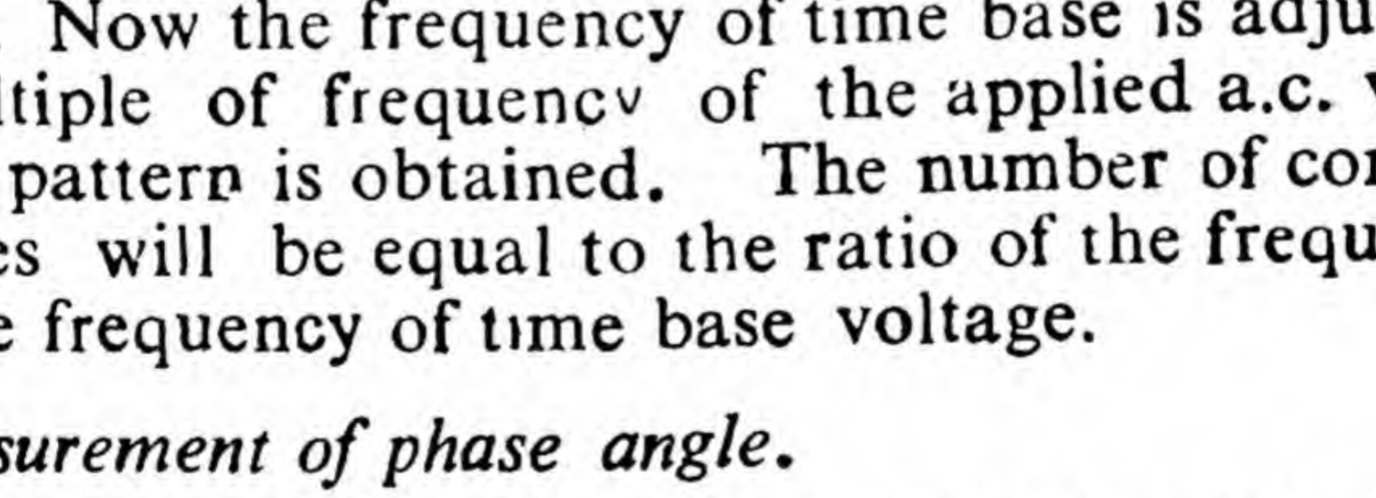


Fig 40. Lissajous figures on the screen of C.R.O.

the two voltages are of equal amplitude and of the same frequency. The orientation of ellipse with respect to coordinate axes depends upon the phase difference between the two voltages. To prove mathematically that the Lissajous pattern is an ellipse, let the voltages applied to the X and Y plates be $e_x = E_1 \sin(\omega t + \theta_1)$ and $e_y = E_2 \sin(\omega t + \theta_2)$ respectively. These expressions can be written as

$$\frac{e_x}{E_1} = \sin \omega t \cos \theta_1 + \cos \omega t \sin \theta_1 \quad \ldots(1)$$

$$\frac{e_y}{E_2} = \sin \omega t \cos \theta_2 + \cos \omega t \sin \theta_2 \quad \ldots(2)$$

The equation of the pattern can be obtained by eliminating the variable ωt. For this purpose we multiply equation (1) by $\cos \theta_2$ and equation (2) by $\cos \theta_1$, *i.e.*,

$$\frac{e_x}{E_1}.\cos \theta_2 = \sin \omega t \cos \theta_1 \cos \theta_2 + \cos \omega t \sin \theta_1 \cos \theta_2 \quad \ldots(3)$$

and $$\frac{e_y}{E_2}.\cos \theta_1 = \sin \omega t \cos \theta_2 \cos \theta_1 + \cos \omega t \sin \theta_2 \cos \theta_1 \quad \ldots(4)$$

Subtracting equation (4) from equation (3), we have

$$\frac{e_x}{E_1} \cos \theta_2 - \frac{e_y}{E_2} \cos \theta_2 = \cos \omega t (\sin \theta_1 \cos \theta_2 - \cos \theta_1 \sin \theta_2) \quad \ldots(5)$$

Similarly multiplying equation (1) by $\sin \theta_2$ and equation (2) by $\sin \theta_1$ and subtracting the two, we get

$$\frac{e_x}{E_1} \sin \theta_2 - \frac{e_y}{E_2} \sin \theta_1 = \sin \omega t (\cos \theta_1 \sin \theta_2 - \sin \theta_1 \cos \theta_2) \quad \ldots(6)$$

Squaring and adding equations (5) and (6), we have

$$\frac{e_x^2}{E_1^2} + \frac{e_y^2}{E_2^2} - \frac{2e_x e_y}{E_1 E_2} \cos(\theta_1 - \theta_2) = \sin^2(\theta_1 - \theta_2) \quad \ldots(7)$$

This is the equation of an ellipse. The orientation of the ellipse with respect to coordinate axes depends upon the phase difference $(\theta_1 - \theta_2)$ between the two voltages. When $(\theta_1 - \theta_2) = \pi/2$, the principal axes coincide with the coordinate axes. When $(\theta_1 - \theta_2) = 0$ or π the equation reduces to straight line. Various cases are shown in figure 40.

The phase difference between the two voltages can be calculated with the help of the pattern of C.R.O. by noting A and B as shown in figure 41. It may be noted that the distance B is the value of e_y when e_x is zero. Therefore, substituting this condition in equation (7), we have

$$B = E_2 \sin(\theta_1 - \theta_2) \quad \ldots(8)$$

in the same manner the value A is the maximum value of e_y and, therefore from the expression of applied voltage to Y plate

$$A = [e_y]_{max} = E_2 \quad \ldots(9)$$

Solving these expressions simultaneously, we get

$$\sin\ (\theta_1-\theta_2)=\frac{B}{A}$$

The phase angle ϕ is given by

$$\phi=(\theta_1-\theta_2)=\sin^{-1}\ \frac{B}{A}$$

(7) *Measurement of frequency* : The unknown frequency (voltage) is applied to one pair of deflecting plates and a known frequency from some variable frequency standard oscillator is applied to the other deflecting plates. The frequency of standard oscillator is varied until a single loop stationary pattern is obtained. The frequency of the sinusoidal voltage is then same as that of the oscillator.

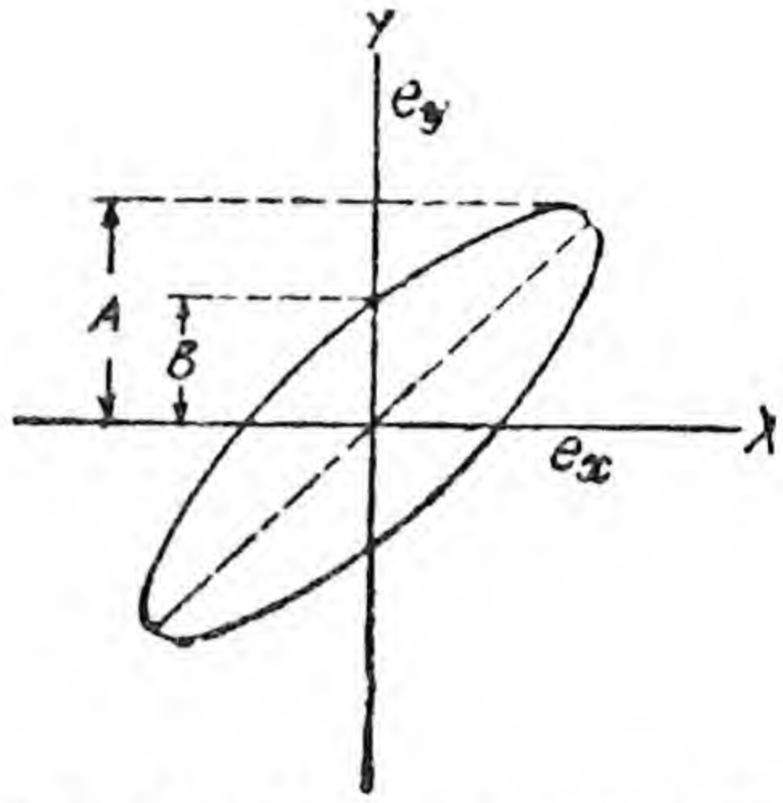


Fig. 41. Measurement of phase angle.

5·8. KLYSTRON :

It is a special ultra-high frequency tube [fig. 42], capable of generating, amplifying and detecting radio waves ranging in frequency from 600 to 30,000 mc/s. Klystron makes use of the velocity modulation principle. They use resonant cavities as tuned

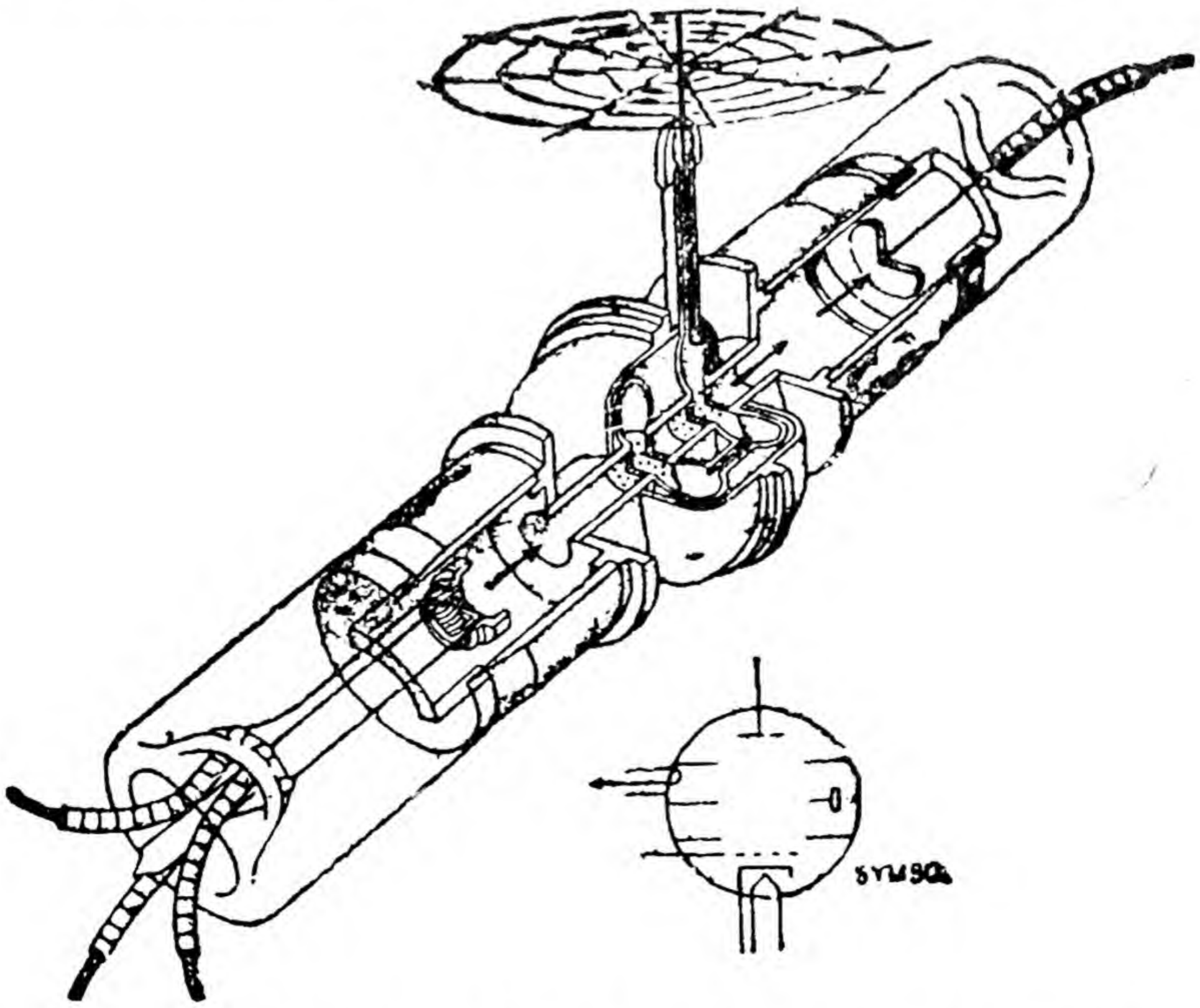


Fig. 42. Cut away view of a two resonator klystron oscillator.

circuits and a velocity modulated electron beam to sustain the energy field within the cavity.

A cut away view of a two resonator klystron oscillator is shown in figure 42. The electron emitted from the cathode are accelerated towards the main body of the tube, where they pass through a tube and then through a set of grids in a cavity resonator. Because of alternating axial field existing within cavity resonator, some of the electrons, while passing through the first resonator, are accelerated and some slowed down. This velocity modulation action causes the electrons to form in bunches by the time they pass through the grids of the second resonator. Due to the impact of the bunches on the grids. kinetic energy of electrons is converted into high frequency electromagnetic energy of the second resonator. The output is drawn from this resonator.

5·9. MAGNETRON :

The magnetron is a vacuum tube whose current is controlled by means of magnetic field strength. The tube is used for generation of ultra-high frequency oscillations. With its special form, it may oscillate at wavelength from 100 to 1 cm.

Generally two forms of magnetron tube are used. Early form of the tube were of the split anode type. Modern super high frequency magnetrons are made in the form of a multianode cavity. The basic structure of such magnetrons is shown in figure 43.

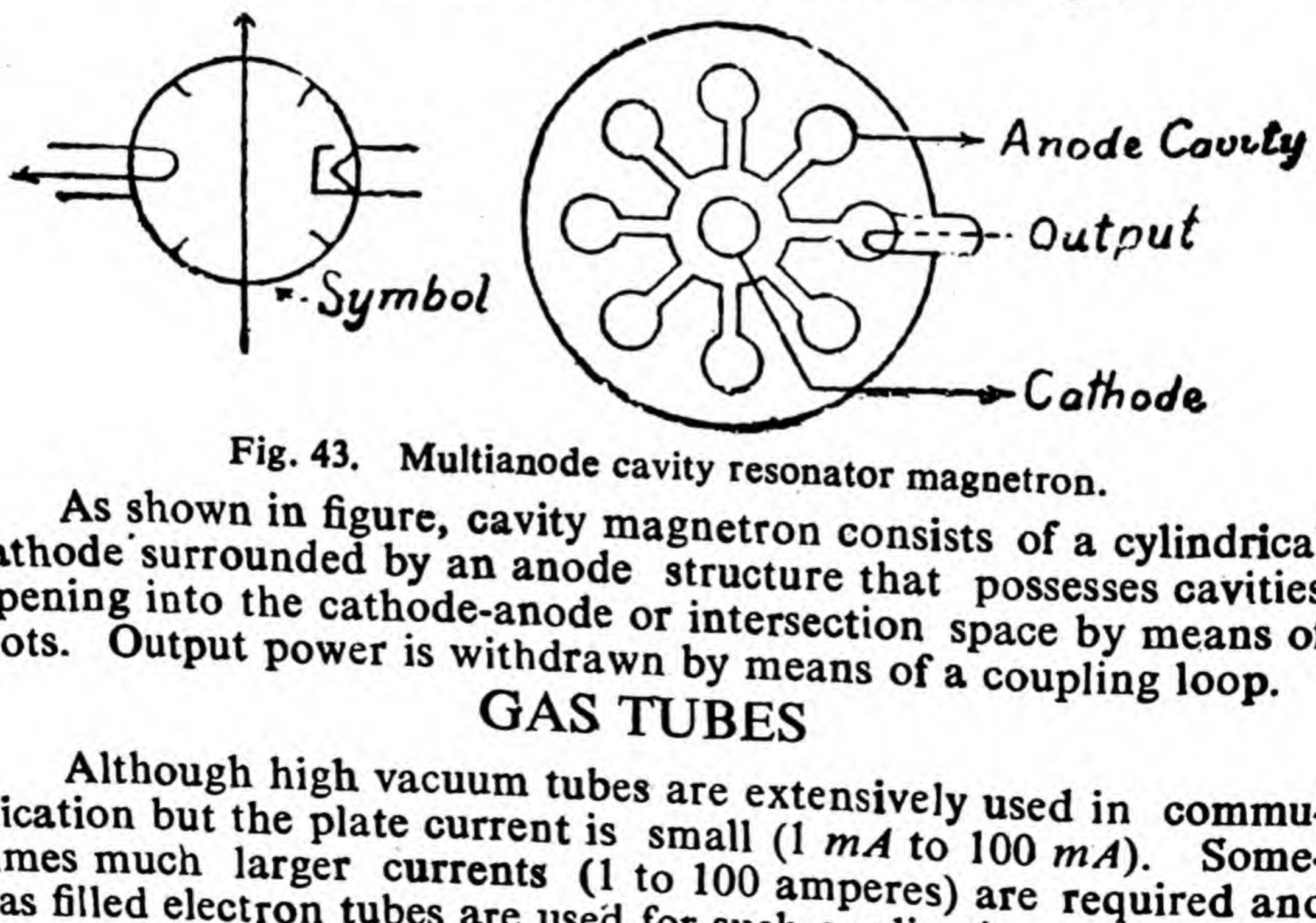


Fig. 43. Multianode cavity resonator magnetron.

As shown in figure, cavity magnetron consists of a cylindrical cathode surrounded by an anode structure that possesses cavities opening into the cathode-anode or intersection space by means of slots. Output power is withdrawn by means of a coupling loop.

GAS TUBES

Although high vacuum tubes are extensively used in communication but the plate current is small (1 *mA* to 100 *mA*). Sometimes much larger currents (1 to 100 amperes) are required and gas filled electron tubes are used for such applications.

The introduction of an inert gas, such as argon, or neon at a low pressure about 10^{-3} to 10^{-5} atmospheres, or a metallic vapour,

such as mercury markedly alters the properties of the electron tube. When high velocity electrons move through the tube, the gas for molecules are ionized due to collision and as a result charged particles are formed. These charged particles move under the influence of voltage across the tube and collide with other gas molecules. This produces further ionization. In this way a large number of positiveiy charged ions are formed which move toward the cathode and neutralize many electrons constituting the space charge. Thus the influence of space charge over the electron flow in the tube is reduced.

In gas diode, when the plate voltage is sufficiently high (greater than the ionisation potential of the gas), a larger plate current is obtained. For a given plate to cathode potential, the plate current in a gas diode is much larger than it would be in a vacuum diode. Thus gas diodes are extensively used in a high power systems in-place of high vacuum tube.

The gas tubes may be classified into the following categories:

(*i*) Thermionic gas diodes.

(*ii*) Thermionic gas triodes (thyratron).

(*iii*) Cold cathode tubes :

(*a*) cold cathode gas diodes.

(*b*) cold cathode gas triodes.

(*iv*) Tubes using mercury pool cathodes.

(*a*) Mercury arc rectifiers.

(*b*) Ignitrons.

(*c*) Excitrons.

5·10. EFFECT OF GAS IN A THERMIONIC DIODE :

The characteristics of a thermionic diode are altered even if the small fraction of a gas is present in it. The presence of oxygen and water vapour affects the cathode (*i.e.* tbey have chemical effects on it) and hence decreases its emission. On the other hand, the presence of mercury vapours and inert gases (helium, neon, argon, krypton and xenon) increases the current as the space charge is neutralised. Let us consider the tungsten filament tube with a small amount of gas in it. The plate voltage and plate current characteristic is shown in fig. 44. The curve 1 corresponds to the characteristic in the absence of the gas while curve 2 refers to the current voltage relation with the gas at low pressure. With the introduction of more gas, the pressure is increased and the curve 3 shows the characteristic.

It is observed from fig. 44 that in the presence of the gas, the

curve rises but reaches the same saturation current. Furthermore the plate voltage also affects the current.

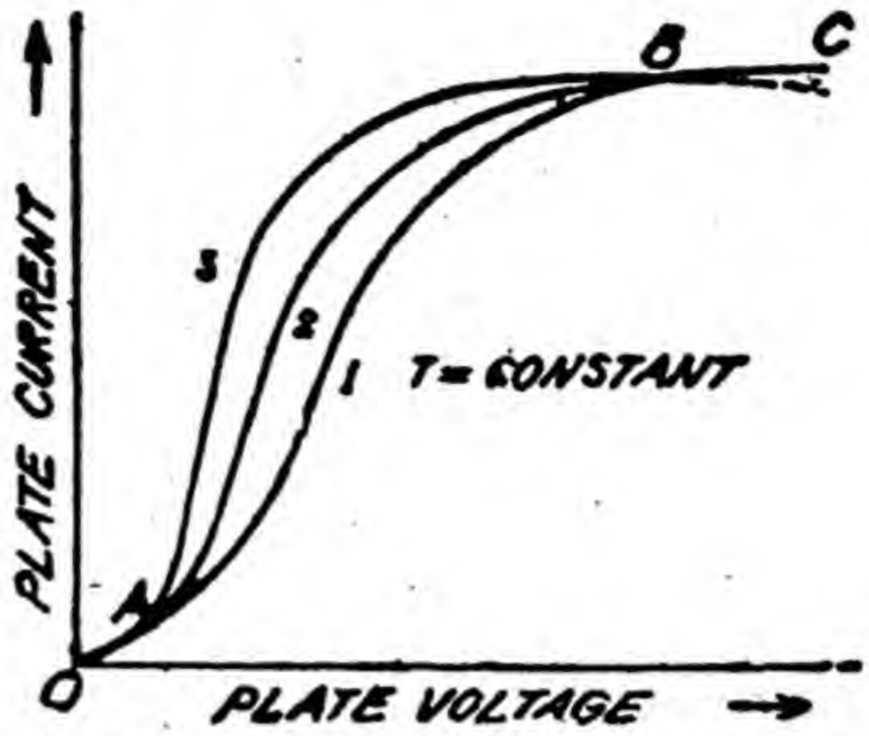


Fig. 44. Current-voltage curve of diode with varying gas pressure

It is also clear from fig. 44 that the gas does not have any appreciable effect in the region *OA* where the plate voltage is less than the ionization potential of the gas The major effect of the gas is in region *AB* where, in the absence of the gas, the current is limited by the space charge. The neutralisation of space charge permits more electrons to reach the plate and the current reaches the saturation value at much lower plate voltage. The region *BC* is changed only slightly. Here it should be remembered that the effect of the gas in region *OA* is small does not mean that there is no departure from high vacuum curve for voltages below ionisation potential of the gas or a departure must occur at the ionisation potential. There are other factors which affect the characteristic in the region *OA*.

5·11 GAS DIODES WITH THERMIONIC CATHODES :

Thermionic gas field diode consists of a thermionic emitter (cathode) and a cold electrode (plate). The two electrodes are enclosed in a sealed glass envelope. The tube is evacuated and then mercury vapour or some inert gas like helium, neon, argon, or xenon at the pressure of about 10^{-5} atmosphere is inserted in the envelope. The cathode is specially designed to give a very large emission to provide large current. This can be achieved by increasing the cathode temperature as well as the area of the emitting surface. Thus the cathodes are constructed such that heat radiation from one section falls over the other section, thereby saving much radiated power Such cathodes are known as *heat shielded cathode*. In gas field tube, ionized region (plasma) extends from plate to within a short distance of the cathode Due to low field strength in plasma region electrons can move in all directions even round the corners, and hence plate need not be close to the cathode, while in vacuum tubes the plate is kept close to the cathode in order to reduce the space charge effect. Figure 45 shows the plate voltage plate current characteristic of a hot cathode gas diode. The curve 1 refers to the vacuum tube while the curve 2 is of the gas diode. The nature of the curve is the same as in a vacuum tube *i.e.* the plate current increases with increase of the plate voltage. We know that the ionisation potential of the gas is usually 1 to 15 volts and at this voltage, the ionisation of the gas takes place and the plate current becomes maximum as is indicated by the steep portion

of the curve 2. This is due to the fact that the positive ions produced by the ionisation of the gas neutralise the electronic space charge completely and thereby eliminate the limitation of the flow of the current in the tube.

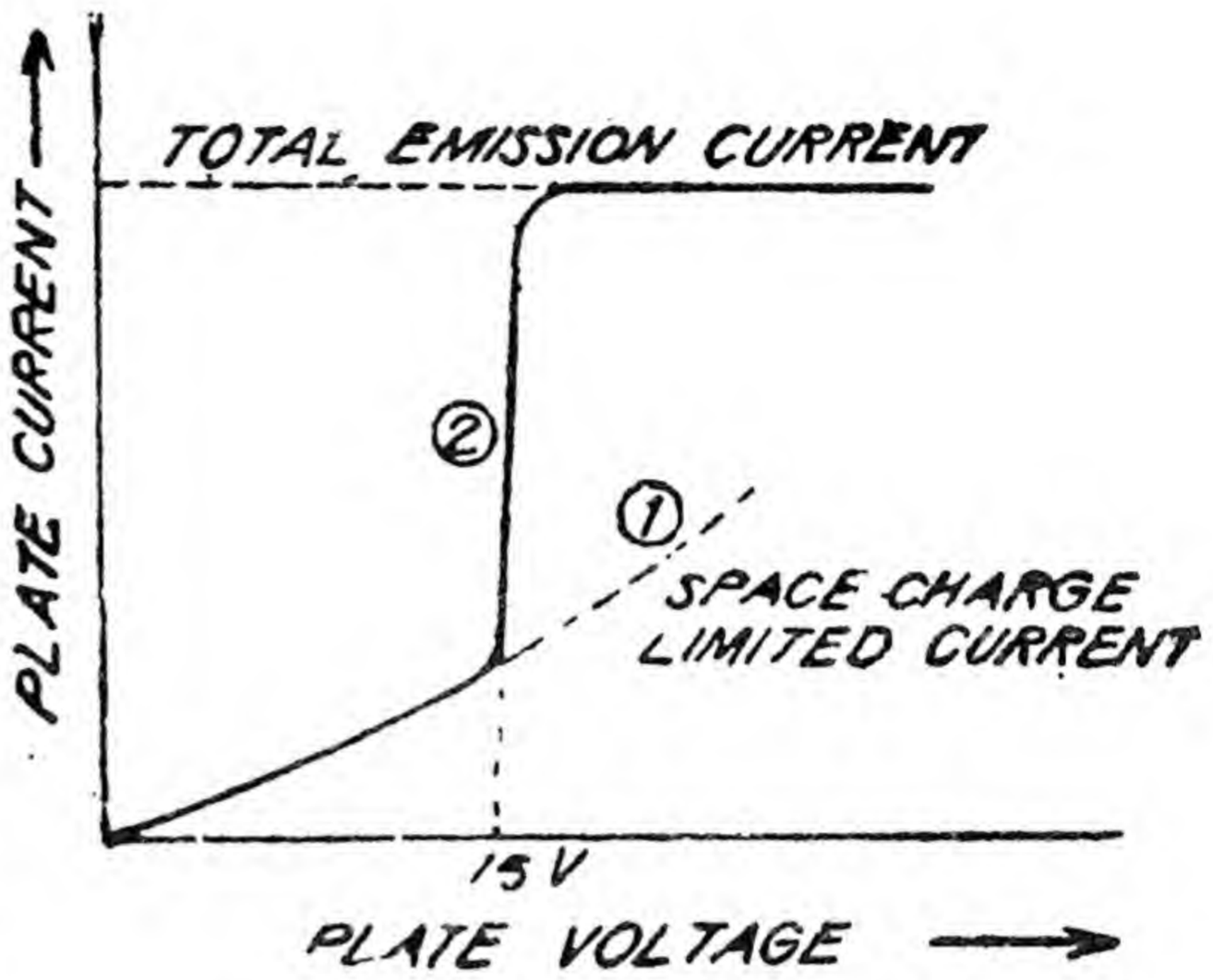


Fig. 45. Current voltage curve of a gas diode.

Sudden increase in curve 2 may be explained as follows. At ionization potential, the ionization of few gas molecules takes place due to the collision of electrons with gas molecules. The electrons so produced by ionization travel towards the plate and thereby increase the plate current. At the same time positive ions travel towards the cathode and neutralise the space charge. This results in a larger flow of electrons towards plate, larger ionization, further increase in current and further reduction in space charge. The process is cumulative. Finally, fully emission current flows without requiring any increase in plate voltage.

In order to neutralise the electron space charge completely, the number of positive ions should be equal to the number of electrons. We know that the rate of emission of electrons is much larger than the rate of ion formation, even then the above mentioned condition is possible in a gas diode. This is due to the fact that the mass of positive ion is much larger compared with the electron and the electric field removes electrons more quickly than heavy positive ions. The current contributed by positive ions is however very small.

Potential distribution in a gas dio e : The potential distribution in a gas diode is shown in fig. 46. Here it is assumed that the cathode and plate of the diode are in parallel plane. The potential distribution curve for free space is a straight line and that for a vacuum tube is such that potential at any point is proportional to the (4/3) power of the distance of the point from the cathode

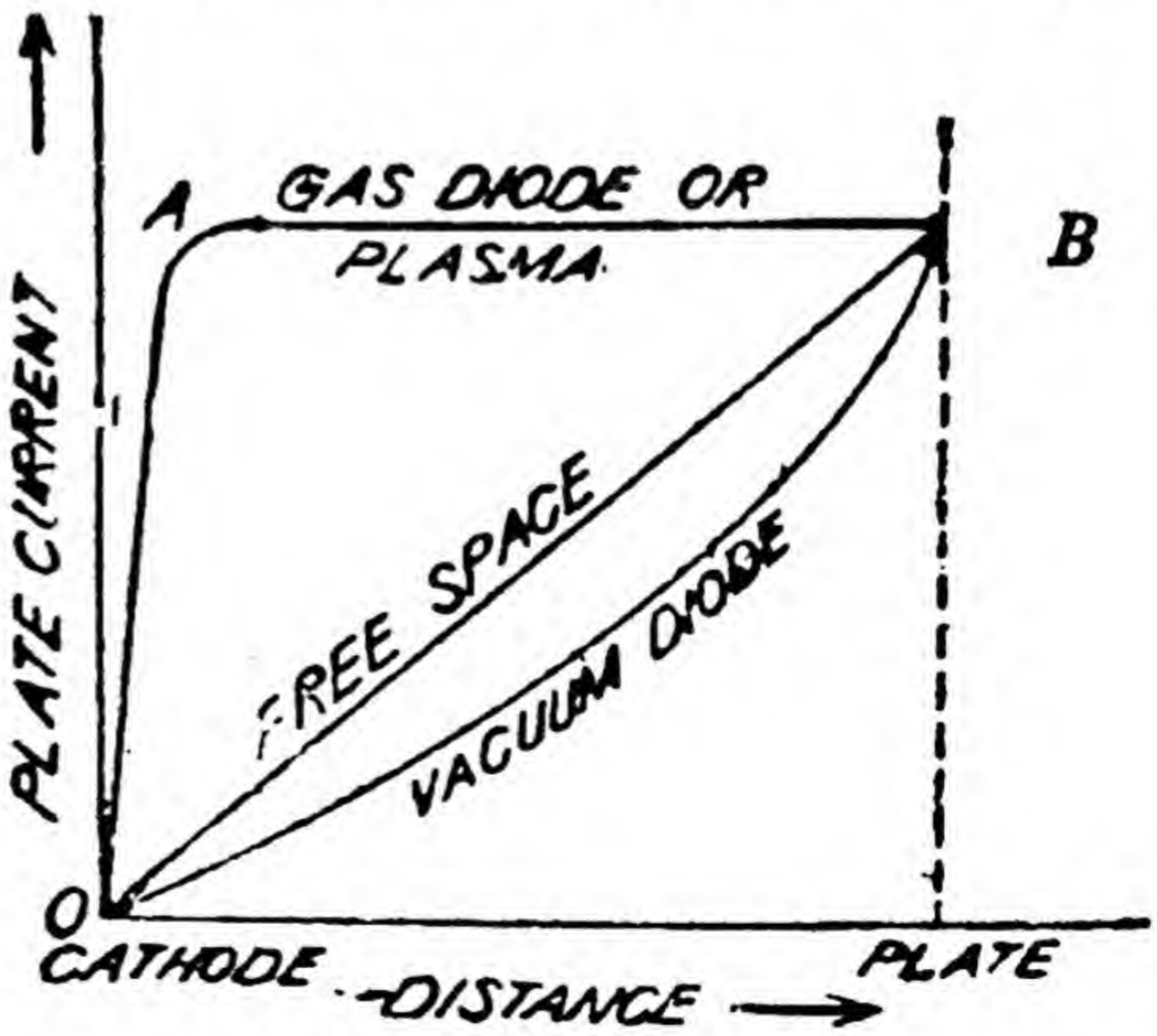


Fig. 46. Potential distribution in gas diode.

i.e. it is expressed as $K(x)^{4/3}$, where K is a constant. The positive ions increase the current, and the greater current increases the production of positive ions and so on, until positive ions are produced at such a great rate that their charges tend to raise the potential at some point in the tube above the plate potential. *Thus plasma is at higher potential to that of the plate due to the excess of positive ions.* Under this situation, the current is no longer limited by space charge and the characteristic of the tube is the typical of gas tube.

The interelectrode space in a gas diode can be divided into two regions. The first is *OA* and is called a *positive ion sheath* or *cathode sheath*. The second is the region *AB* and is called the *Plasma.* In the positive ion sheath, electrons emitted from cathode are accelerated by electric field and are projected into the plasma. The positive ions from plasma reach the sheath and are finally drawn to cathode. Due to the constant potential of plasma, the voltage drop across the sheath is essentially equal to that across the tube. This sheath voltage is often called the cathode voltage drop.

The velocity of the positive ion is less than that of electron as positive ion is heavier. Due to this fact, the space charge in region *OA* is due to positive ions There is almost zero space charge in this region. The space charge neutralisation in the gas tube is responsible for special type of electrode structures.

The plasma may be thought of a mixture of (i) an electron gas, (ii) a positive gas and (iii) a molecular gas The length of the plasma adjusts itself so as to fill the tube, except for the short space required by the positive ion sheath about the cathode.

5·12. EFFECT OF GAS PRESSURE ON THE CHARACTERISTIC OF A GAS DIODE :

The pressure of the gas affects the characteristics of the gas diode in the following three ways :

(i) the rate of evaporation of the cathode,

(ii) the drop of voltage while conducting current. and

(iii) the breakdown voltage in the reverse direction.

If the pressure of the gas is high enough, the rate of evaporation of the cathode decreases and the cathode can operate at high temperature safely. If the pressure is too high, breakdown followed by a glow or arc discharge may take place in a reverse direction during the inverse part of the cycle of alternating voltage across the tube. Under these circumstances, the tube will conduct in both directions and ceases to rectify. When the pressure is very low, then the gas molecules may not be sufficient to neutralise the space charge. In this way for any gas filled in the thermionic tube, there is a definite value of pressure above which the allowable peak inverse voltage of the tube is lowered, and below which the peak anode current is decreased.

In order to increase the inverse voltage rating of a gas tube it is desirable to reduce the gas pressure. Too low pressure is also undesirable because the ionization of the gas will be very small and consequently the number of positive ions required to neutratlise the space charge will be small. Thus for every gas field thermionic tube there is an operating range of pressure.

In hot cathode diodes, generally mercury vapour is used because it remains at an equilibrium pressure depending upon the temperature of the liquid mercury. The pressure thus depends upon the temperature of the condensed mercury vapour. The variation of allowable peak inverse voltage and peak plate current of a hot cathode mercury vapour tube as a function of temperature of the condensed mercury is shown in figure 47. It is seen from the figure that at a temperature about 80°C, the peak inverse voltage becomes very small (about 2 kv) and as the temperature is reduced to 25°C, the peak plate current becomes small. The peak current that may be drawn from the tube is less than the rated peak current of 600 amps, since otherwise the tube drop would exceed the disintegration value of 22 volts. Mercury vapour pressure may be maintained within narrow limits and thus mercury vapour is used in high voltages gas tubes.

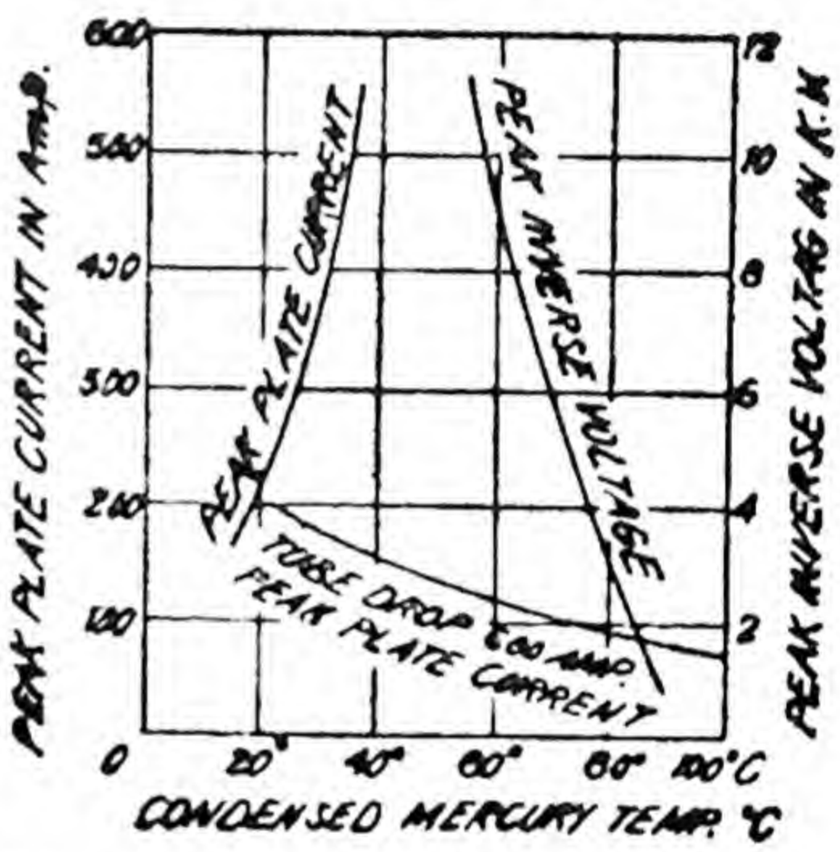


Fig. 47. Variation of peak plate current and peak inverse voltage as a function of condensed mercury temp.

Inert gases can be used in low voltage gas tubes only where large variation in gas pressure is permitted. In case of inert gases, the pressure does not depend upon temperature, hence they have an advantage over mercury vapour tube.

5·13. GAS FILLED THERMIONIC TRIODES OR THYRATRONS :

Thyratron is the commercial name for a thermionic gas filled triode. The configuration of different electrodes in a thyratrons is different than that in a vacuum triode. In thyratron the filament is either filamentary or indirectly heated. The plate is in the form of a disc of nickel or it is in the form of a block of graphite depending upon the current rating of the tube. The grid is in the form of a cylinder surrounding the plate and provides an electrostatic shielding between the plate and the cathode.

Although the control grid in thyratron acts in a similar way as in vacuum triode but its control action is widely different due to the presence of the gas. When grid is highly negative, the tube does not conduct because the electrons flow and consequently, the gas ionization is prevented even at high plate voltage. Now keeping the plate voltage constant as the grid potential is varied in a positive directions, a point is reached when the tube starts conducting.

The plate voltage now reduces to a value equal to the ionization potential. The grid voltage at which thyratron starts conducting is known as cut off value. The ionization results in the abrupt development of the plate current and the plate current assumes the full value limited by the external resistance in series with tube. When the tube starts conducting, the control grid loses all controls which may be explained in the following way :

When thyratron conducts positive ion sheaths are formed around athode and negative grid, and the remaining space of the tube is filled with plasma. In plasma, the number of positive and negative ions are equal and the potentials is almost constant which is equal to the plate potential. The grid attract the positive ions of the plasma towards it and a sheath of positive ions is formed at its surface. Under the equilibrium conditions, the net charge in the sheath is zero because the total positive charge of all positive ions is equal to the negative charge on the grid. Now through the plasma, positive ions flow to cathode and electrons to plate without experiencing any influence of the grid. In other words, the positive ions near the grid completely shield the electric field of the grid irrespective of voltage *i e*. grid has no effect on plate current.

The conduction in the tube does not cease even the control grid is made highly negative. The current flow can be stopped only by lowering the plate potential. When ionization ceases, the control grid regains its control and it must again be raised to firing point before a plate current flow can occur. The characteristics make the thyratron useful as a power switch, wherein a small impulse applied to the grid can control a very large amount of power in the plate circuit. A large resistance is included in series with the grid to limit the current in the grid circuit because due to the positive ions, the grid current may become quite large and causes wastage of power.

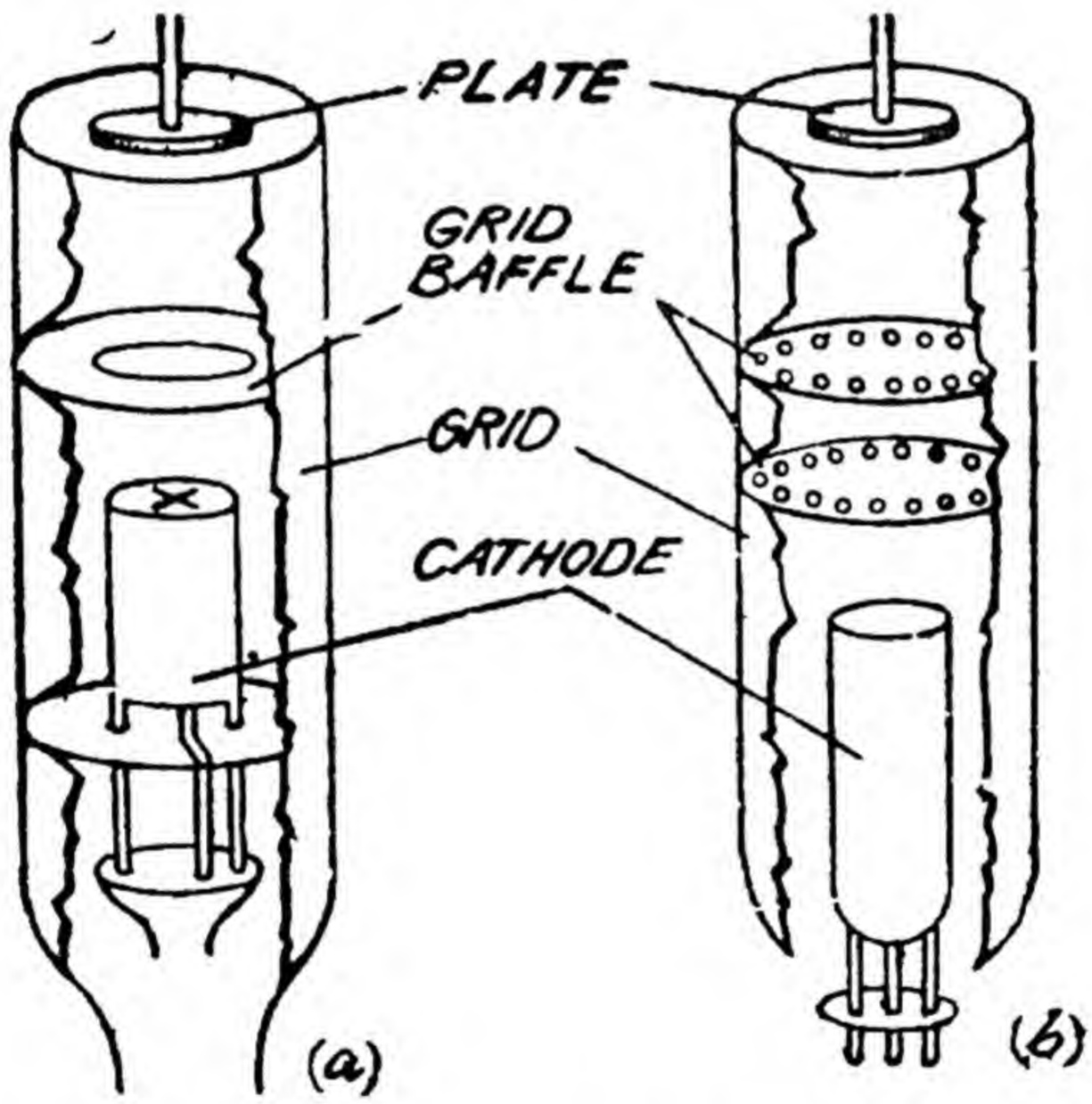
In thyratron, once the breakdown occurs and arc is formed, the potential between anode and cathode falls to ionizing potential.

A baffle or a number of baffles having small holes are introduced between plate and cathode. There are two types of thyratrons :

(*i*) *negative control thyratron* : It has only one baffle with large hole in it which is a part of grid. The breakdown occurs with negative grid voltage. This is shown in fig. 48 (*a*).

(*ii*) *positive control thyratron* : It has one or more baffles with large number of small holes so as to provide better shielding. The breakdown occurs when the grid is positive. This is shown in figure 48 (*b*).

The grid in negative grid thyratron controls the maximum current density instead of average current density across its area. When the current density reaches at a critical value the tube fires. When there are a large number of holes in the grid of unequal size, then the current through the largest hole controls the starting and other holes play no part. Due to this fact thyratron grid is constructed in the form of a particular disc with a single large hole. For the same reason, the critical current density may occur along a path surrounding the grid disc instead of grid hole, hence thyratron grid is constructed as to surround almost either the plate or cathode. It is desirable that the grid should shield both the plate and cathode from the glass walls. If the grid shields only one electrode then space between the grid and the other electrode is influenced by the electric field produced by the charges on the glass walls. The charges are fluctuating and so influence the starting characteristics.



(*a*) Electrode structure of a negative control thyratron. (*b*) Electrode siructure of a positive control thyratron.

Fig. 48.

Thus the difference between a vacuum triode and thyratron is that in a vacuum triode the plate current is controlled by grid voltage while in cases of thyratron, the grid voltage can only initiate the arc and then loses its control over the flow of plate current after the arc is produced.

The circuit arrangement for drawing the control characteristics of a thyratron is same as that of a vacuum tube. For a particular value of plate potential, the value of grid potential required for conduction just to begin the

arc is known as critical grid potential. For negative grid thyratron, control characteristic curves with temperature as a parameter are shown in figure 49. It is observed from these characteristics that these change with temperature. The temperature dependence is much pronounced with mercury vapour as the pressure of mercury vapour changes considerably with temperature.

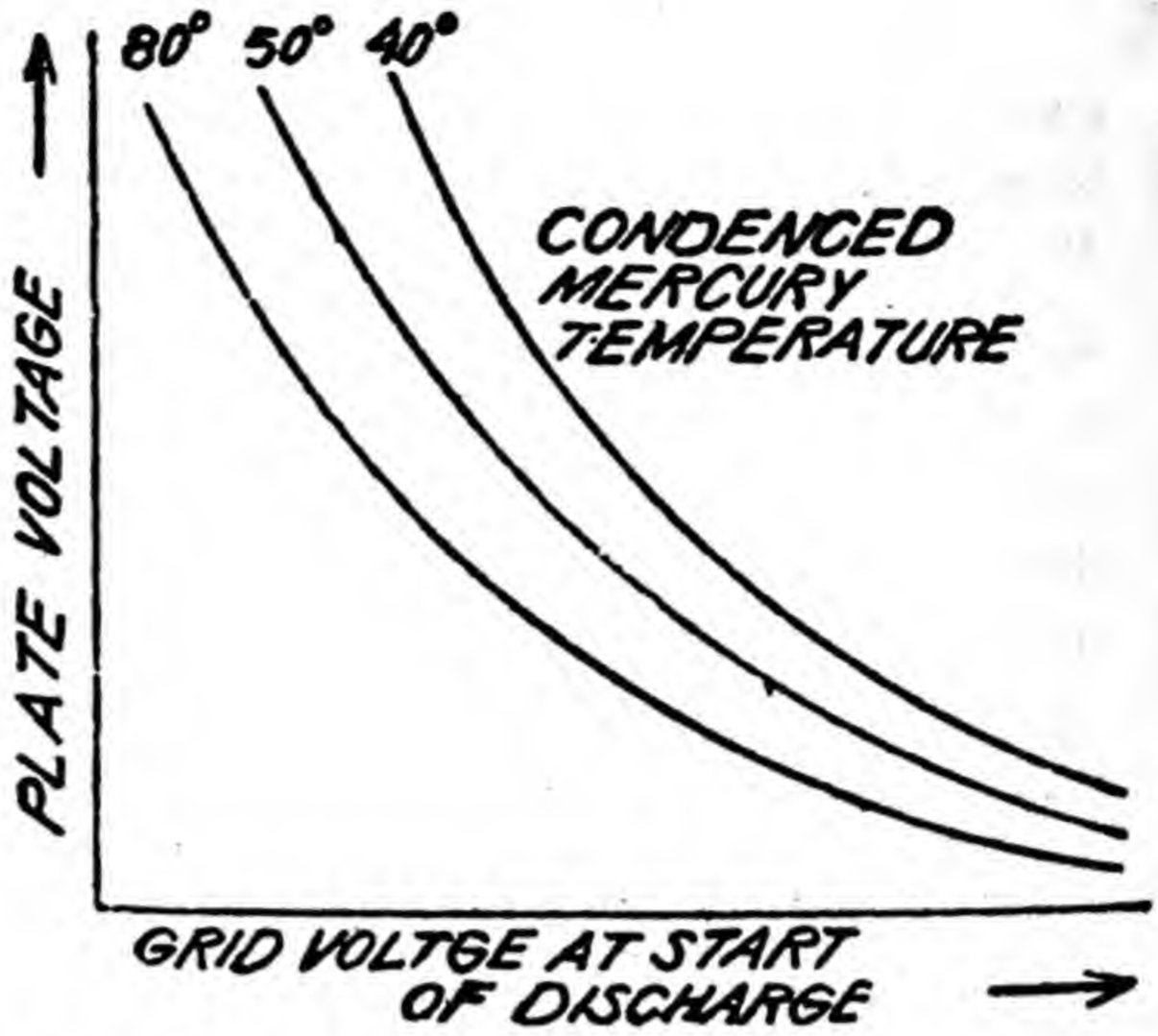


Fig. 49. Characteristics of a negative thyratron.

Selection of a gas tube :

While selecting a suitable gas in a gas tube, following two points are kept in view :

(*i*) the gas should not react chemically with the electrodes. For this purpose inert gases are most suitable, and

(*ii*) there should be a suitable ionization potential of the gas. It is observed that when the gas is conducting, the voltage drop across the tube is usually of the order of ionization potential and the ionization potential decreases with the atomic weight. Thus for high efficiency, mercury vapour or one of the heavier inert gases is preferred.

Tube containing mercury vapour can carry large current and, hence they are suitable for industrial applications such as controlled rectifiers. For high speed working one of the inert gases is used.

Ionisation and de-ionisation times of thyratron :

Ionisation time is the time for the plasma to form and anode current to reach its final value after the application of anode voltage. It is a function of tube structure, electrode voltage, pressure of the gas, and associated circuit. It varies from a micro second to several microseconds.

De-ionisation time is the time required for the plasma to disappear and grid to gain control after the removal of plate voltage. It varies few micro-seconds to several hundred micro-seconds. It depends upon the same factors as for ionisation time. Deionisation time is small in tubes with small grid holes and hence positive grid thyratron are generally used for applications requiring small de-ionisatron time.

5·14. COLD CATHODE GAS DIODE : VOLTAGE REGULATOR TUBE

Cold cathode (no heating arrangement for cathode) gas diode

or V-R Tube is a common glow tube which is designed to keep the voltage across it constant irrespective of supply voltage or the load current variation. It consists of two electrodes in sealed glass envelope with neon or some other inert gas at a pressure of about 0·1 m.m. of mercury. The cathode is a hollow cylinder which has large area and maintains the constancy of the operating voltage. The anode consists of a cylindrical wire placed along the axis of the cathode fig. 50 (a). The pressure of the inert gas is so kept that the breakdown voltage is only a few volt greater than the drop across the tube after the breakdown. To keep the breakdown voltage as low as possible a small wire projects out from the cathode to within a small fraction of an inch of the anode. Initially the breakdown occurs at the end of the wire and the anode. The projected wire is known as starter. The starter serves two purposes : (i) it reduces the breakdown potential ; so a low potential is required to start the discharge, and (ii) it maintains the glow spot when the current is small. Once the discharge starts, it will spread to the other portions of the cathode depending upon the current.

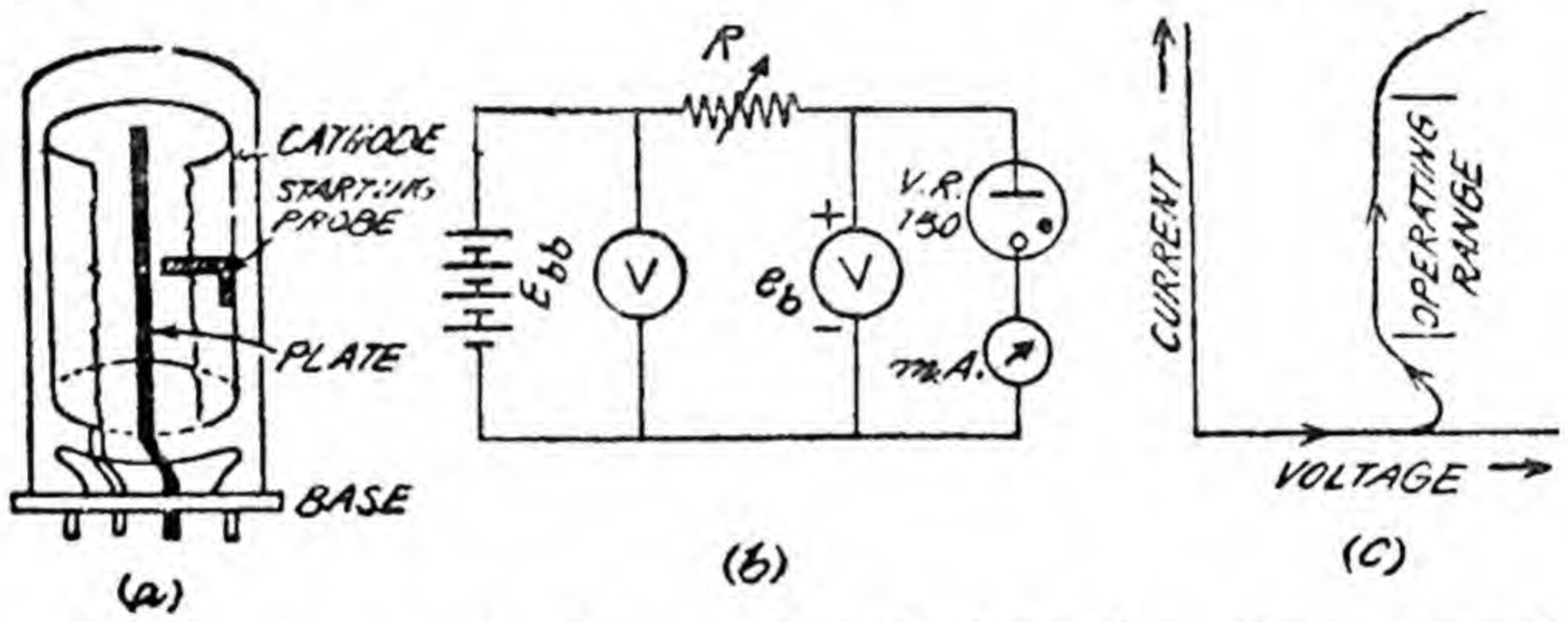


Fig. 50. Electronic structure of V-R tube. Circuit for finding the characteristic of V-R tube. Current voltage characteristic of V-R tube.

The circuit for finding the characteristic of V–R tube is shown in fig. 50 (b). The supply voltage E_{bb} is slowly increased, and current through the tube as well as voltage across it are measured. Till the voltage across the tube does not reach the breakdown potential, the milliammeter shows zero reading. Near the breakdown potential, a glow appears at the tip of the starter and runs to the cathode. Now the milliammeter shows a current. The voltage across the tube slightly falls from the breakdown voltage as is clear from 50 (*c*). In other words, the glow is initiated at a voltage slightly higher than the normal operating voltage.

When the tube fires, the current in the circuit is controlled by *R*. Now *R* is slowly changed and current and voltage across tube are measured. The curve is almost a vertical line showing thereby that the tube is working in the grow region where the voltage across the tube is substantially constant.

The tube finds application in electronic regulated power supply to provide constant reference voltage. The voltage that may be regulated by a V-R (150) tube is 150 volts.

5·15. TUBES WITH MERCURY POOL CATHODES :

When large currents are to be conducted, mercury pool cathodes are used in gas tubes. The mercury pool is not itself a source of electrons but by means of auxiliary electrode an arc is started. The arc concentrates on a small bright spot which moves erratically over the surface of the mercury pool. The mercury pool serves two purposes ; (*i*) it supplies the electrons, (*ii*) it supplies mercury vapour through which the current is conducted.

Figure 51 shows a tube with a mercury pool cathode. Now-a-days it is obsolete but due to its simple construction, we shall describe it here. The tube is in the form of a large evacuated glass bulb with a pool of murcury (*K*) which serves as the cathode. Near the mercury pool *K* there is another pool of mercury (*A*) in lower side arm known as the starter electrode. Initially there is no connection between these two mercury pools. The tube has two anodes of graphite sealed through the two arms of the glass bulb, and are situated at the upper side. When the anode voltage is applied the cathode is unable the supply the electrons. The supply of the electrons can be carried out by the starting electrode *A* in the following ways :

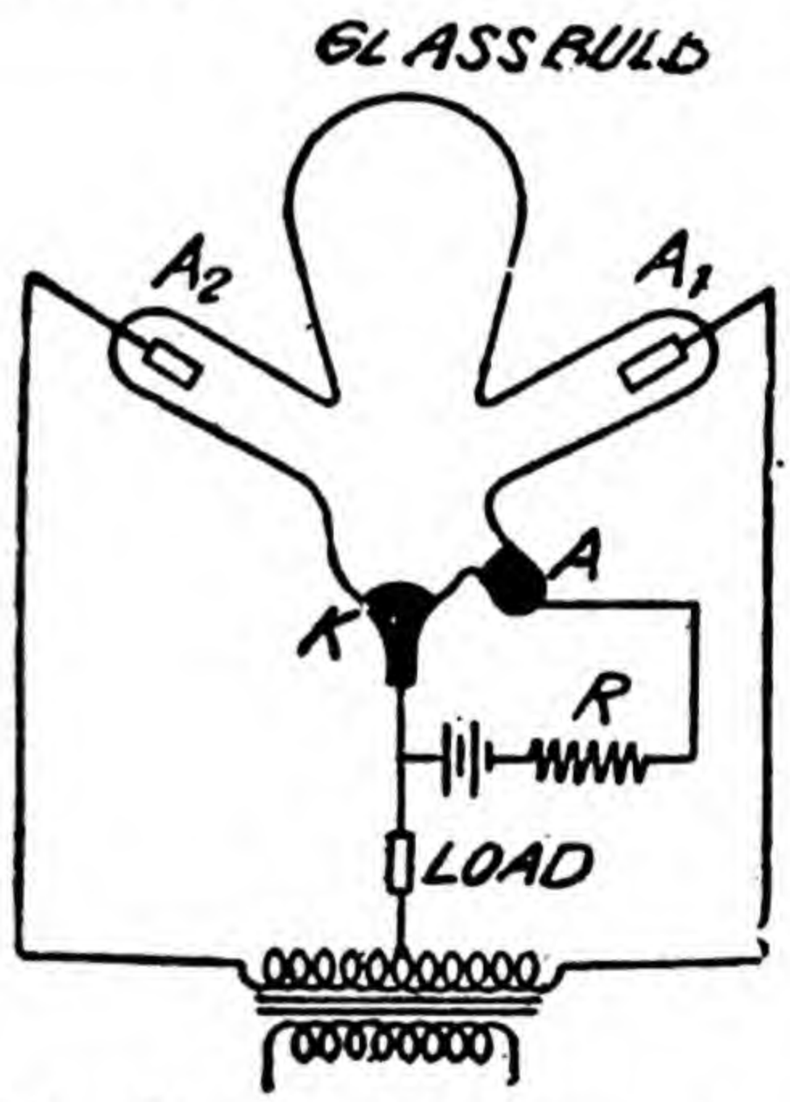


Fig. 51. Tube with mercury pool.

By tilting the tube, a contact is made between cathode *K* and starting electrode *A*. Now a current limited by resistance R flows between cathode-starter circuit. When the tube is set in upright position, the contact between cathode and starter is broken and a small arc appears. The heat of arc vaporises mercury which fills up the glass tube. The arc produces an incandescent space at the surface of mercury pool, which is called a cathode spot. This hot spot is the source of electron emission. The high velocity electrons on their way to anode produce ionisation of mercury vapour. The positive ions liberated during ionisation bombard the pool and maintain the temperature of the electron emitting cathode spot. In this way a continuous supply of electrons is maintained. Thus the cathode spot is a source of tremendous emission current density of the electrons and it is estimated to be about 4,00 amperes per square centimeter. When total current is large, several spots often appear, each carrying about 30 to 40 amperes.

The potential distribution in the arc of the tube is shown in fig. 52. The arc can be divided in three parts ; the cathode sheath

(shown as region *OA*), the plasma (shown as region *AB*) and the anode sheath (shown as region *BC*).

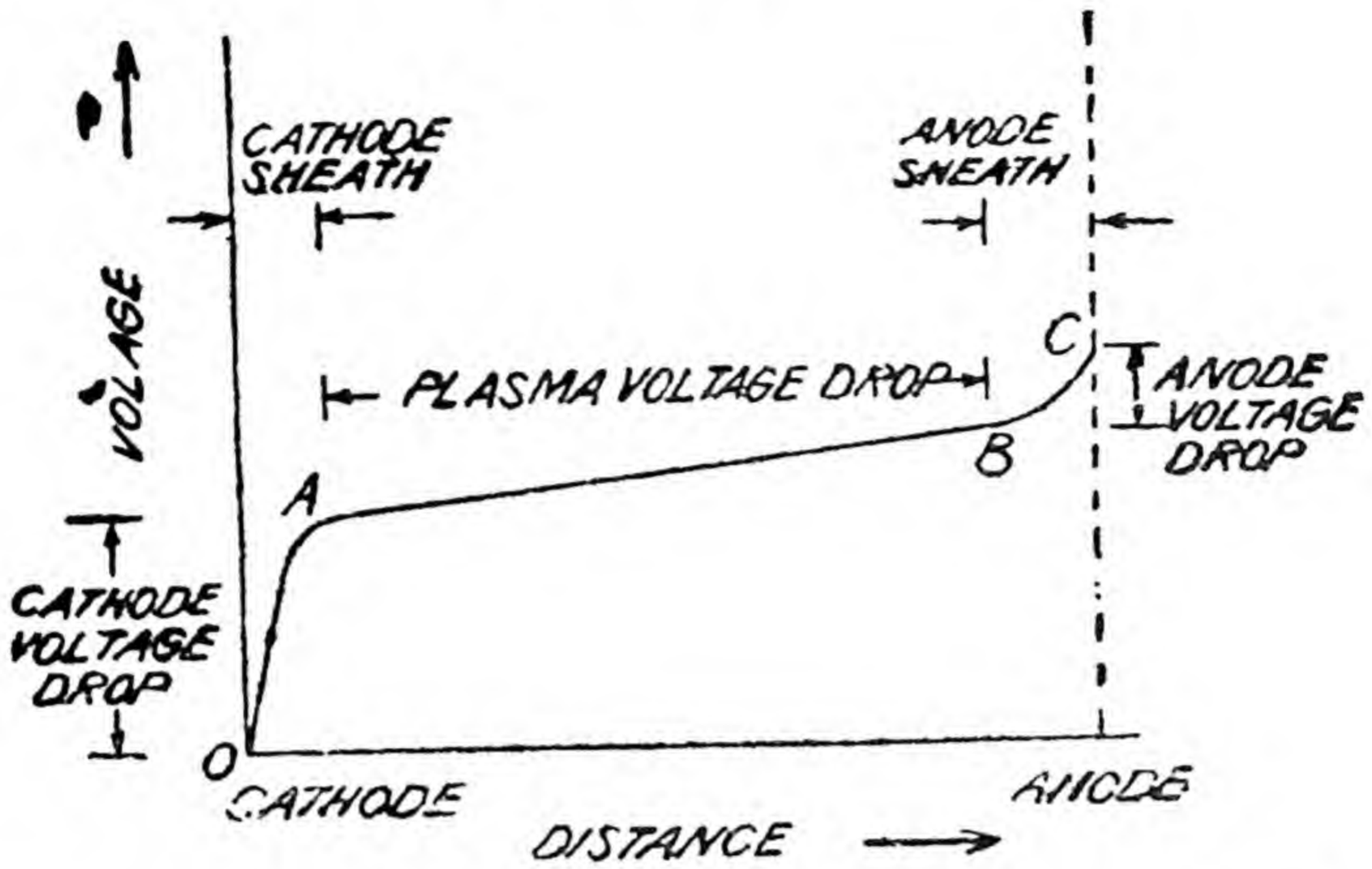


Fig. 52. Potential distribution in arc of mercury poo! cathode tube.

The cathode sheath surrounds the cathode. It consists of electrons moving from cathode to plasma and positive ions moving from plasma to the cathode. The voltage drop in this sheath is about 9·9 volts. Since the work function of mercury is 4·5 volts, and hence about 4·5 volts is used up in liberating the electrons and rest 5·4 volts is used to accelerate the electrons to produce ionisation.

The plasma in mercury arc is a region of low potential gradient. It contains positive ions and electrons which are in equal numbers. The potential gradient is about 0·05 and 0·2 volt per centimeter. The plasma chiefly serves a highly conducting path for the main length of the arc.

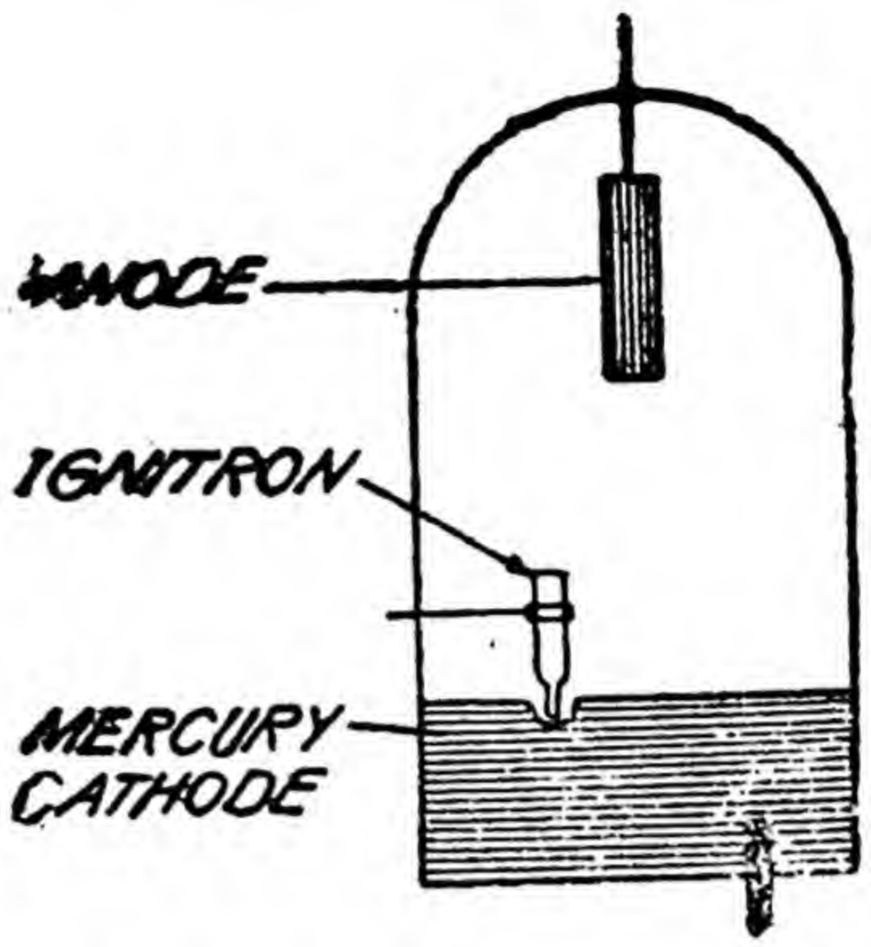
The anode sheath is a region surrounding the anode. Very few positive ions are found in this sheath and it has mostly electrons. The anode sheath has a high potential gradient. The electronic space charge surrounding the anode opposes the motion of electrons from plasma to anode. The anode voltage drop is the principal reason for an increased over all arc voltage drop at high currents.

The common types of arc tubes with mercury pool cathodes are (1) mercury arc rectifier; (2) ignitron and (3) capacitron.

5·16. IGNITRON :

Ignitron is a single anode mercury arc rectifier with a pool type cathode. Figure 47 shows the arrangement of Ignitron. In the Ignitron, the cathode consists of a large pool of mercury at the bottom of the tube. The starter is the Ignitron whose tip is immersed in the mercury pool cathode. The Ignitron is a rod made of semiconductor such as boron carbide or silicon carbide which has a relatively high electrical resistivity. Mercury does not wet these

materials and form a mensiscus at the point of contact of the ignitron. When a current near about 25 amperes, from an auxiliary source of about 200 volts is passed to the pool through ignitron a cathode spot is formed at the surface of the pool. The tube is fired and arc is formed if the anode is positive. When an a.c. potential is applied between the anode and the cathode the arc should be extinguished in each negative half cycle provided the arc is initiated regularly. Ignitrons are used widely in controlling the heat current in resistance wel ers since the valve gives current in short pulses of large peaks.



Flg. 53. Showing the arrangement of Ignitron.

PHOTO ELECTRIC DEVICES

5·17. CLASSIFICATION OF PHOTO ELECTRIC (light sensitive) DEVICES :

The function of photo-electric devices is to convert light or radiant energy into electric energy. There are three types of photo-electric devices : (*a*) photo-emissive, (*b*) photovoltaic and (*c*) photo-conductive.

(a) Photo-emissive device : In this device the emission of electrons takes place from the surface of photo sensitive metals when exposed to light waves.

(b) Photo-voltaic device : This device generates power *i.e.*, an e.m.f is generated when two substances are in close contact and the boundary separating the two is exposed to light radiation.

(c) Photo-conductive device : Devices which change their resistance by the action of incident light radiation are known as photo-conductive devices.

All three devices are used for industrial systems.

5·18. PHOTOEMISSIVE CATHODES :

The emission current from any metal is a function of the frequency of incident radiation. It is therefore useful to consider the relative response for different wavelengths. Before considering the relative response, we would like to define the *sensitivity*. The sensitivity is defined in terms of current per unit of radiant power striking the metal surface. This is denoted as microamperes per micro watt of incident power at a particular frequency. Light is the radiant elcetromagnetic energy having wavelengths ranging from 100 Å to 100×10^4 Å. The extremes of the wavelength's range lie in the far ultra violet and the far infrared. The visible range extends from 4000Å to 7000Å which is in between the ultra violet and infrared. Now we come to the consideration of relative

response for different wavelengths. Figure (54) shows the relative

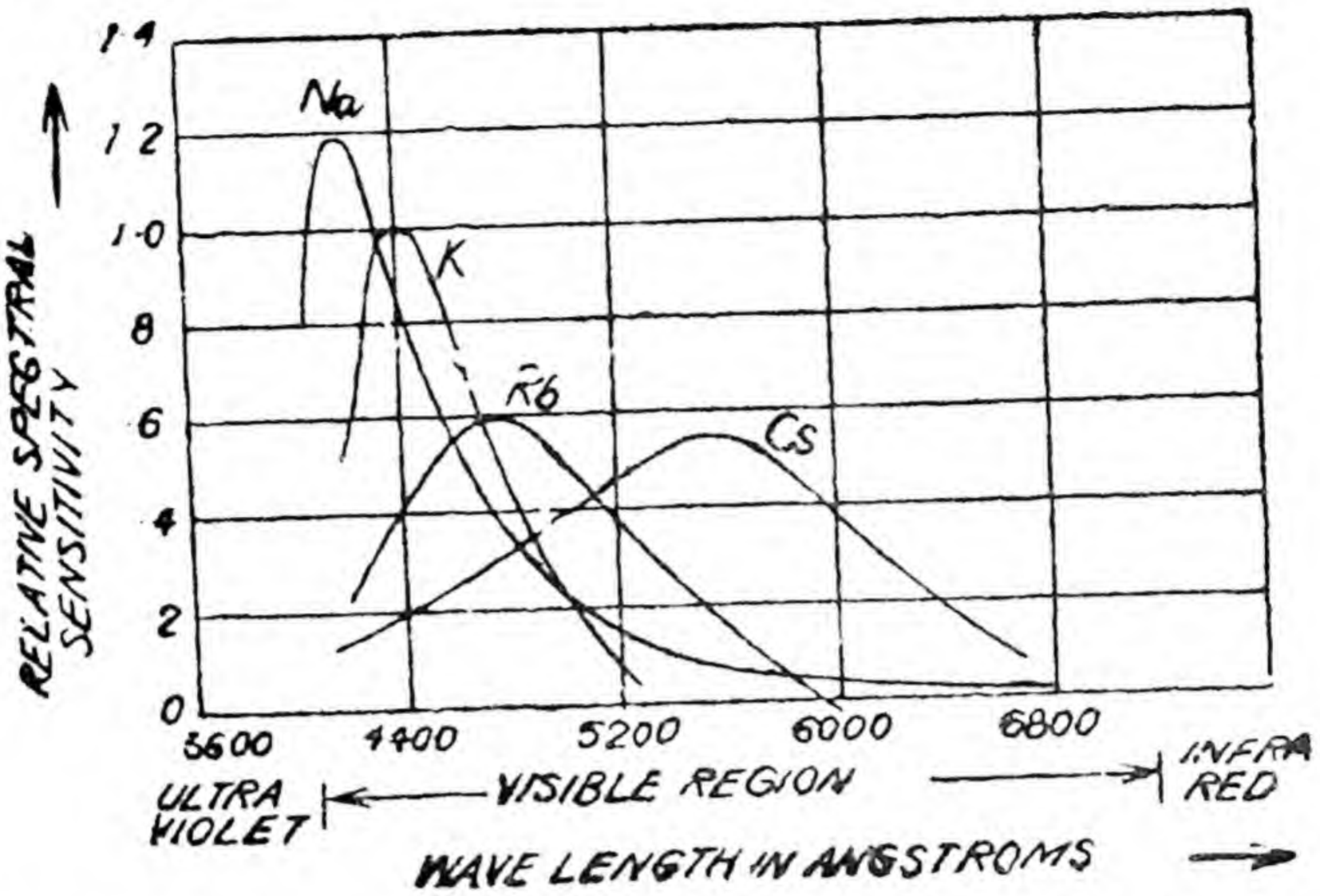


Fig. 54. Relative response of certain alkali metals.

spectral sensitivity (emission current per watt) as a function of wavelengths for photo emissive materials which are generally some alkali metals. From this figure we can select the optical frequency band over which the particular material is photo sensitive.

It is clear from the figure that cesium gives the best emission for visible radiation. The alkali metals are most commonly used for photo-emissive cathode because the frequencies at which emission takes place (threshold frequencies) lie in the visible spectrum. From Einstein equation of photo emission we conclude that lower is the work function of the metal, lower will be the wave length of radiation which may cause photo-electric emission. The work function may be reduced by using composite surfaces and hence they are used for photo emissive cathodes. The preparation of surfaces of these materials involves evaporation of metal in vacuum to produce condensation of thin film on the cathode. Curves for typical surfaces are given in figure 55.

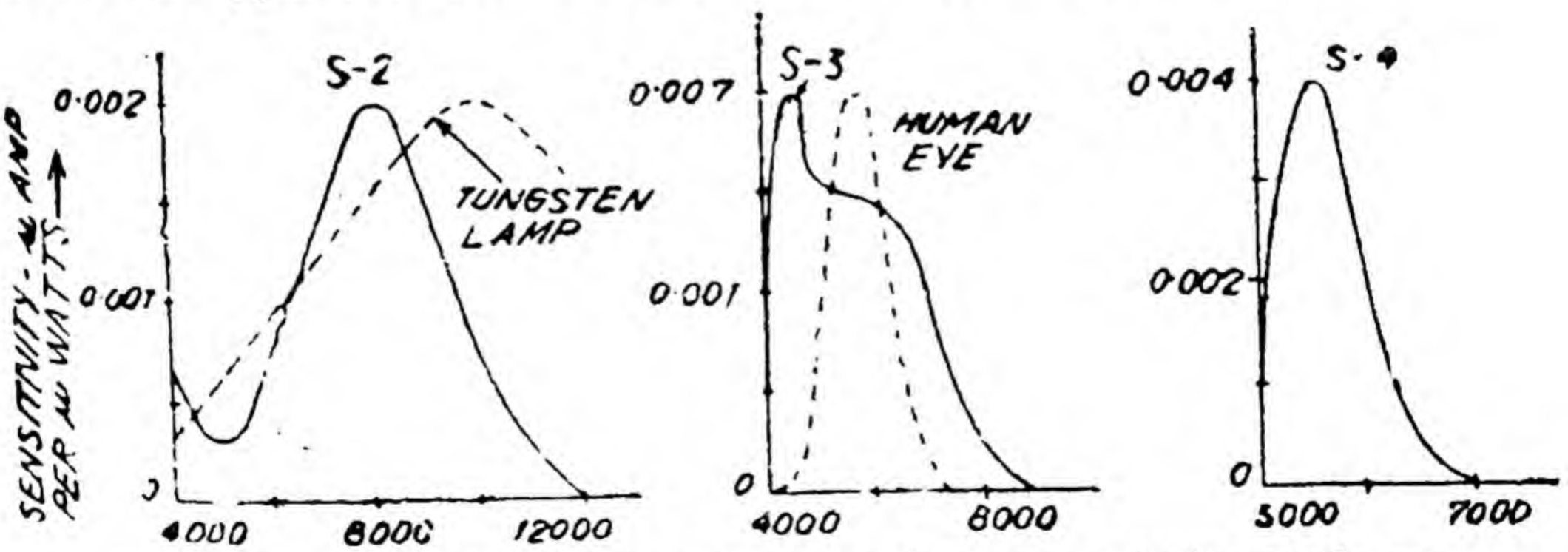


Fig. 55. Colour sensitivity of several photo emissive surfaces.

The surface shown by *S*-2 is cesium—cesium oxide surface which has reasonable sensitivity throughout the visible region. The tungsten lamp operating at 2870°K such that its spectrum of radiant energy is that of a black body, can be used as a standard source for comparison purposes. The comparison with *S*-2 surface shows

that *S*-2 surface is well adapted for use with a tungsten source. The *S*-3 surface employs rubidium which has greater sensitivity in the blue end. *S*-4 surface is of cesium which is highly sensitive and have peak response 20 times greater than the peak of *S*-3 surface.

5·19. PHOTO-EMISSIVE CELLS AND THEIR APPLICATIONS:

Photo-emissive cells may be classified into two categories :

(a) Vacuum photo cells : They are evacuated to a high degree so that the ionisation is almost absent.

(b) Gas photo cells : A certain amount of inert gas at low pressure is introduced in the tube to increase its sensitivity.

Now we shall discuss photo tubes.

5·19-1. PHOTO TUBES :

Photo tube is a two element tube. Cathode is in the form of a half cylinder coated with some photo sensitive material. Different light-sensitive materials for coating can be used to enable the photo tube to respond to the light of different colours or even to invisible radiations. The electrons are emitted when light falls upon the

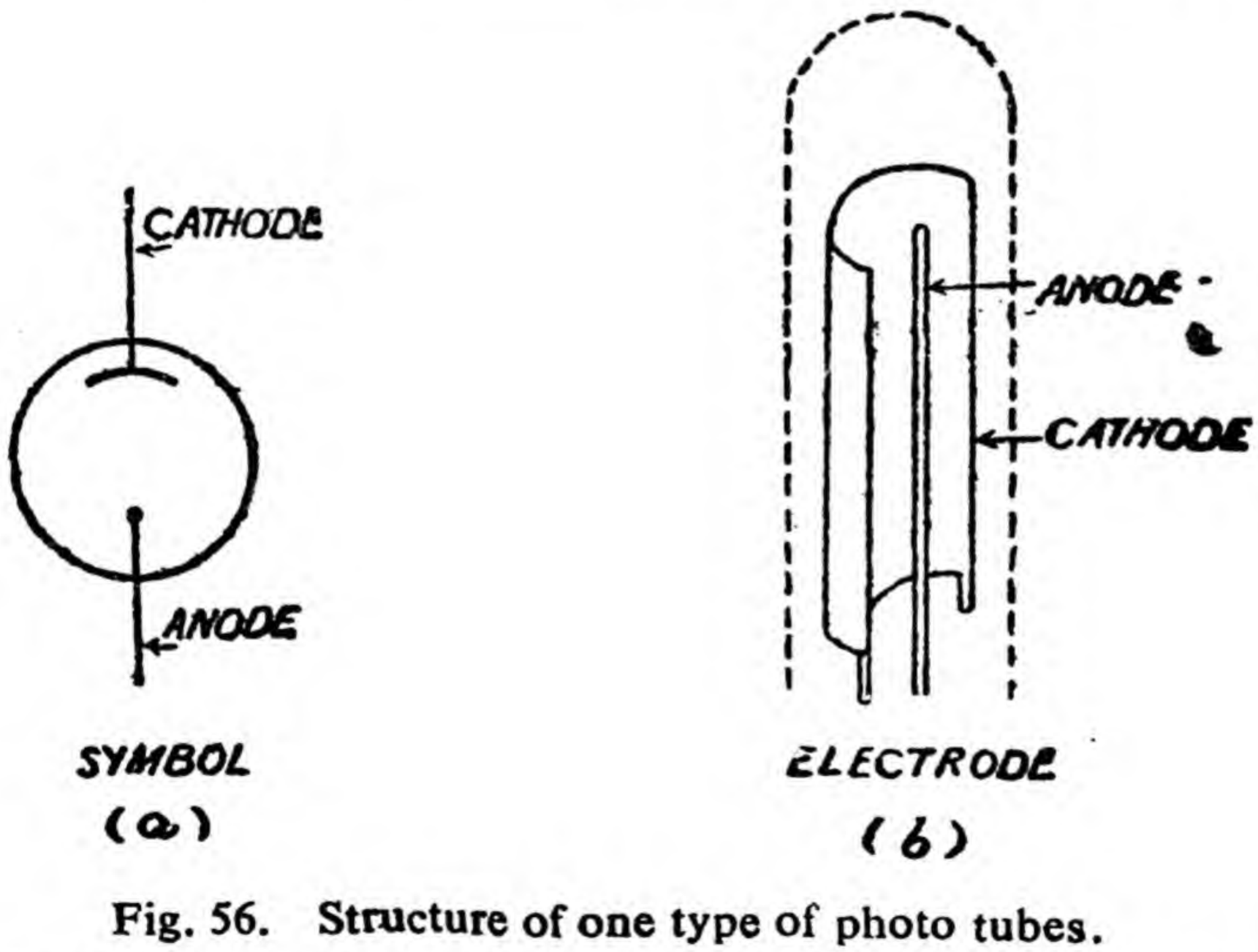


Fig. 56. Structure of one type of photo tubes.

surface of photo-emissive cathode. The emitted electrons are then drawn to positive anode and thus produce a plate current.

Photo tubes are classified as

(*i*) *Vacuum photo tubes.*

(*ii*) *Gas photo tubes.*

Vacuum photo tubes : In such tubes type evacuation is so perfect that their characteristics are quite unaffected by gaseous ionisation but primarily depend upon the nature of the emissive surface and upon the transmission characteristics of the vacuum envelope. Static current voltage characteristics of a typical photo

tube 917 for various levels of illumination† are shown in figure 57. In fig. 58, curve of anode current versus luminuous flux is plotted.

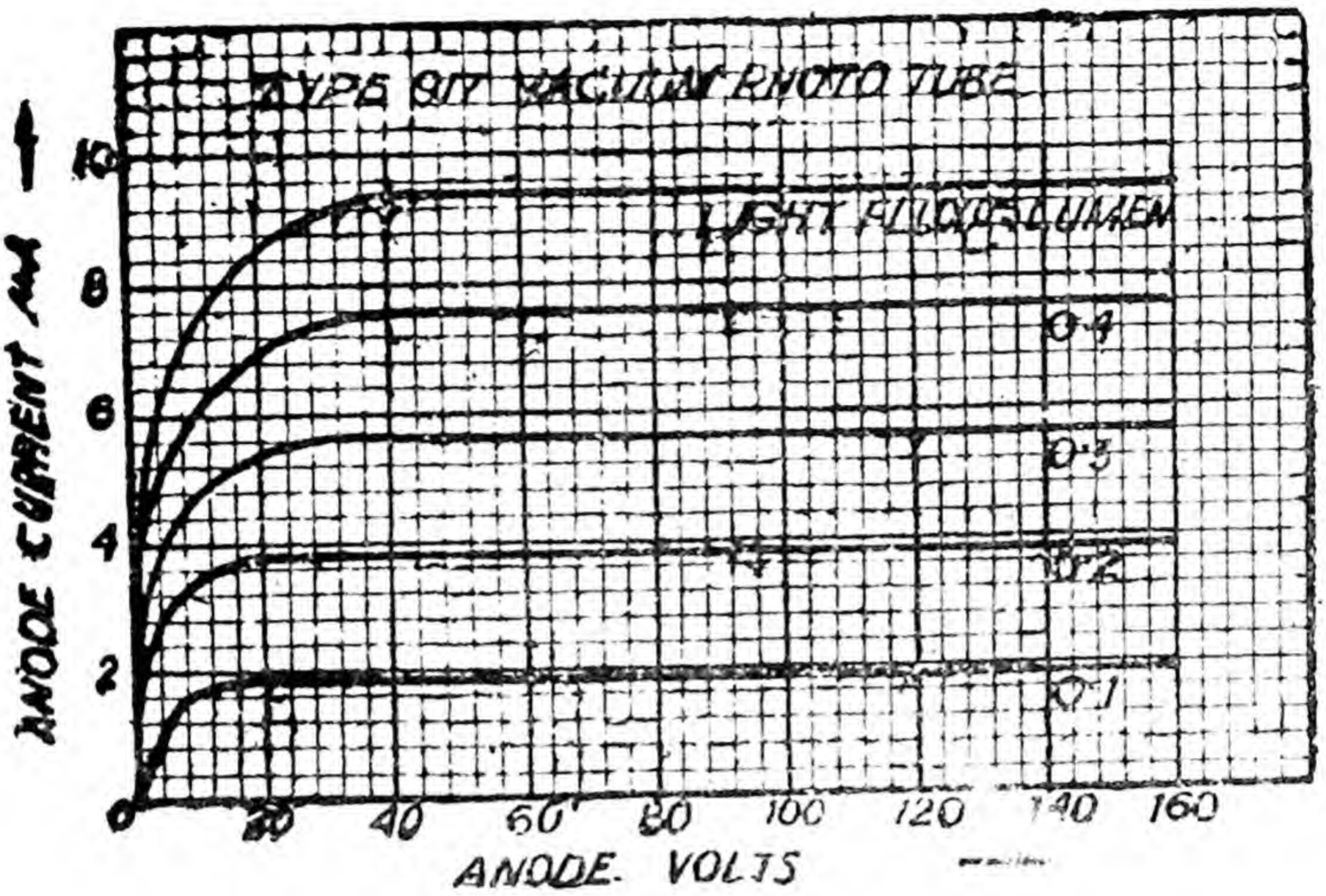


Fig. 57. Static current-voltage characteristics.

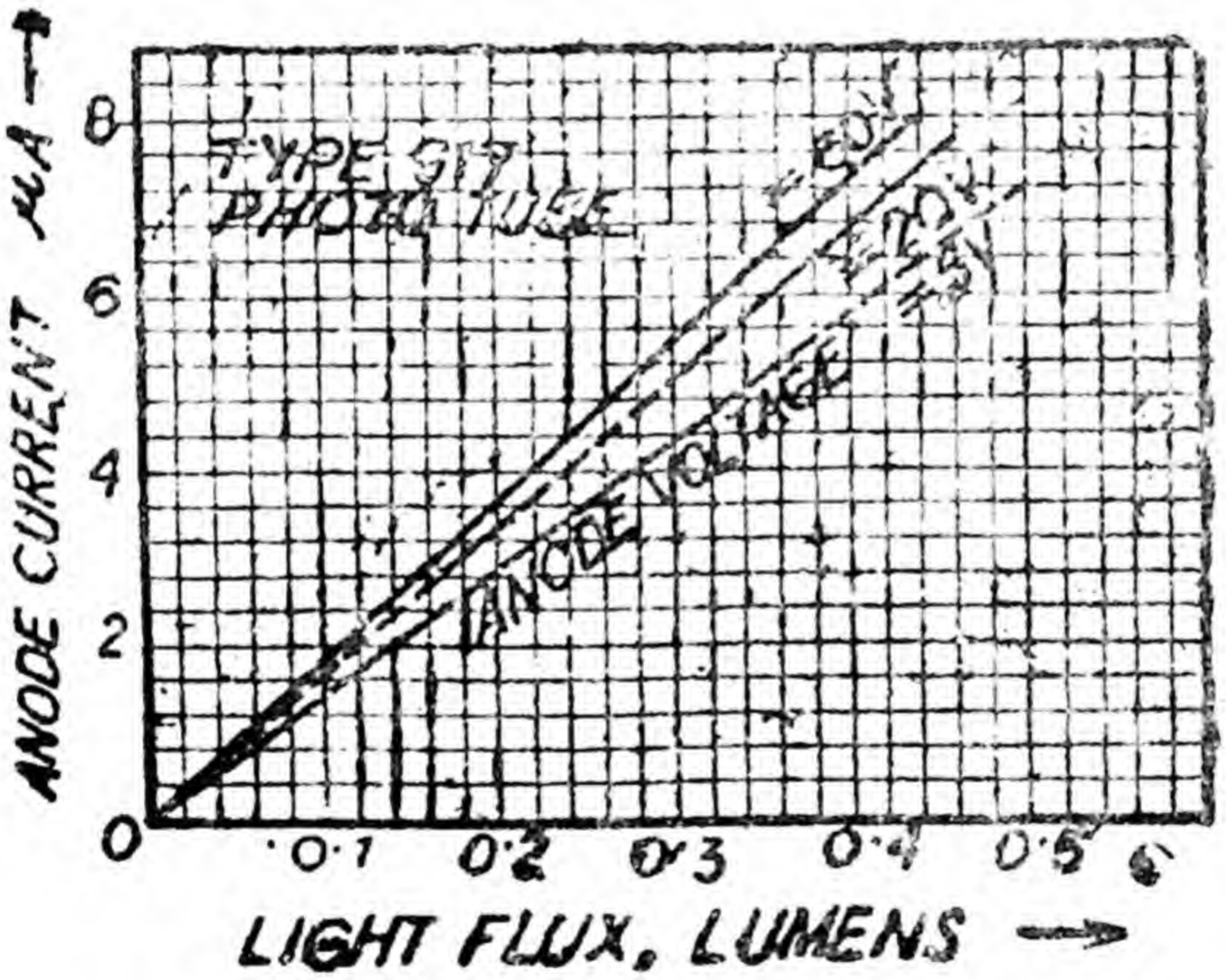


Fig. 58. Curve of anode current vs. luminous flux.

Curves of both figures strictly correspond to the light from a tungsten filament at a colour temperature 2870°*K*. With the radiation of different energy distributions, the form of the curves is the same but current range becomes different.

Solid curve in fig. 58 applies to all anode voltages above saturation. Dotted curves are for anode voltages below saturation.

We see from fig. 57 that current-voltage characteristics are similar to those of a vacuum pentode. *i.e.* like pentode, vacuum photo-cell acts as a constant current generator.

Curves show that, even with the cathodes of high sensitivity, the current obtained from vacuum photo tubes are small.

Gas photo tubes : In such type of tubes, some quantity of gas, which does not react with cathode coating film, is introduced to increase the sensitivity.

When photo electrons travel from cathode to the anode, then, in their way, they strike with some of the gas molecules. The result is the ionisation which splits gas molecule into a free electron and a positive ion. This free electron moves towards the anode and thus it ionises other gas molecules giving rise to move free electrons. The free electrons, thus released, pass to anode, increasing the effective current of photo tube. Positive ions are also drawn to the cathode and hence constitute a current. Therefore at every point between the cathode and anode current is increased, the amount of which depends upon the number of collisions per second between cathode and anode. In addition, photo-tube current is also increased by the secondary emission arising due to the impact of positive ions with cathode. This whole process, by which the effective photo-tube current is increased, is called *gas amplification*.

The gas amplification that can be realised in photo tube depends upon the gas *pressure* and the *voltage* involved. Increase of anode voltage increases the number of primary electrons which make ionising collisions. With sufficiently high anode voltage, the primary electron after its one collision and the newly released free electron due to collision may make one or more collisions in their journey to anode. The current increases very rapidly if the voltage is raised above ionisation potential of the gas.

The gas pressure is also very important. If the gas pressure is too low then the gas molecules will obviously be far spaced and thus mean free path may become so large as to avoid any probability of collision between the primary electron and gas molecules. As the pressure increases, the number of collisions made by an electron in moving from cathode to the anode increases ; but if the pressure is too high then an electron may stop without being sufficiently accelerated so as to gain energy to cause ionisation.

---

† For the interest of the readers few definitions are listed below :

1. Luminous flux is the time rate of flow of light or radiant flux of particular energy distribution, lying entirely in the visible range.

2. Lumen is the unit of luminous flux. It is equal to the flux through a unit solid angle from a uniform point source of one candle.

3. Illumination is the density of the luminous flux on surface.

4. Luminous intensity is the luminous flux on a small surface normal to the direction divided by the solid angle (in steradians) which the surface subtends at the source of light. The unit of luminous intensity is candle.

†When alkali metals are used in the cathode, the insert gases helium neon, argon, krypton, and xenon are filled in the envelope since they do not react with alkali metals.

Thus an optimum pressure, which is found to be 0·2 mm. of Hg, is

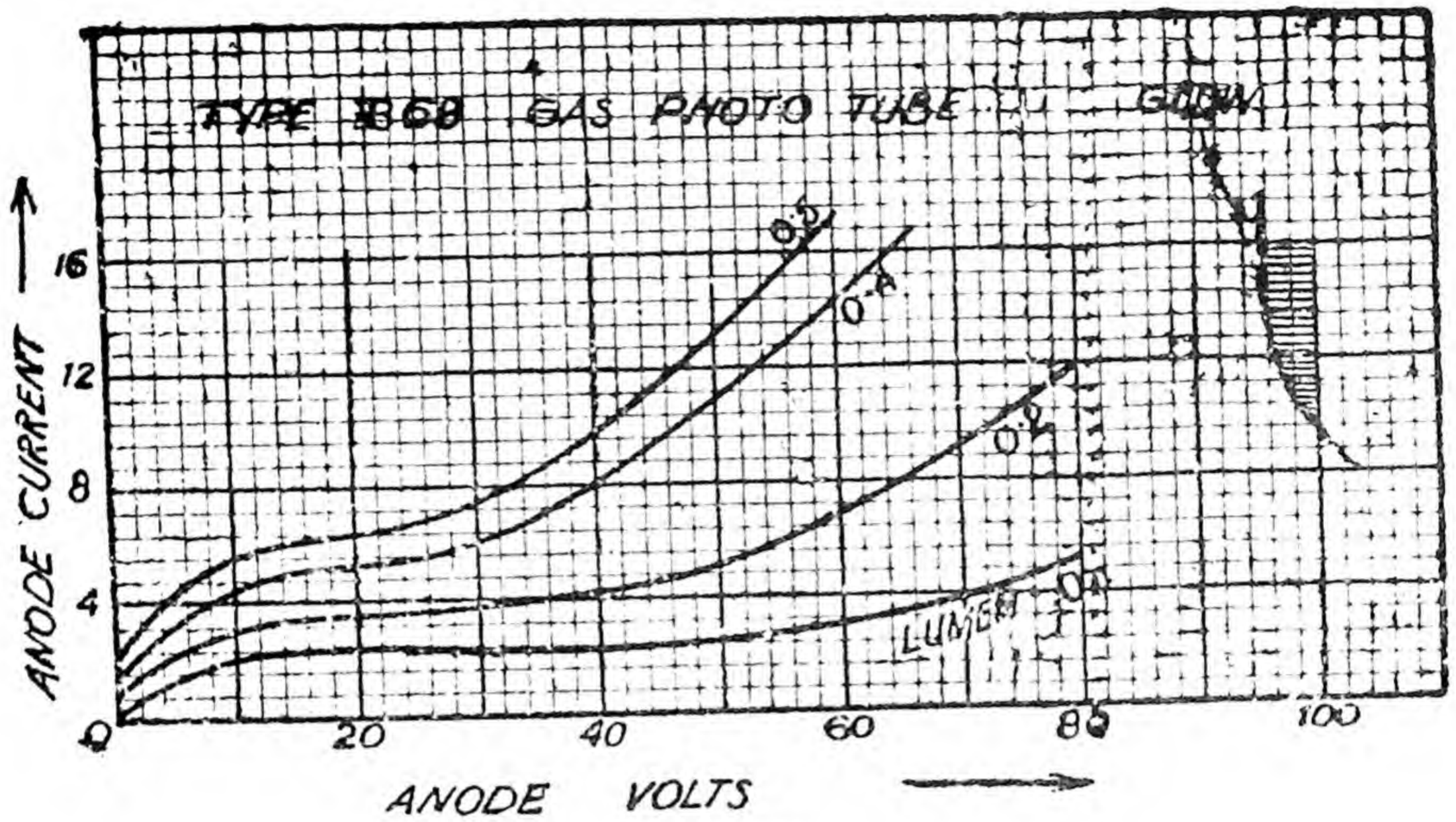


Fig. 59.

commonly used in the gas photo tubes. This corresponds to a pressure of 268 dynes/cm.^{2} and a mean free path of 0·23 cm. The ratio of the sensitivity of the gas photo-tube at a given voltage and flux to the sensitivity obtained at the same voltage and flux without ionisation resulting from the presence of gas is termed as gas *amplification factor*.

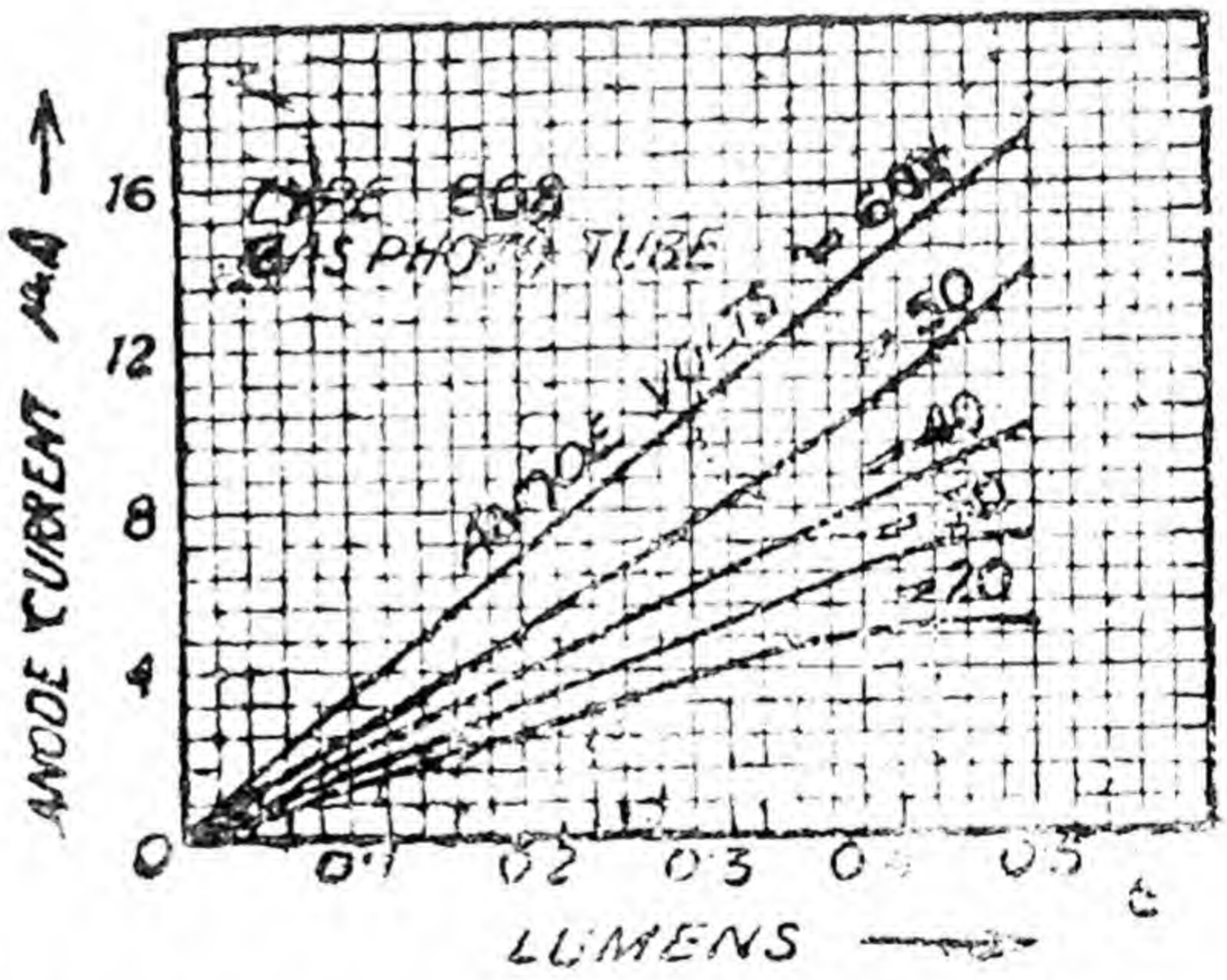


Fig. 60. Gas photo tube characteristics.

In figure 59 anode current has been plotted against anode voltage and the figure 60 curve of anode current versus luminous flux is shown. Both curves correspond to the type 868 gas photo-tube.

The form of current voltage curves of figure 59 below 15 volts for gas photo tube quite resembles with the form of the same curves for vacuum photo tube [fig. 57] which in fact should be the case since a very little ionisation takes place below 15 volts. When the voltage exceeds 90, a glow discharge may take place. The dotted curve in fig. 59, shows that how the breakdown voltage falls with the increase of flux.

Disadvantages of using gas : There are three main disadvantages which affects the gain in sensitivity :

(*i*) Anode current is no more a linear function of flux.

(*ii*) For voltage more than 90, glow may take place which is objectionable due to two reasons. Firstly, it causes the current to be independent of illumination and secondly, it reduces permanently the sensitivity of the tube. Therefore anode voltage must be kept below 90.

(*iii*) Because of the large mass of the positive ions the time taken for a positive ion to pass from the vicinity of anode to the cathode is appreciable. When the light source is cut off, the current does not drop to zero immediately for there are still positive ions floating about between electrodes, *i.e.* there is a preceptible time lag in the response to a change in flux. In vacuum photo-tube photo-electric current ceases in less than 10μ sec. after the illumination is cut off and hence no time lag.

Phototubes are extensively used for counting and sorting devices. They may be used in image transmission system for the conversion of light intensity variations into electrical variations.

5·19–2. PHOTO ELECTRIC RELAY :

The vacuum or gas filled photo emissive cell has high internal resistance and provides a high voltage and low current output. Due to this property, it can be connected in the grid circuit of an amplifier. Fig. 61 shows a simple vacuum tube amplifier where the photo-emi sive cell is connected in the grid. When no light is incident on the photo-emissive tube, a potentiometer in the grid may be adjusted in such a way that no plate current flows. In the figure, i_b determines the value of light required to cause the relay. When light strikes the photo emissive cell, conduction through the photo cell takes place which causes a positive grid signal. As a result, the plate current increases and closes the relay.

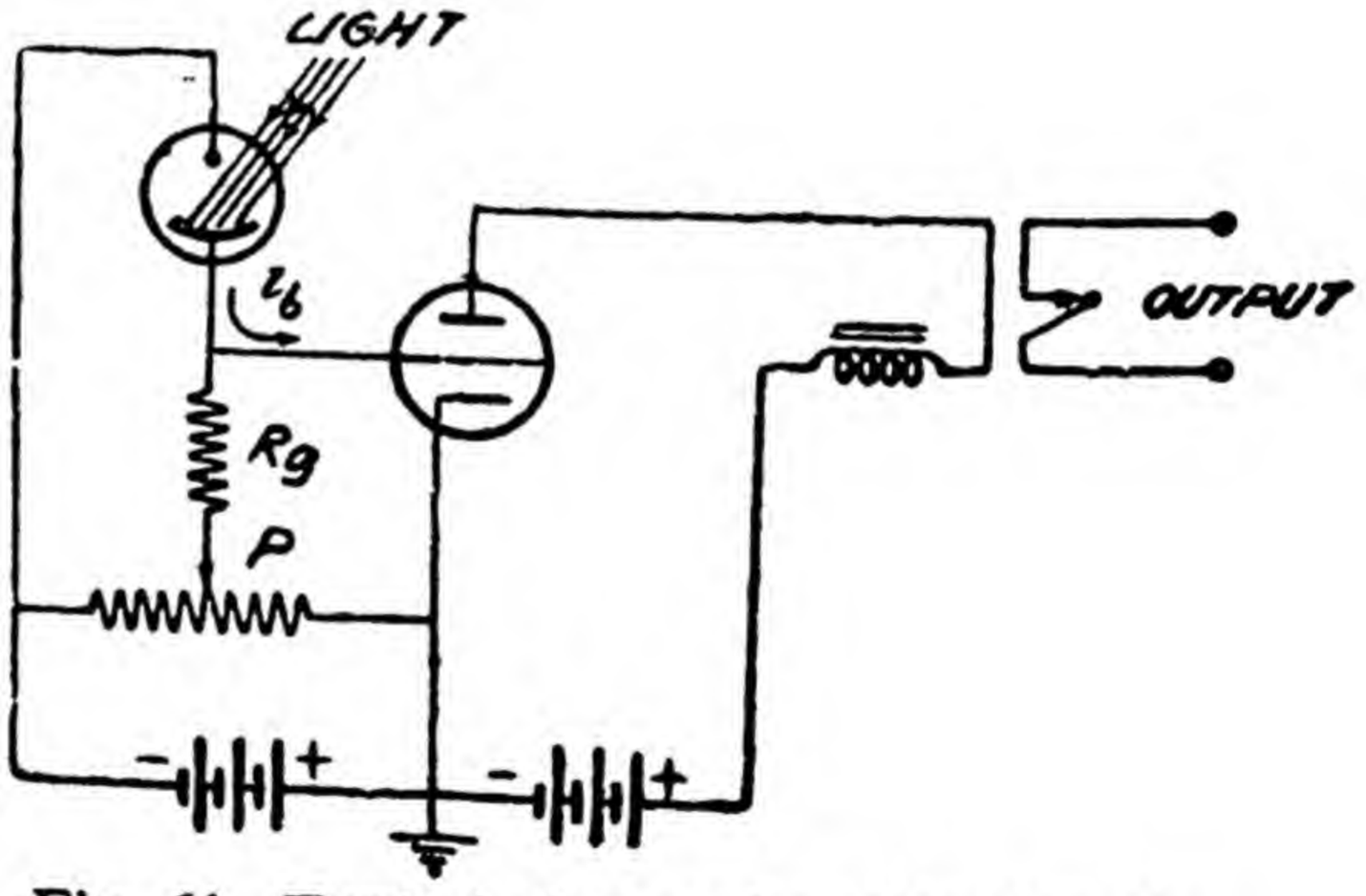


Fig. 61. Typical photo tube relay circuit.

The circuit may also be used as relay drops by changing the polarity of photo-emissive cell and applying initially such a voltage to the grid that relay takes place. In this case, when light strikes, it causes a negative grid signal and the plate current drops the relay.

5·19-3. PHOTOMULTIPLIERS (ELECTRON MULTIPLIERS) :

The drawback of the photo tubes in a small plate current which

is about $10\mu A$. In many applications, like reproductions of sound recorded on the film, in television etc., the output current of a phototube should be amplified which can be done by using the few stages of amplifier The current may be amplified by the device called electron multipliers which is based on the principle of secondary emission. The multiplier tube has an average current of about 1 mA and can provide a gain or amplification of over 20,000 or more.

A photomultiplier tube consists of a light sensitive photocathode K, a system of nine or ten secondary emission electrodes, called dynodes D and a collector anode A, as shown in fig. 62.

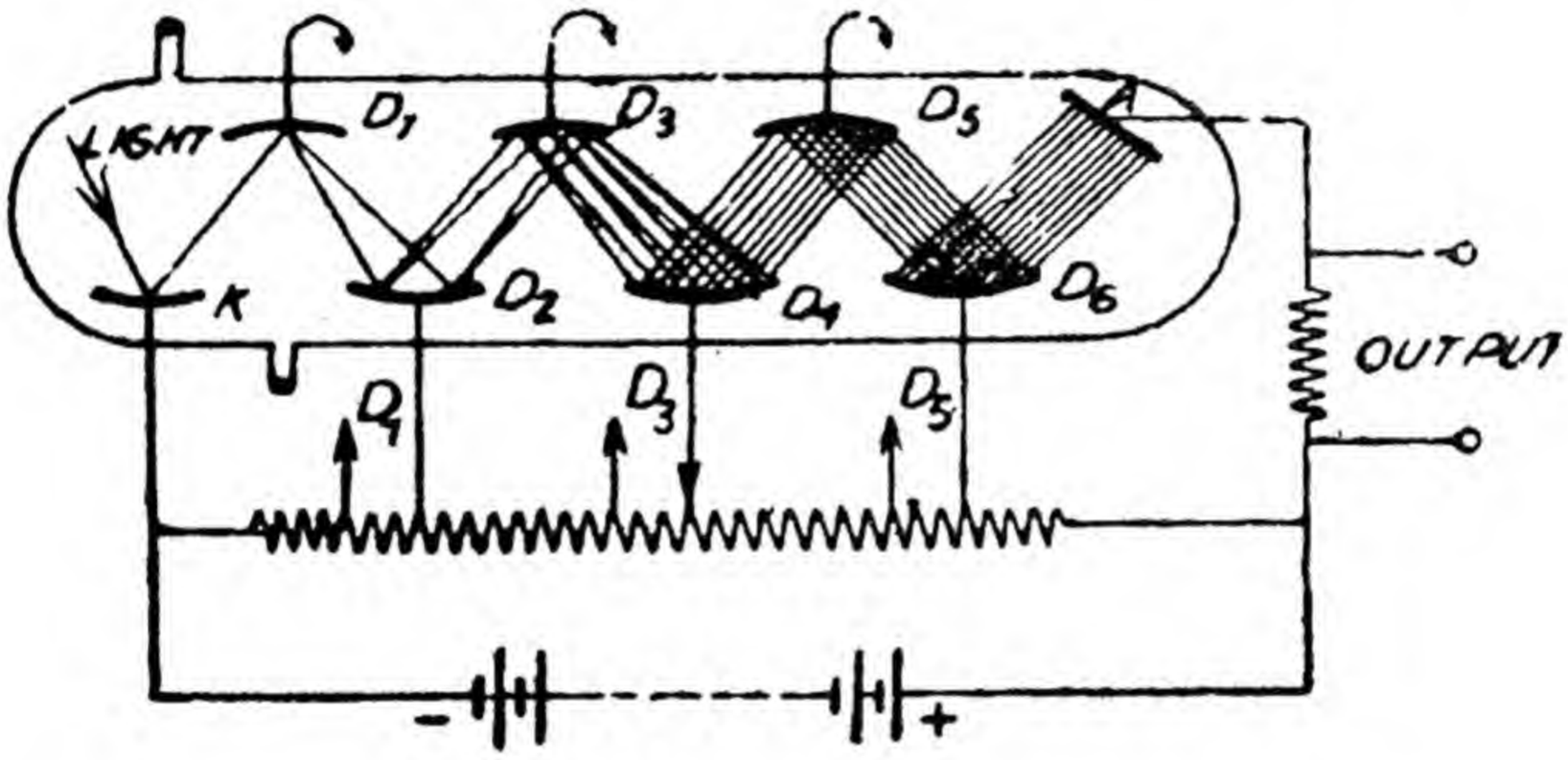


Fig. 62. Photo electric secondary emission multipliers.

Each dynode is having successively higher positive voltage and is coated with a material (like silver magnesium) which have a secondary emission ratio N from three to five.

When light falls on photo cathode, electrons are liberated and are attracted towards the first positive dynode D_1. Secondary emission ratio, N electrons are emitted which are accelerated towards dynode D_2. D_2 is similar to D_1 and the potential difference between D_1 and D_2 is the same as the potential difference between D_1 and K. Thus the secondary emission ratio at D_2 is again N *i e.* N^2 electrons are emitted from D_2. The process is repeated for other dynodes *i.e.* if n be the number of dynodes then one electron at the input will give N^n electrons at the output. The output current is collected by the collector anode A. The current is thus amplified N^n times

Many electrons pass straight down the tube and thus reduce the gain. In the developed form the electrons are focussed successively on one target after another, thus preventing spreading of the beam. The dynodes and the collector anode are arranged along a circle rather than along a straignt line for economy of space.

A typical photomultiplier circuit is shown in figure 63. The use of the individual bleeder resistors between each dynode ensures a constant voltage ratio between them and the anode at all times. This is essential for maximum gain and stable performance. The high voltage can be obtained with power supply. Signal output is taken as shown in the figure 63.

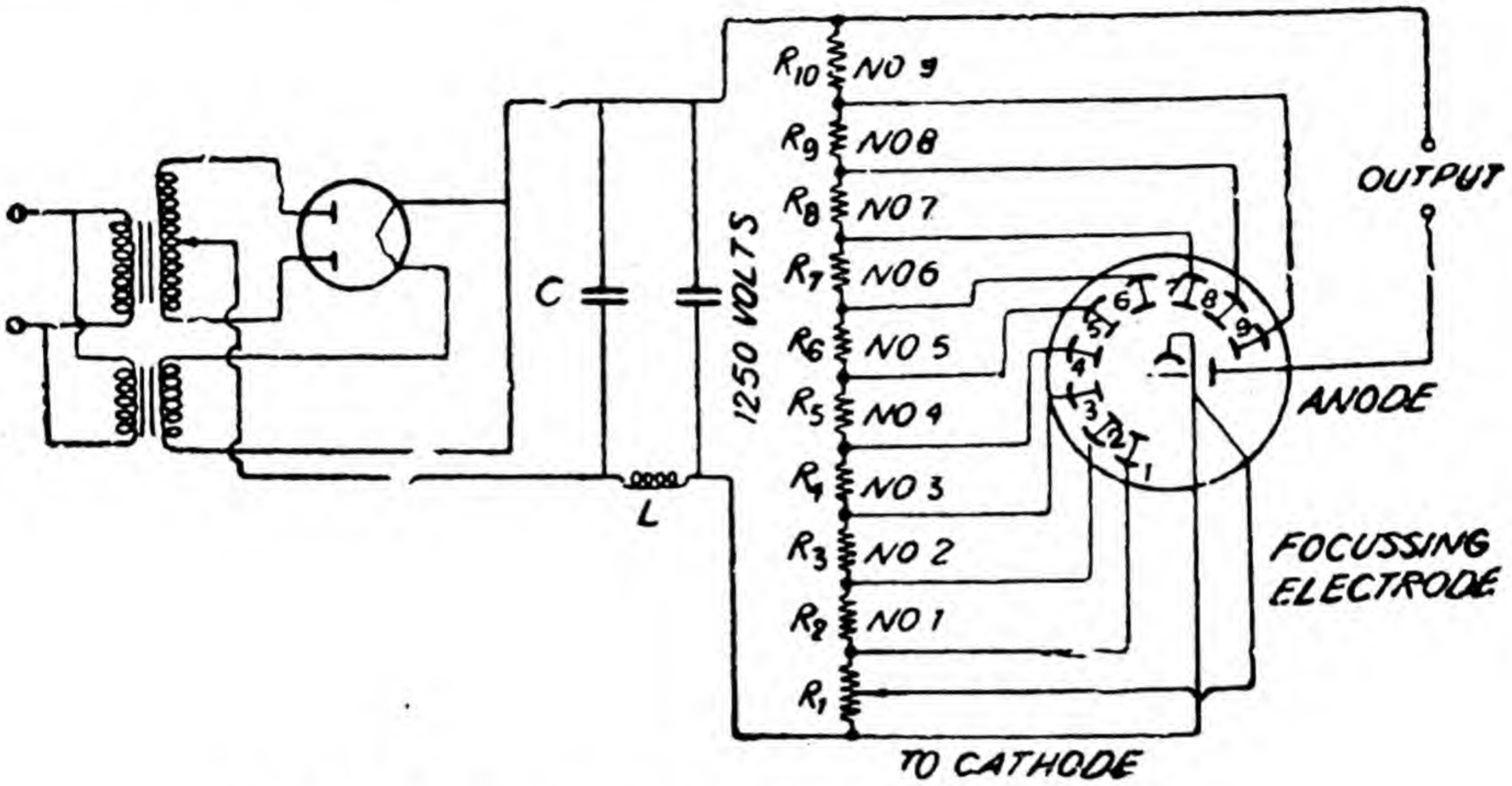


Fig. 63. Photomultiplier with full wave rectifier.

5·20. PHOTO CONDUCTIVE DEVICES (PHOTO DIODES AND PHOTO TRANSISTORS) :

Photoconductive devices are thin layers of semiconductors such as selenium, silicon, cadmium sulphide, lead sulphide etc, whose resistance decreases with light, mounted in a sealed glass. Modern units employ germanium or silicon in the form similar to point contact or junction diodes or transistor as shown in figure 64.

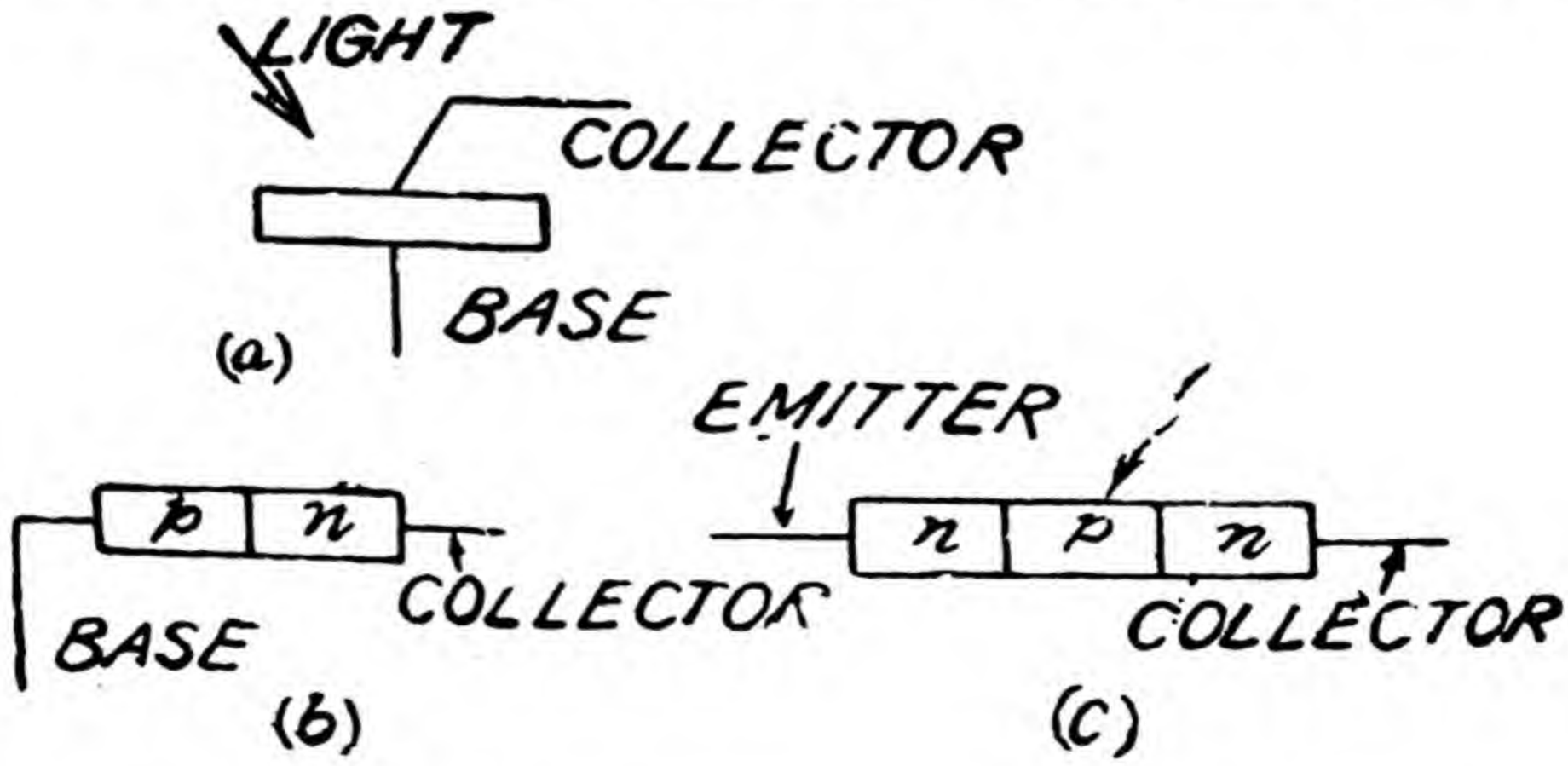


Fig. 64. (*a*) Point contact. (*b*) P-N Junction.

To understand the mechanism of photo current, let the light energy be allowed to incident upon the *P-N* or *NPN* junction. The

energy is absorbed in raising the electron from the top of the filled band into the conductivity band. Due to the escape of the electron from filled band a hole is left which may be swept out by the applied potential. Thus hole and electron are able to contribute to photo current. Fig 65 shows the simple circuit arrangement with a voltage *E* impressed across the photoconductive device and a resistor *R* in series. The photoconductive device is a junction diode with positive voltage applied to collector and the negative voltage to base side. With no light, the photo-diode has a high resistance which limits the current flow. When light falls on this device, the resistance decreases and the series current increases. Current flow through *R* can be amplified for control. It is also possible to replace *R* by a sensitive relay which closes when the current flow increases and thus controls a larger amount of power through relay contacts.

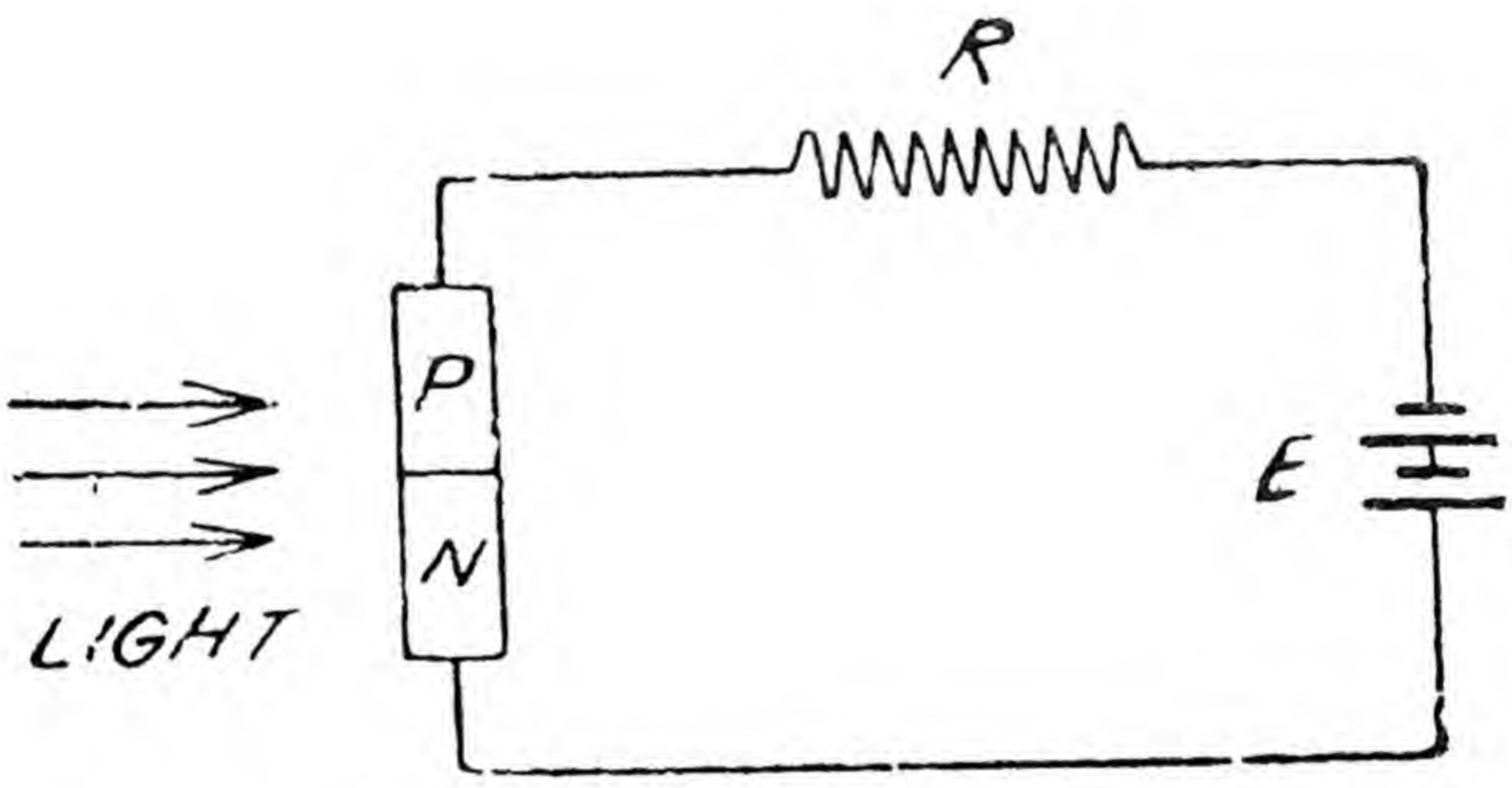


Fig. 65. Photoconductive circuit.

Under perfectly dark conditions, there is no current in photo tubes while an appreciable current exists in photo-diodes or photo transistors. This is expected because some electrons from the impurity atoms would exist in the conduction band at normal temperatures and in darkness. These have been raised to conduction band from donor impurities by thermal energies, and produce conductivity in the dark. As long as the current is small, it will not interfere with the photo current. Fig. 66 shows the characteristics of a photoconducting cell as a function of voltage.

From figure, it can be observed that the sensitivity is proportional to voltage applied across the device but the voltage is limited so that no overheating and subsequent break down takes place.

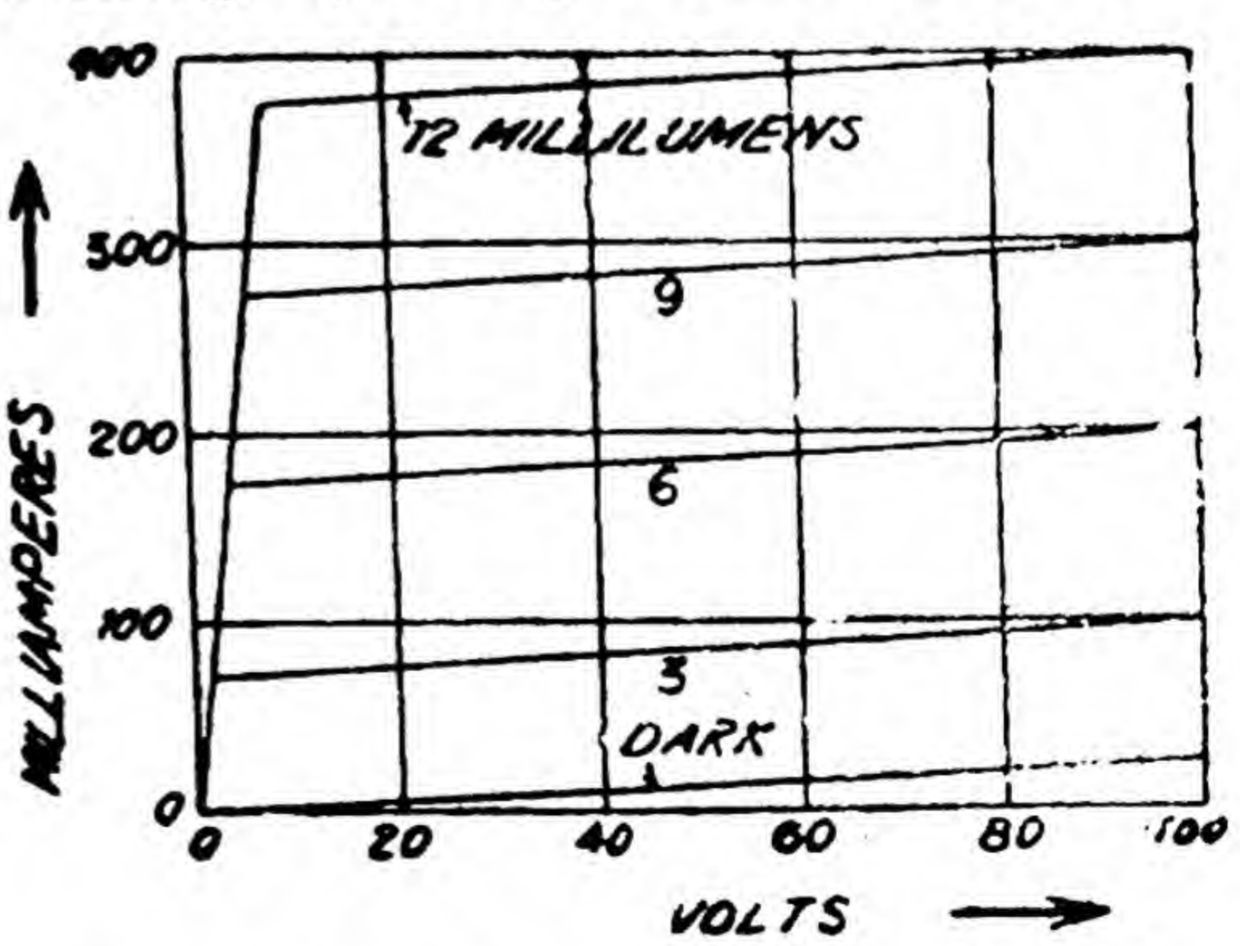


Fig. 66. Approximate characteristics.

Usually the current does not exceed a few milliamperes but it may be increased by suitable design.

Germanium and other semiconductors give a response farther into infrared than is obtainable from photoemissive surfaces.

5·21. PHOTOVOLTAIC CELLS :

The photovoltaic cell generates an *emf* under the illumination of a barrier layer. This is also-called as 'barrier layer cell." The cell consists of layer of semiconductor on a metal base plate. Usually the copper oxide is deposited on the base plate of copper, or iron selenide is deposited on the iorn base. The circuit connections are shown in figure 67. When the barrier layer is illiuminated

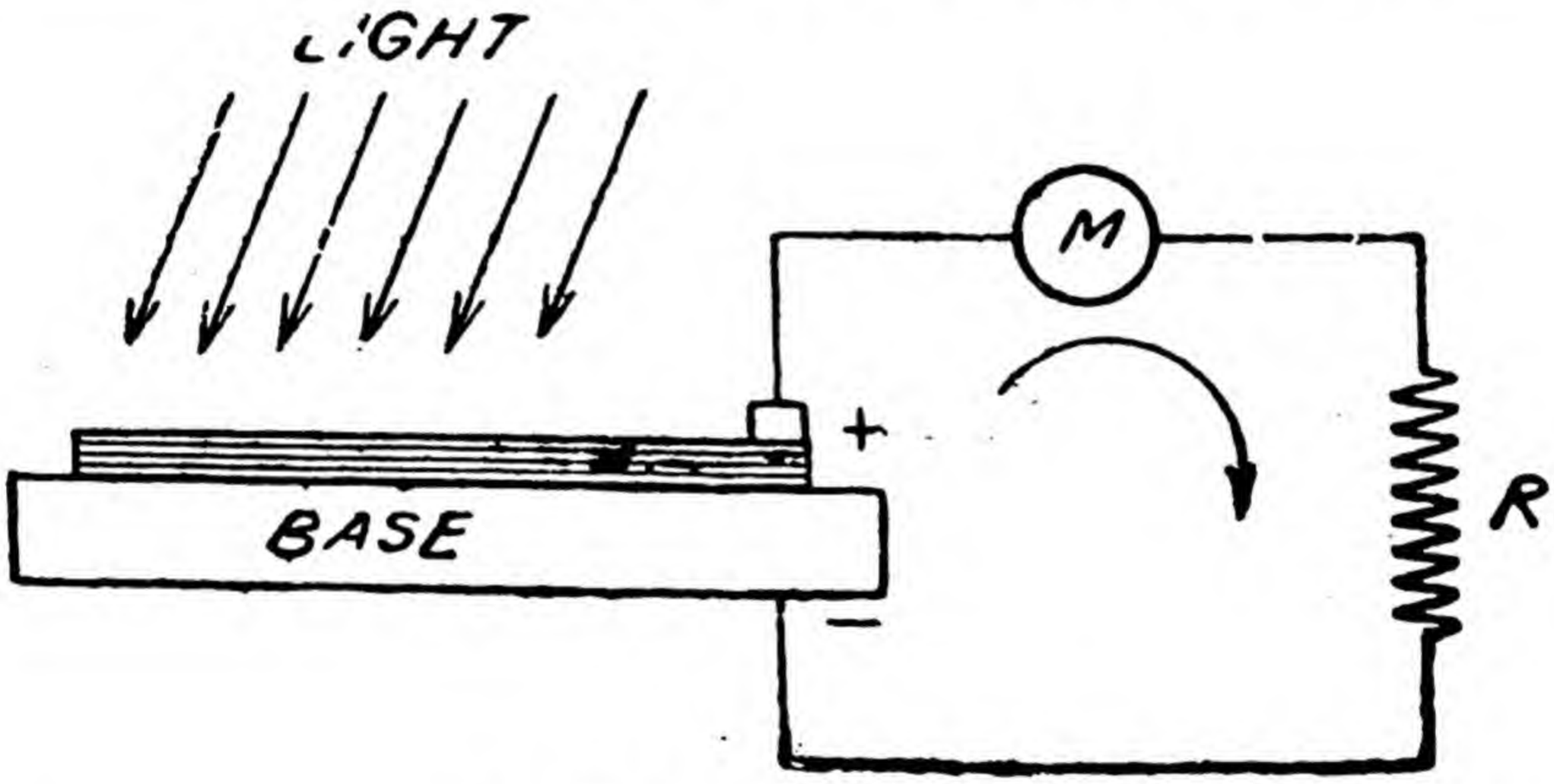


Fig. 67. Photovoltaic cell.

by light energy, current flows in the circuit which can be explained as follows : when light passes through the semiconductors, electrons may be raised to conduction band because they are given enough energy to pass across the forbidden energy gap. Thus electrons flow from semiconductor to metal and conventional current flows in the opposite direction. The reverse flow of electrons from metal to semiconductor is prevented by the surface barrier set up in the semiconductor. As long as the illumination continues, electrons continue to flow in the conduction band and thus resulting in a continuous current.

Figure 68 shows the output current as a function of illiumination in footcandles. These characteristics are in an iron selenide cell. The cell resistance is small and in sunlight currents of a few milliamperes may flow. It is clear from the graph that for external resistance equal to zero. the circuit current is directly proportinal to the incident light flux, but when the resistance is introduced in the circuit the linear result does not hold good. It is found that both *emf* and internal resistance of cell are functions of temperature and hence short circuit current is almost independent of temperature changes. Cells of this nature are sensitive to high

temperature, for example the selenide cell is damaged if it operates above 55°C.

The voltage generated in photo voltaic cell is d. c. and hence the amplification by vacuum tube is not suitable except by chopper amplifier The output current may be increased by connecting the several photovoltaic cells in parallel, because the total area and the total received light flux increases. They may also be connected in series for working into high resistance loads.

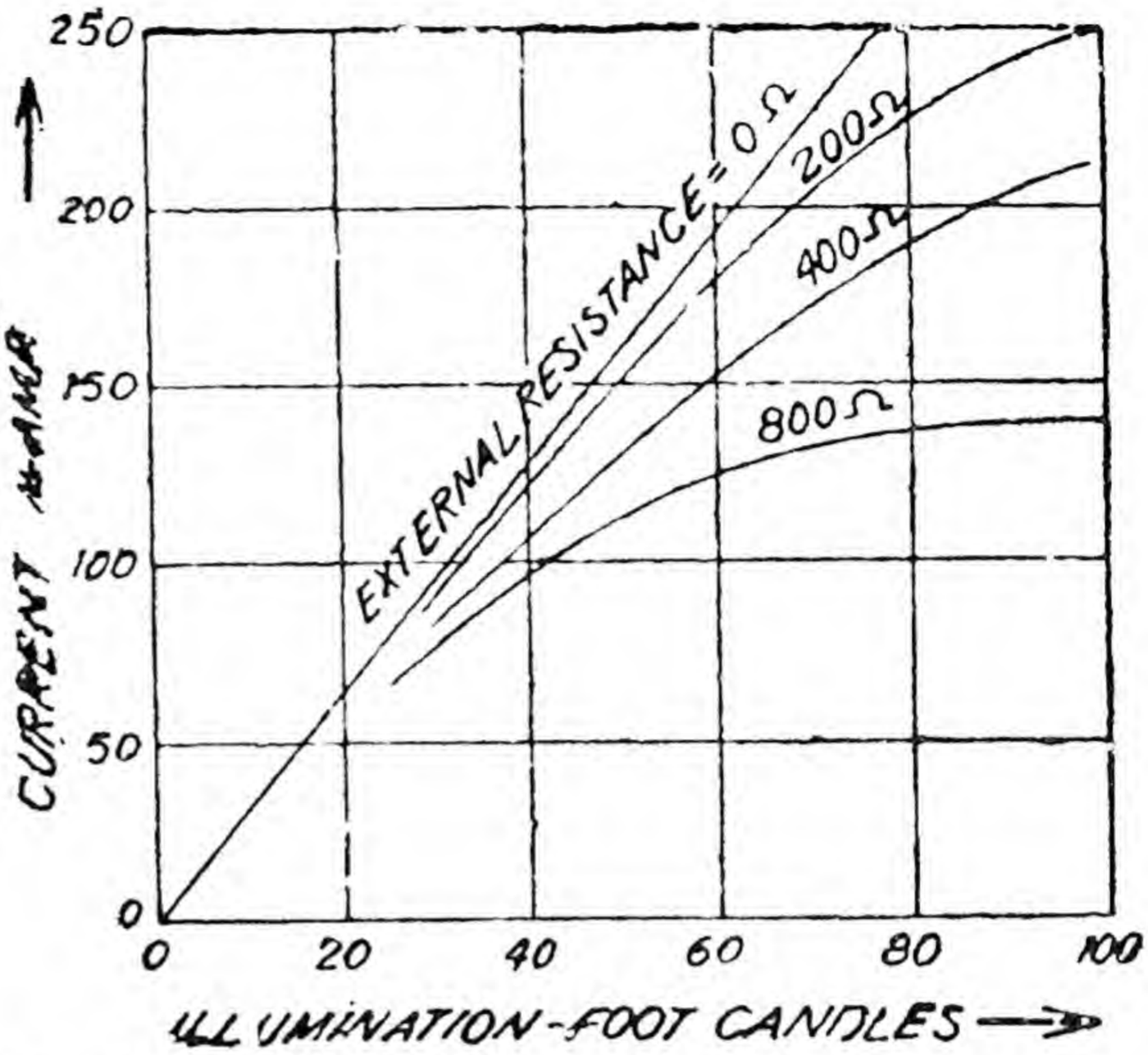


Fig. 68. Characteristics of iron selenide cell.

The advantage of photovoltaic cells over the photoconductive cells and photoemissive cells is that no external voltage source is necessary. But disadvantage is that the output voltage is small compared even with a vacuum photo tube.

EXPERIMENTS

Expt. No. 1 : To measure the inter-electrode capacitance of triode.

In system involving more than two electrodes, it is possible for a capacitance to exist between each pair of terminals. Triode is a three-electrode system. The three capacitances existing between the different electrodes are shown in fig. 69, C_{gp} is direct capacitance between plate and grid, C_{pk} is direct capacitance between plate and cathode, C_{gk} is direct capacitance between grid and cathode.

Inter-electrode capacitance plays an important role at high frequencies, C_{pk} is not very much effective since it shunts the internal as well as external plate circuit but C_{gp} is most effective because, by means of it, any action in the external plate circuit, can be fed back into the grid circuit thus makingt he tube, to a certain extent, a reversible four pole, the grid cathode terminals being considered as two input poles and plate cathode termlnals as the two output poles.

Procedure : The direct capacitance between a pair of electrodes in a cold cathode condition can be measured by connecting the electrode in the unknown arm of a

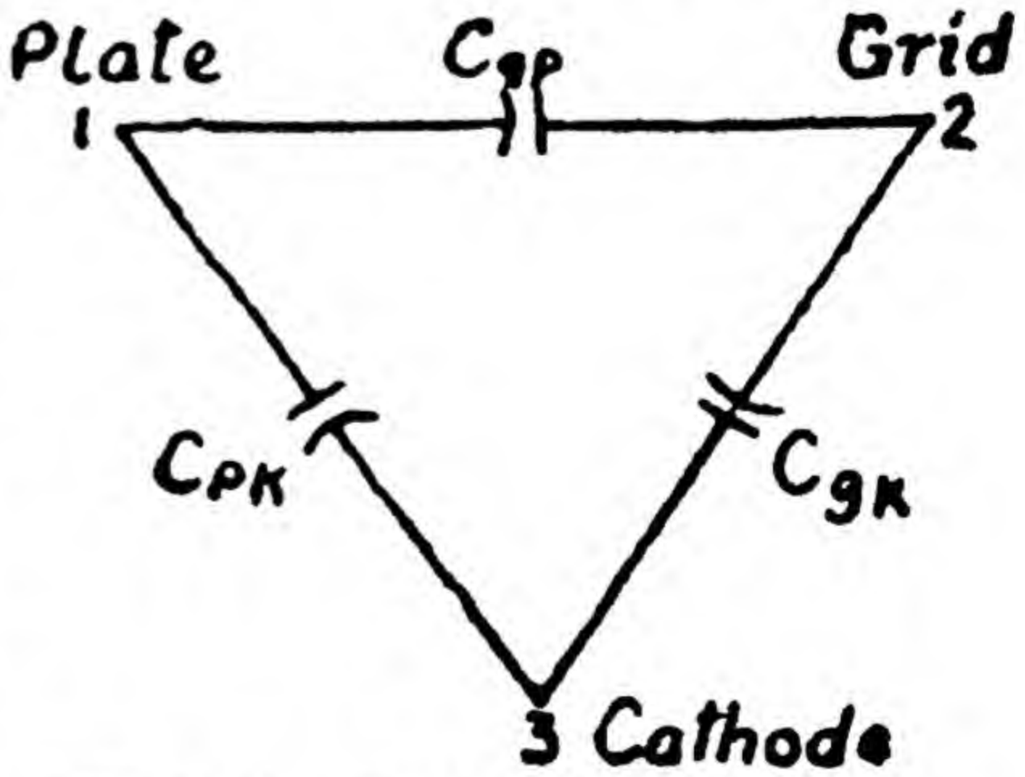


Fig. 69. Inter-electrode capacitances of a triode valve.

capacitance bridge in the usual manner as for measuring an unknown capacitance. The third terminal is connected to one of the two remaining corners of the bridge. The circuit arrangement is shown in figure 70.

As shown in the circuit, the capacitance to be measured is C_{gp}, and the third terminal is returned to corner B of the bridge. This causes capacitance C_{gk} to be shunted across the oscillator where it cannot affect the bridge behaviour At the same time, capacitance C_{pk} is shunted across the resistance R_2. This requires the additional compensating capacitance C_1, such that $C_1/C_{pk}=R_2/R_1$ in order to achieve the balance and maintain the bridge ratio at the value R_2/R_1 defined by the resistance arms. Thus both C_{pk} and C_{gk} being removed from the consideration, the direct capacitance C_{gp} can be measured as an ordinary capacitance by substitution method, without the necessity of determining the associated capacitances C_{gk} and C_{pk}.

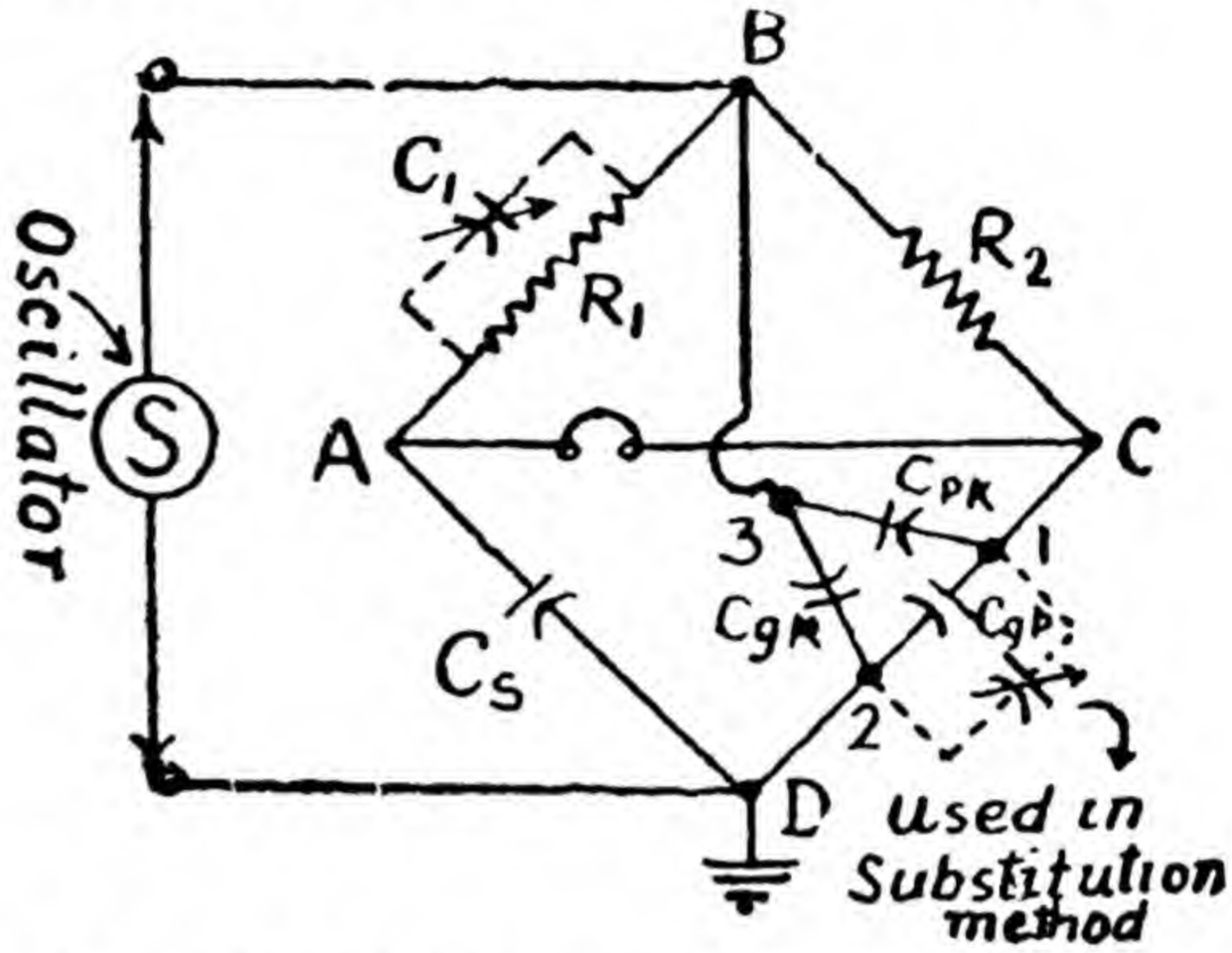


Fig. 70. Direct-capacitance measurement by capacitance bridge.

Exp No. 2 : To measure the dynamic constants of a triode valve.

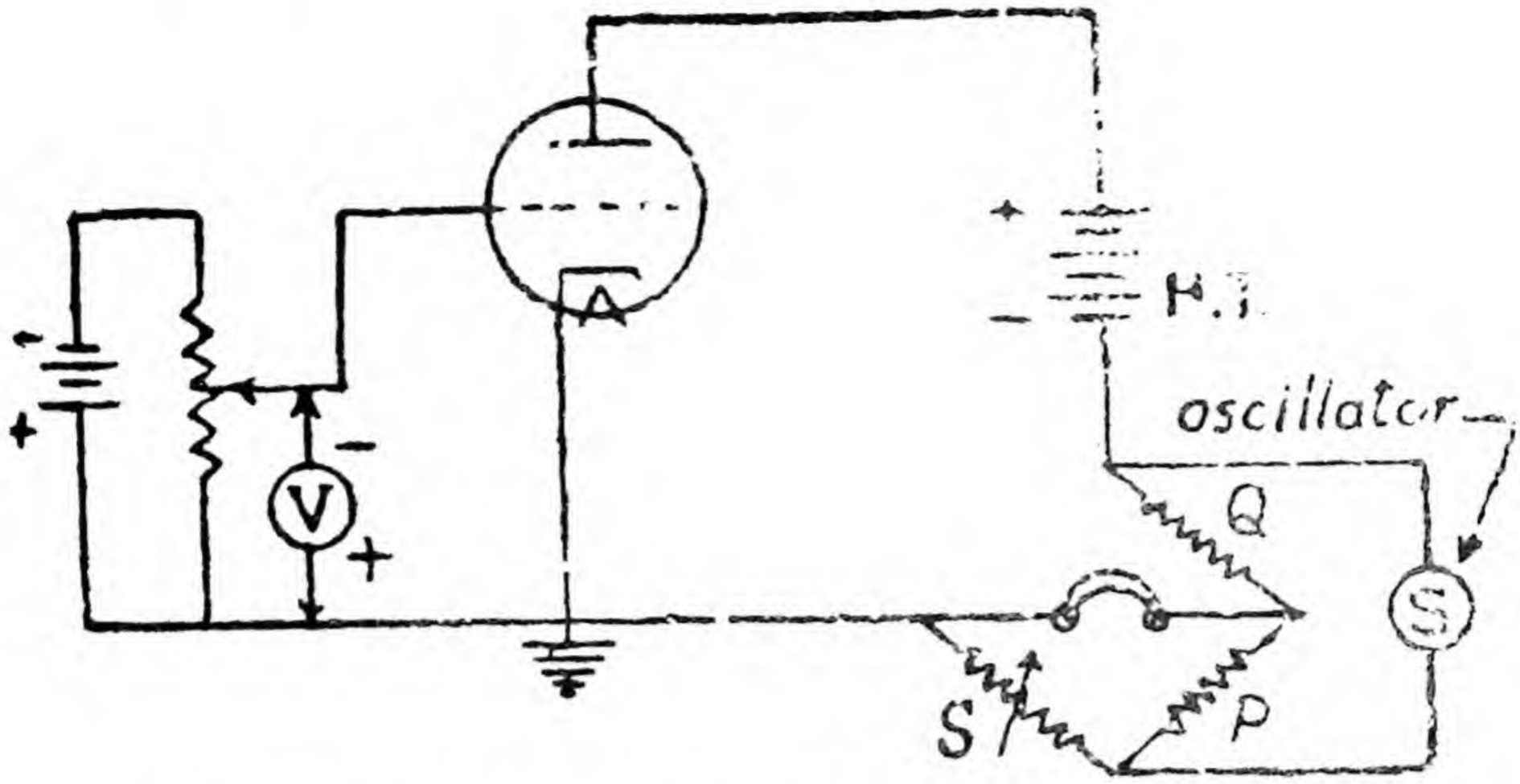
Procedure : Simple circuit arrangement for the determination of the three constants, plate resistance, amplification factor, and transconductance are shown in figure 71.

Triode is connected in an a.c. bridge circuit, P, Q, S are known resistances.

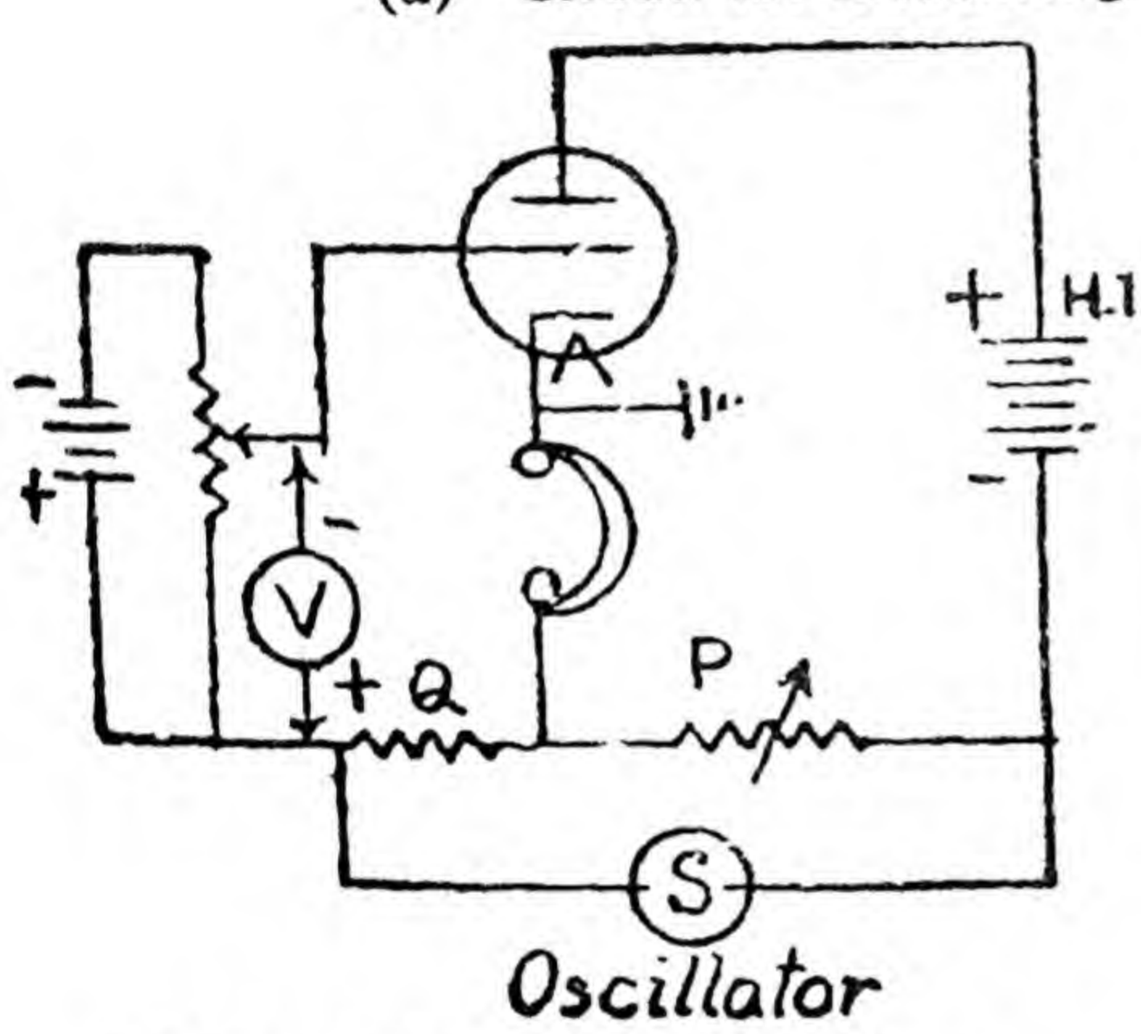
(i) **Plate resistance r_p :** Circuit is shown in fig. 71 (*a*). The plate is kept at about 200 volts. Starting with zero potential on the grid, the bridge is balanced and the values of P, Q, S are noted ; then the dynamic plate resistance is calculated from the relation

$$r_p=\frac{Q}{P}S.$$

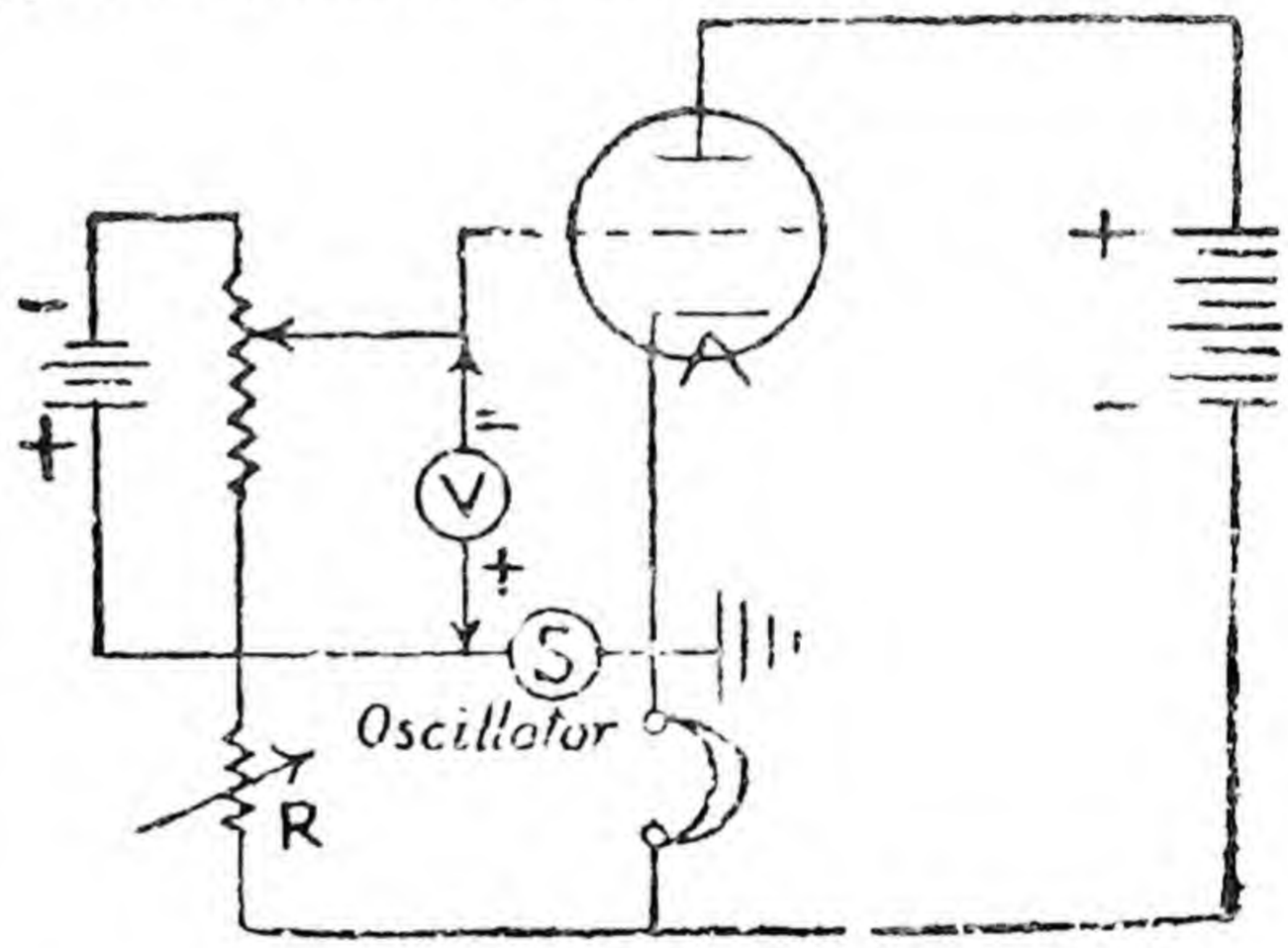
For perfect balance it may be necessary to connect a capacitor across Q or S to balance the valve internal capacitances. The test is repeated for various increasing negative grid voltages. Curve is drawn of the dynamic plate resistance to a base of negative grid voltage as shown in figure 72 (*a*).



(a) Circuit for measuring dynamic plate resistance



(b) Circuit for the measurement of dynamic amplication factor.



(c) Circuit for the measurement of dynamic mutual conductance.

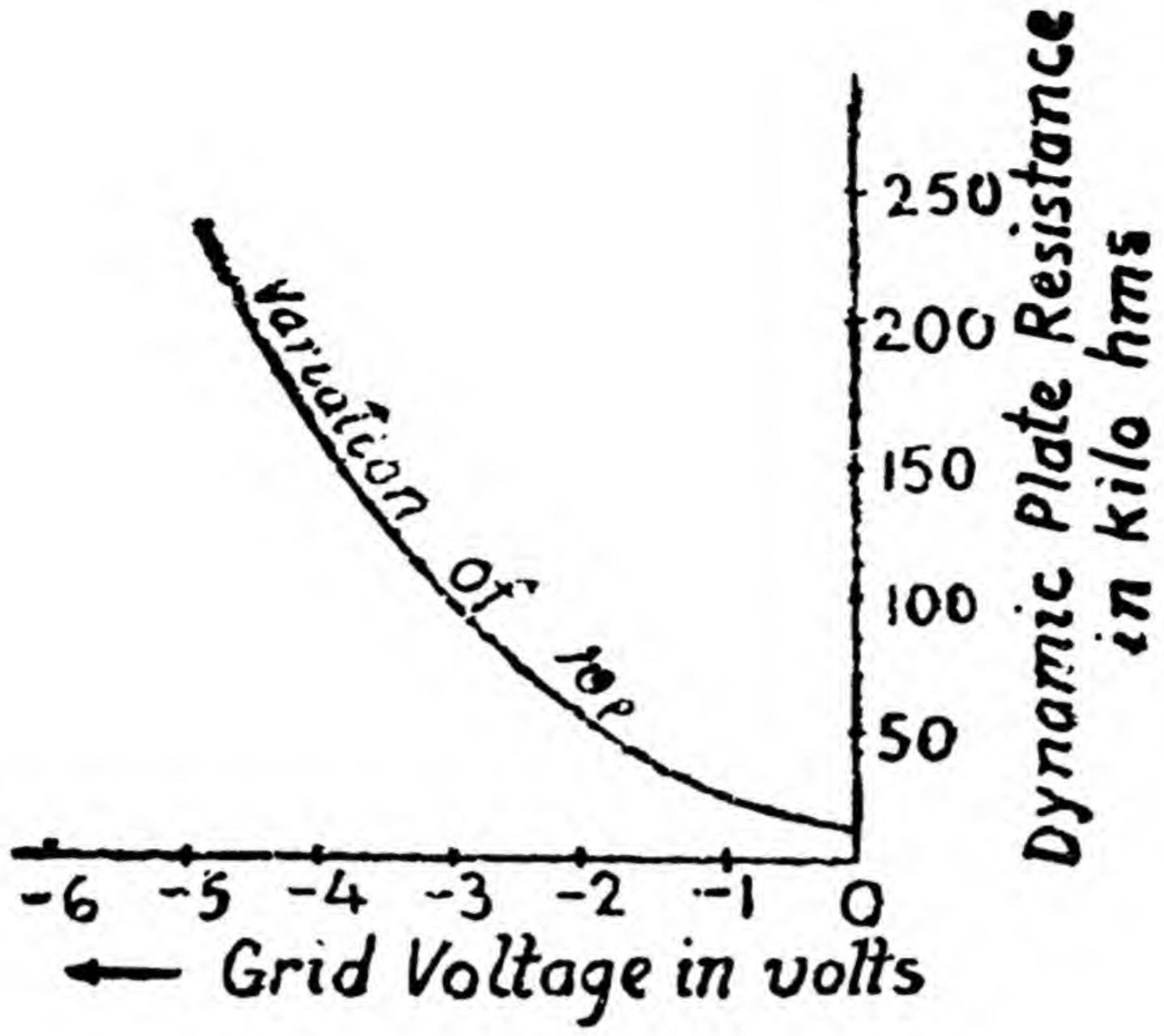
Fig. 71. Circuit arrangements for the measurement of dynamic tube constants.

(ii) **Amplification factor :** A triode of high amplification factor is connected in the circuit as shown in figure 71 (*b*).

Commencing with normal voltage applied to the plate and a low negative voltage applied to the grid, *P* is adjusted for minimum sound in headphone. This is repeated for various increasing values of negative grid voltage. At balance the amplification factor μ is given by $\mu = P/Q$.

The curve of amplification factor to a base of grid voltage is drawn as shown in figure 72 (b). Plate voltage is 120 volts.

(iii) **Mutual Transconductance :** Circuit is shown in figure 71 (*c*). An audio-frequency oscillator of low output resistance is connected in series with grid battery. Headphone connected should have resistance less than the impedance of the valve. The oscillator output should be kept low in order to prevent the grid from becoming positive at low values of negative grid potential.



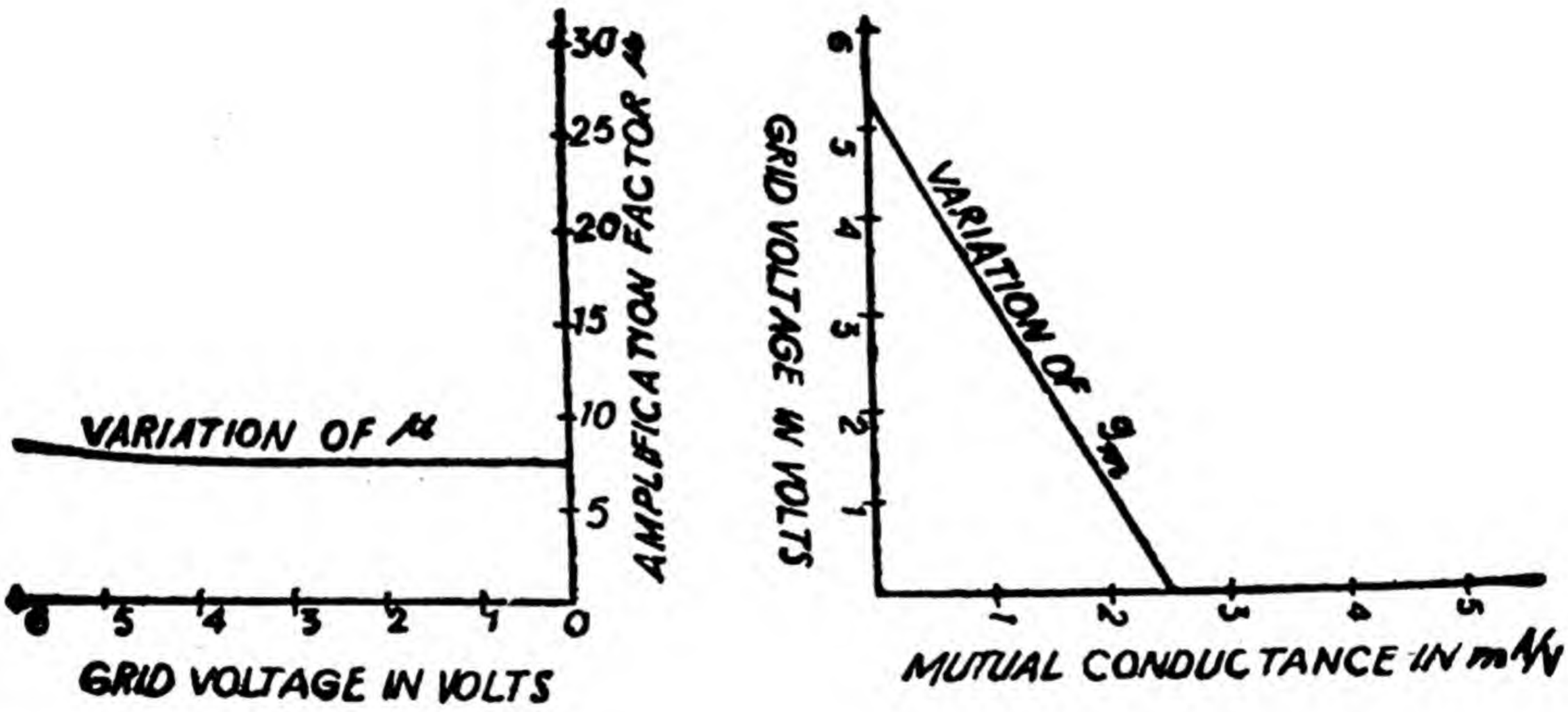


Fig. 72. Graph illustrating
(a) variation of r_p, (b) variation of μ with grid bias,
(c) variation of g_m with grid bias.

Starting with low value of negative grid voltage, R is adjusted for minimum sound in headphone. At balance, g_m is given by $1/R$. The test is repeated for increasing values of negative grid potential. Curve of mutual conductance to a base of grid bias voltage is shown in figure 72 (*c*).

Expt. No. 3. To draw the characteristic curves of gas photo cell 922*.

---

*R. C. A. 868, 918, 921, 933, are gas phototubes while 917, 922, 923 are vacuum photo tubes.

Procedure. The circuit for sketching the characteristics curves of photo-electric cell is shown in figure 73. Instead of micro-ammeter a ballistic galvanometer is inserted in the plate lead of the cell. The photoelectric current is, therefore, recorded interms of deflection.

Light from sources is allowed to fall on the cell through a shutter after passing through a filter of definite wavelength. The two characteristics are in figs. 74 and 75 :

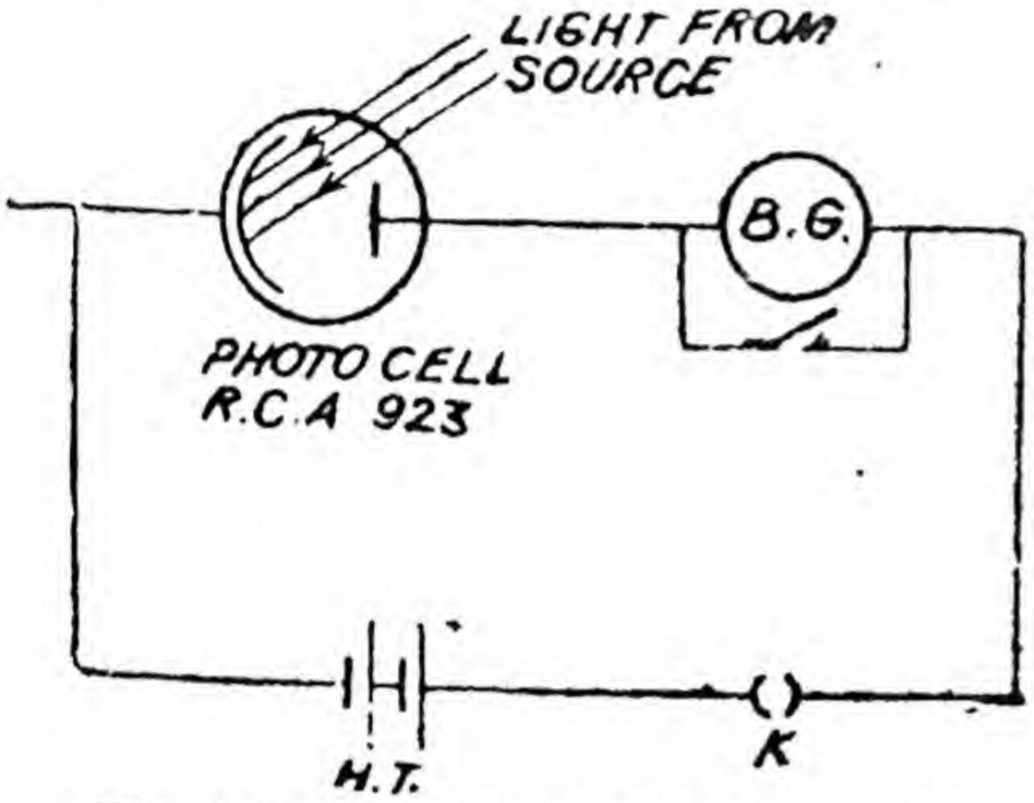


Fig. 73. Circuit for photo cell characteristics.

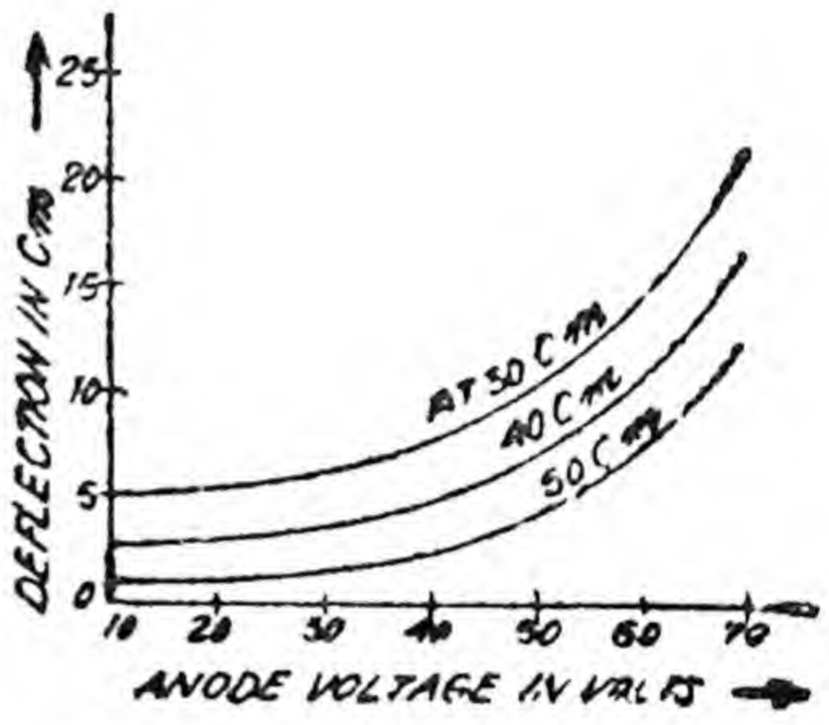


Fig. 74. Constant intensity c aracter-istics of photo cell.

(*i*) For first characteristics, we keep the intensity of incident light on the photo cell constant for which the distance between source and cell is kept unchanged. Shutter should also be of constant aperture. We vary the anode potential of photo cell and corresponding readings of photo current in terms of deflection are noted. The test is repeated for two or more distances between cell and light source. A graph, shown in fig. 74, between the anode potential and deflection is drawn for two or more fixed values of the distance between photo cell and light source.

(*ii*) For second characteristics, the anode potential of cell is kept constant, while the intensity of incident light is varied by varying the distance between the light source and photo cell. For different distances, deflections are noted. The test is repeated for other values, for different plate voltages, in deflection and $1/d^2$ are drawn where d is the distance between cell and source as shown in fig. 75.

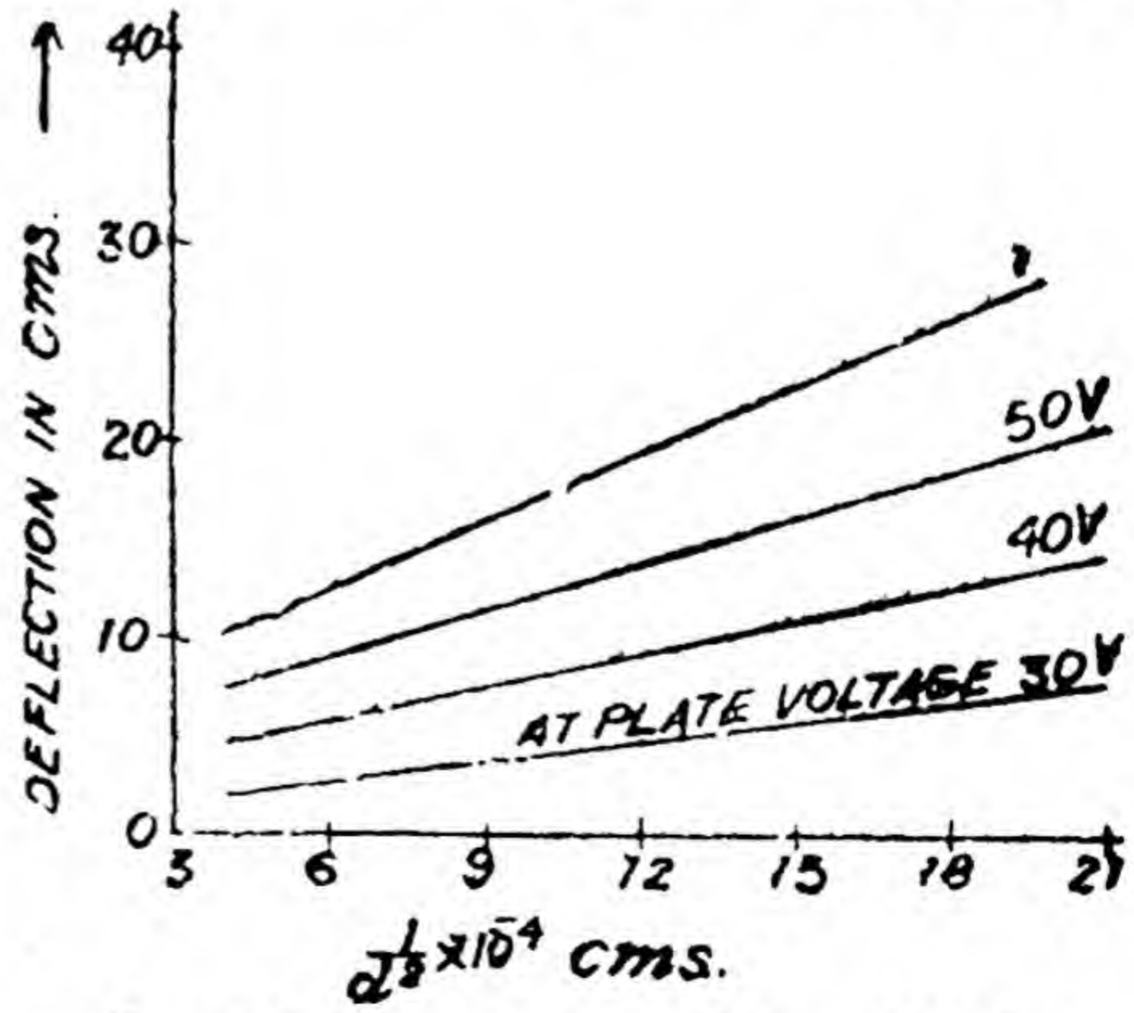


Fig. 75. Curves in $1/d^2$ and deflection at different anode voltages.

Expt. No. 4. To determine the value of Planck's constant and the photo-electric work function of the metal with the help of a photo cell.

Theory : Einstein in 1905 predicted that energy of the fastest photo-electrons would be directly proportional to the frequency of the incident light. He stated this in the form of the equation

$$\tfrac{1}{2}mv^2_{max.}=h\nu-\phi,$$

where m is the mass of electron, $v_{max.}$ is the velocity of fastest electron, ν the frequency of light, h is Planck's constant and ϕ is the energy required for an electron to escape from the surface of the emitting material.

The energy $h\nu$ of each photon depends directly on its frequency. A part of this energy, given over to an electron, is utilised in its removal from the metal surface and rest energy remains in the kinetic form.

The maximum velocity v_{max} with which the electrons are ejected can be found by putting a negative potential on the anode with respect to the cathode of photo cell (vacuum type) and determining when no current is recorded by a series elcctrometer (in experiment we use a glavanometer). The circuit arrangement is shown in figure 76. If the negative retarding potential at which this occurs is V_0, then

$$V_0e=\tfrac{1}{2}mv^2_{max}$$

Since v_{max} is a function of the frequency of incident light, therefore V_0 will also depend upon the latter. If we use filters of two different frequencies ν_1 and ν_2, then Einstein relation can be written as

$$h\nu_1=V_0'e+\phi,$$
$$h\nu_2=V''_0e+\phi,$$

where V_0' and V_0'' are retarding potentials for light of frequencies ν_1 and ν_2 respectively. If we substract the second equation from first, we get

$$h\ (\nu_1-\nu_2)=(V_0'-V_0'')\cdot e,$$
$$h\Delta\nu=\Delta V_0.e,$$
$$\frac{h}{e}=\frac{\Delta V_0}{\Delta\nu}.$$

In electrostatic units,

$$\frac{h}{e}=\frac{\Delta V_0}{300\Delta\nu}.$$

If we calculate ΔV_0 corresponding to $\Delta\nu$, the value of Planck's constant can be known readily.

Procedure : The circuit arrangement is shown in figure 76. The different retarding potentials are given to the anode of the

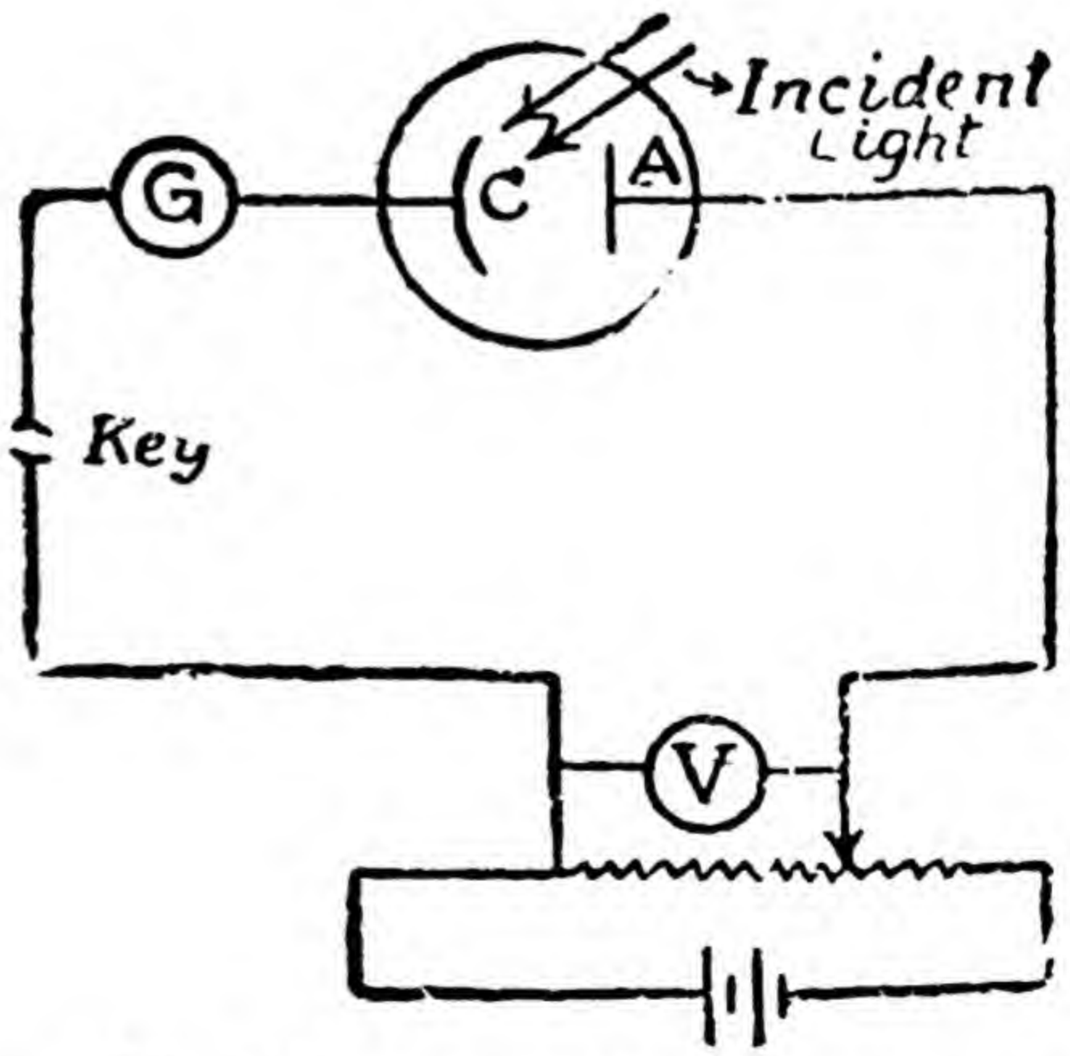


Fig. 76. Circuit for calculating retarding potentials.

photo cell. A galvanometer is inserted in the circuit to indicate the photo current with different light filters. The corresponding retarding potentials for no deflection in galvanometer are noted.

The critical retarding potentials, proportional to the maximum velocities of electrons, for sodium are plotted against the frequency of incident light as shown in fig. 77.

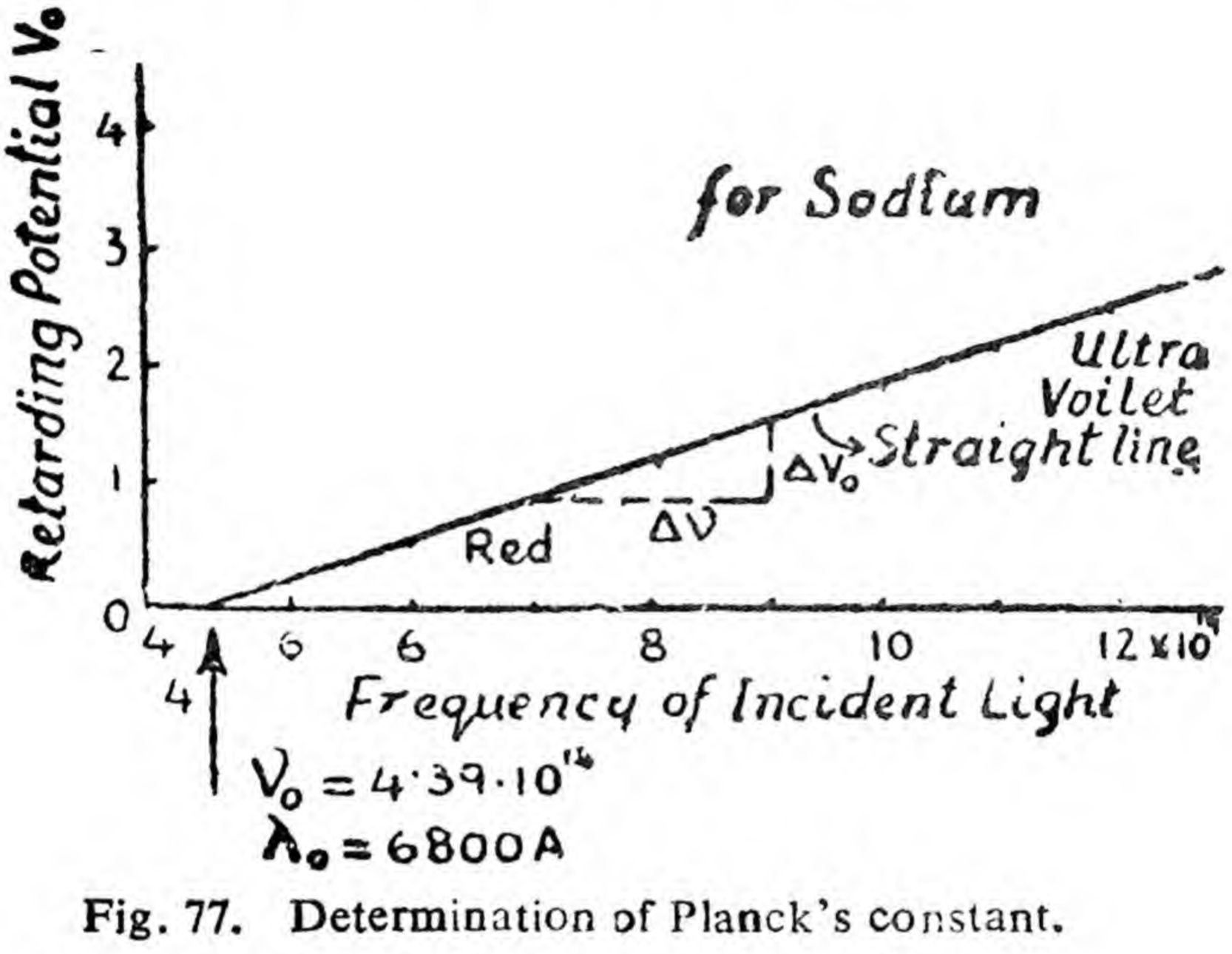


Fig. 77. Determination of Planck's constant.

From the equation,

$$V_0 e = h\nu - \phi,$$

it is obvious that graph plotted in retarding potential V_0, and frequency ν of incident light will be a straight line. From this straight line, value of ΔV_0 corr esponding to $\Delta\nu$ can be calculated. Therefore, slope of this line $\frac{\Delta V_0}{\Delta \nu}$ will give the value of $\frac{h}{e}$. Then the value of Plancks constant is calculated with the help of

$$h = e \cdot \frac{1}{300} \cdot \frac{\Delta V_0}{\Delta \nu} \text{ ergs-sec.}$$

A threshold frequency ν_0 and corresponding threshold wavelength λ_0 will exist. Below this frequency ν_0, no electron will be escaped from the metal depending upon its work function, by the photo-electric effect. Thus putting V_0 or $v_{max} = 0$ and $\nu = \nu_0$ in photo electric equation, we have

$$h\nu_0 = \phi$$

or

$$\frac{hc}{\lambda_0} = \phi.$$

By this relation photo-electric work function of the metal can be determined.

Expt. 5. (a) Striking potential of a gas filled triode.
(b) Ionisation potential of a gas filled triode.

Apparatus : Tube 884, dry batteries, voltmeter, milliameter, suspended type galvanometer.

Procedure : (a) Striking potential of a gas-filled triode.

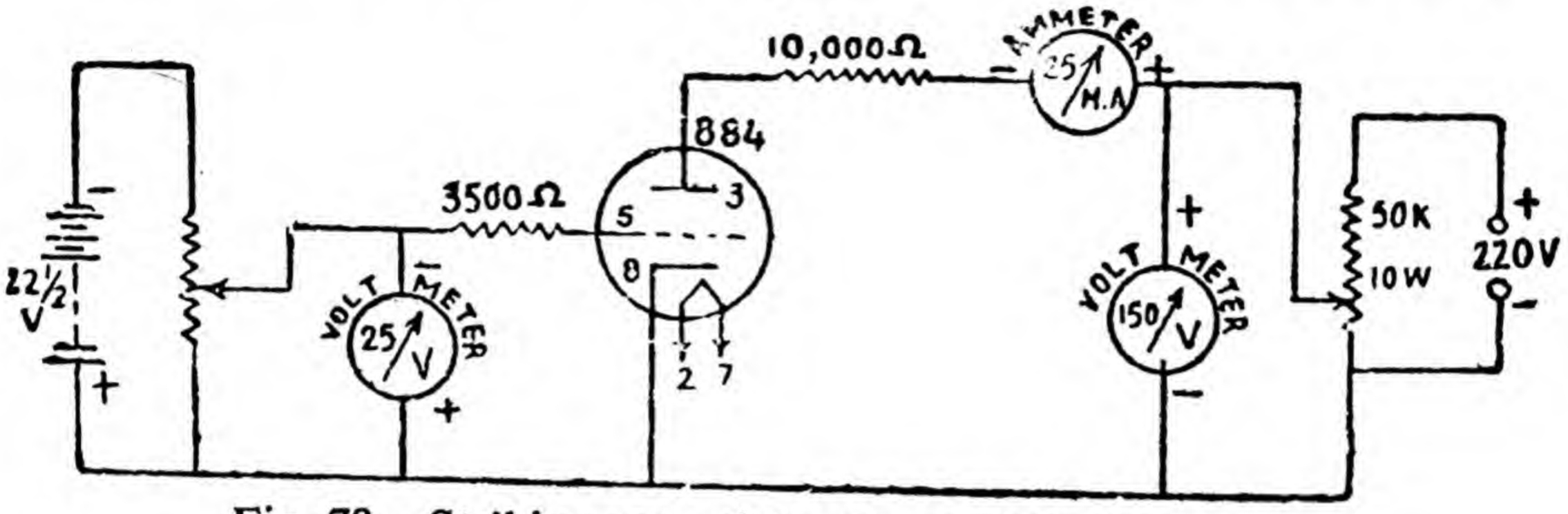


Fig. 78. Striking potentials of a gas-filled triode.

Consider the apparatus as shown in the figure 78. Apply the maximum negative voltage at the grid (greater than 15 volts, say). Turn on the full value of the plate voltage (say about 150 volts). Now change the grid voltages to make the grid less negative and note its value when the plate current just starts. Determine similar critical grid voltages for lower plate voltages. Plot a graph between critical negative grid voltage and plate voltage. The value of grid voltage is the striking or firing voltage.

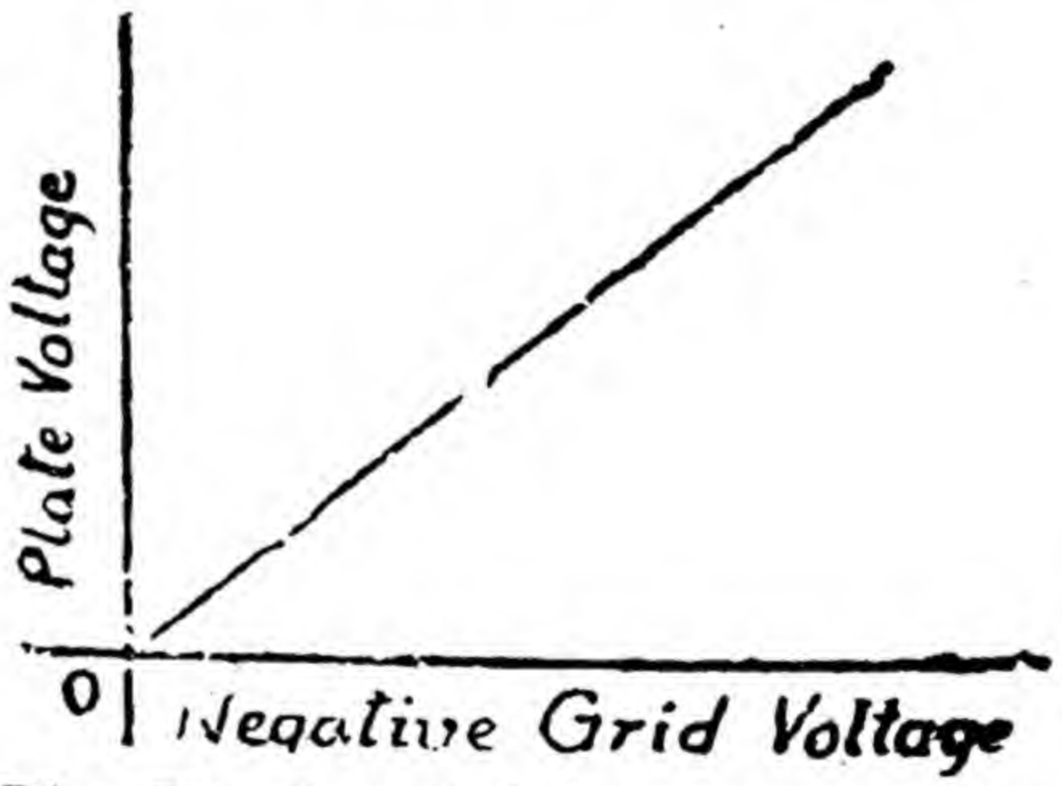


Fig. 79. Graph between negative grid voltage and plate voltage

(b) Ionisation potential of a gas filled triode.

For the ionisation potential make the connections as shown in the figure 80.

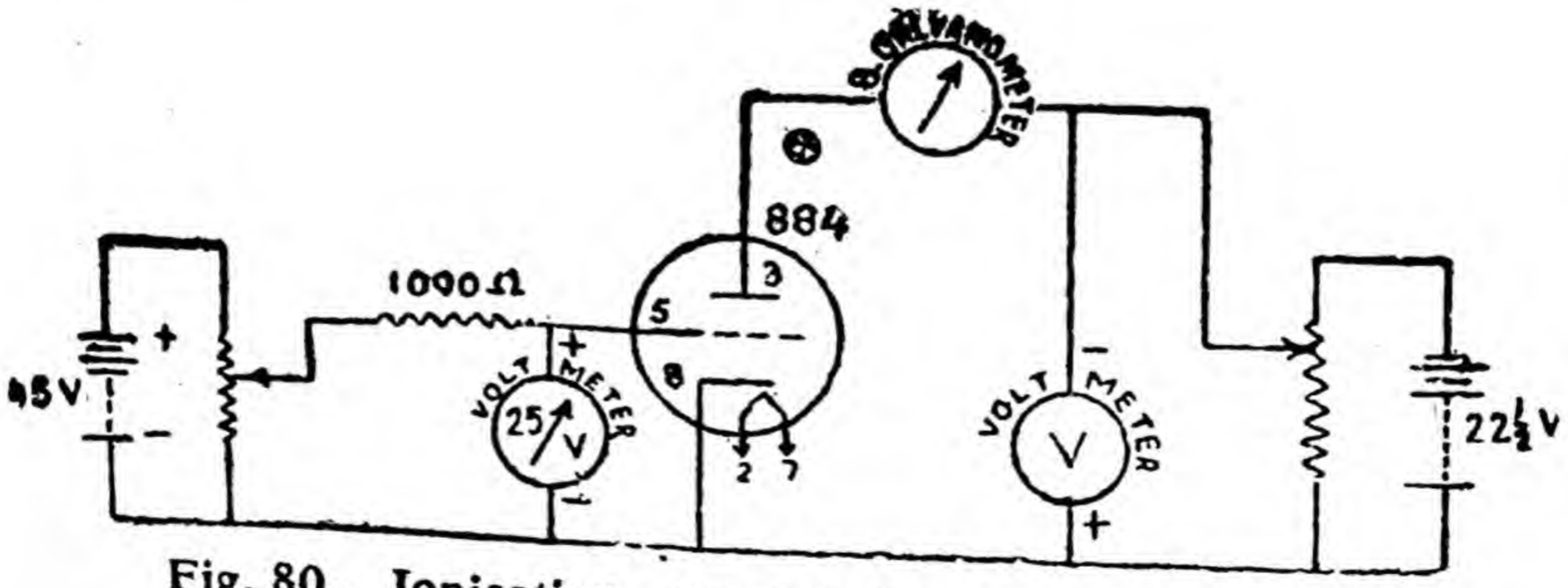


Fig. 80. Ionisation potential of a gas-filled triode.

(*i*) First of all the grid voltage is made zero, and the plate voltage is adjusted to such a value that the shunted galvanometer gives no deflection. At this stage, the plate and the cathode are at the same potential and due to the initial velocity of the emission of the electrons, the plate has to be at negative potential with respect to the cathode to keep the electrons away from the plate. The flow ceases and we get no current in the galvanometer.

(*ii*) Now the grid is maintained at positive potential with respect to the cathode and voltage is increased in steps Note the corresponding deflection in the galvanometer. The positive grid voltage accelerates the electrons and a point comes when inelastic collisions take place and the gas gets ionised. As soon as this happens, the current flows to the plate and we get a marked galvanometer deflection. This potential at which the current flows would indicate the ionisation potential of gas.

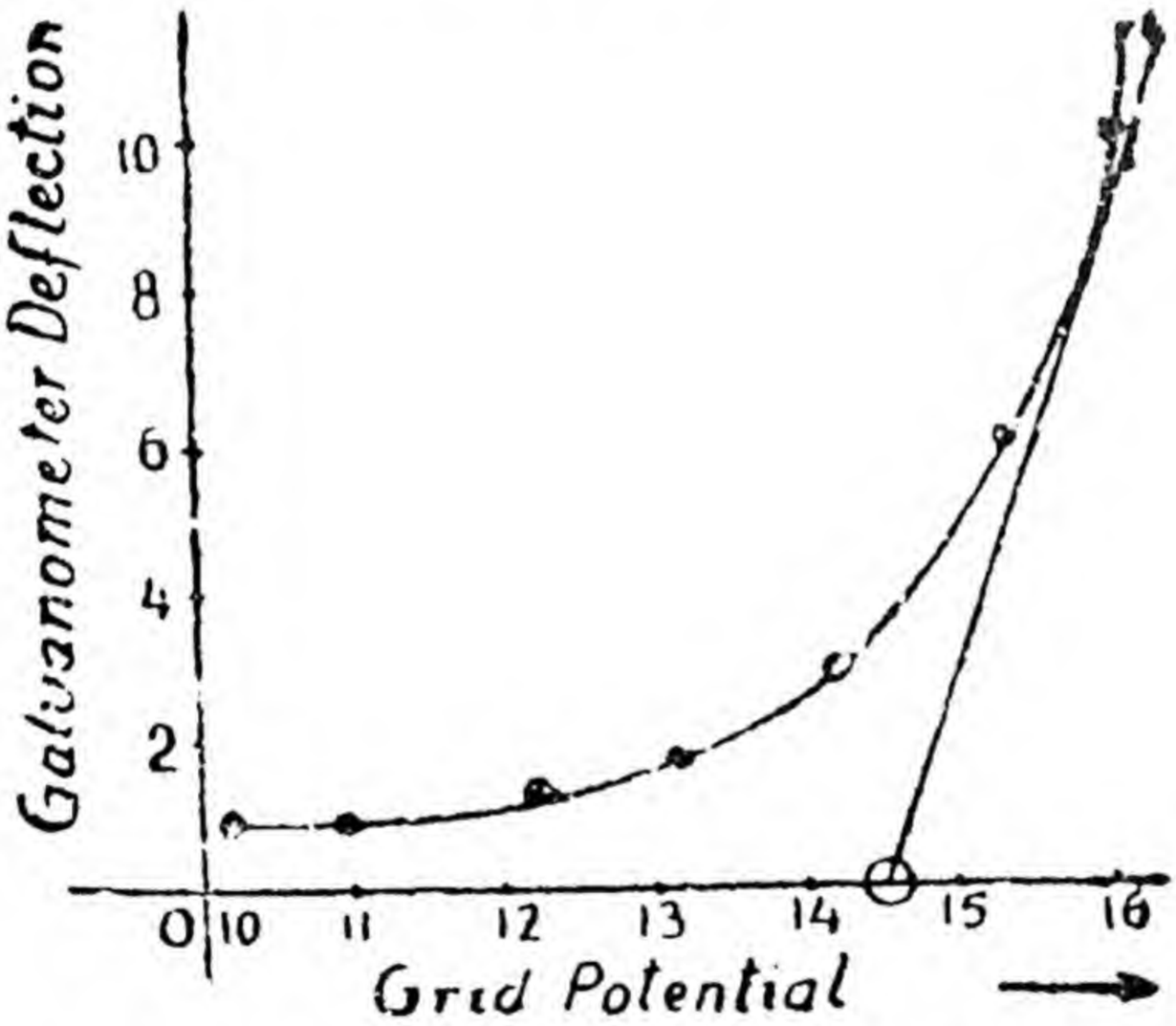


Fig. 81. Graph between grid potential and galvanometer deflection.

Draw a graph between positive grid potential and the galvanometer deflection. Calculate the value of the ionisation potential from the graph.

(*iii*) Now remove the galvanometer from the plate circuit and connect it as shown in fig. 82.

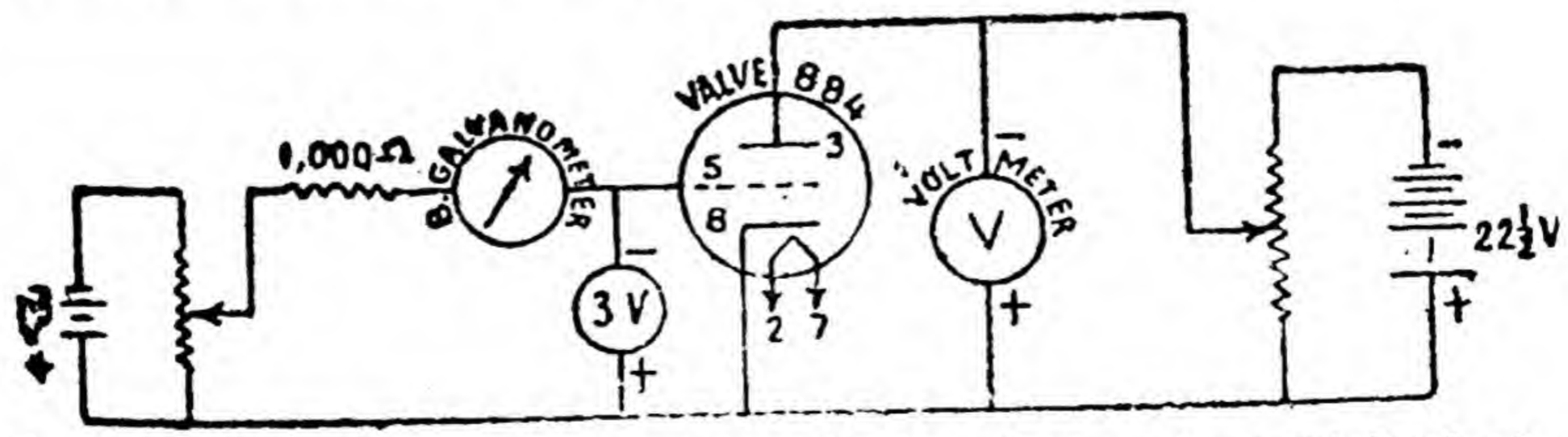


Fig. 82. Circuit arrangement to study the effect of the initial velocity of electron on ionisation potential.

Keeping the plate voltage at the same value, apply negative potential to the grid and increase it slowly till galvanometer deflection becomes zero. This negative grid potential is just sufficient to keep away the electrons which come towards the grid due to their initial velocity.

The true ionisation potential is the sum of these two potentials of the grid

Expt. No. 6. Measurement of e/m by Helical method.

Apparatus : A cathode ray tube ; a solenoid of proper dimension, in the interior of which a cathode ray tube can be placed ; a control which contains under it a power supply and controls (*a*) to operate the tube. (*b*) to operate the solenoid, (*c*) to provide variable a.c. voltage for deflecting plates, an ammeter (range d.c. one ampere), voltmeter (range 1·5 k-volts) one commutator.

Theory. Electrons from the hot filament of the cathode ray tube are accelerated and collimated so as to focus on the centre of the screen. The time t which they require to travel the distance l is equal to l/v, where v is their velocity along the axes. The later may be expressed in terms of the accelerating voltage V given by the equation $\frac{1}{2}mv^2 = e\,V \times (10^8)$, hence

$$t = \frac{l}{\sqrt{\left(\frac{eV \times 10^8}{m} \times 10^8\right)}} \text{ sec.} \qquad \ldots(i)$$

For a given deflecting force, the electrons travel in circles whose radii are small at the left or entering edge of the plate, larger at the right or emergent edge and constant magnitude from there to the screen. The velocity of an electron in one of the circles is equal to the transverse velocity, v_t, then

$$Hev_t = m\frac{v^2}{r}$$

or

$$\frac{v_t}{r} = \frac{He}{m}$$

where v_t/r is equal to the angular velocity, ω, of the electrons moving in the circles.

If T is the time for an electron to make one complete revolution, then

$$\omega = \frac{2\pi}{T} = \frac{He}{m} \text{ or } T = \frac{2\pi}{He/m} \qquad \ldots(ii)$$

For a given accelerating voltage, the magnetic field is varied until the time T for the electrons to make one rotation is equal to the time t for them to travel forward from the condenser plates to the screen (a distance l). All electrons will again be on the axis. Denoting the critical field strength by Hc and equating equations (*i*) and (*ii*), we have

$$\frac{e}{m}=\frac{8\pi^2 V.10^8}{H_C{}^2 l^2} \text{ e.m.u /gm.} \qquad \ldots (iii)$$

The field H at the centre of the solenoid is given by the equation

$$H=\frac{4\pi NI}{10L}\cos\theta, \qquad \ldots (iv)$$

where N is the total number of turns of the wire, I is the current in amperes and L total length of winding in centimeters, θ is the angle as indicated in figure 83. From eqs. (*iii*) and (*iv*),

$$\frac{e}{m}=\left[5\times 10^9\left(\frac{L}{Nl\cos\theta}\right)^2\right]\frac{V}{I^2} \text{ e.m.u./gm.}$$

Procedure :

(*i*) Record the constants of the solenoid and tube.

(*ii*) Place the solenoid such that its axis lies in the east-west direction. Mount the cathode ray tube inside the solenoid at the centre. The power unit should be kept as far away as possible to avoid the stray magnetic field.

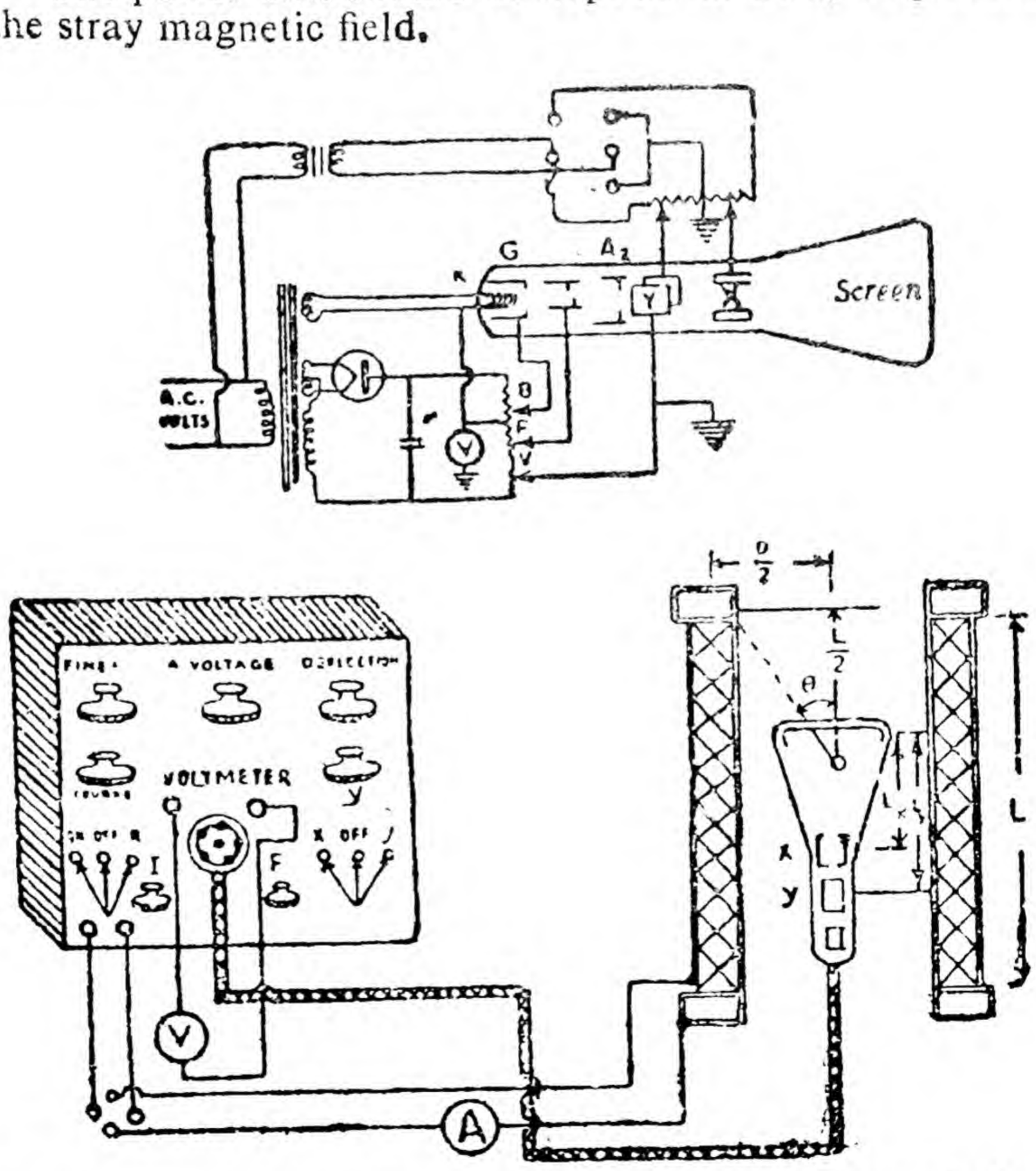


Fig. 83. Vacuum tube circuit and solenoid used in Helical method.

(*iii*) **Switch on the power supply unit. Turn the potentiometer marked "Accelerating voltage", and adjust the voltage to any desired value. With the help of *F* (fine) and *I* make a fine and clear spot on the cathode ray tube.**

(*iv*) Apply a.c. potential to one set of plates. A deflection of 2 or 3 cm. is adequate for the experiment.

(*v*) Turn on the solenoid current and increase the current till the line is reduced to a small point. Reverse the solenoid current and readjust the control to a fine point. The average of these two currents in amperes is *I*.

(*vi*) Take simultaneous values of *I* and *V* using *X* and *Y* plates.

(*vii*) Now repeat with three other voltages ; it will be necessary to refocus the spot on the tube at each voltage.

Observations :

(*a*) 1. Distance between the edge of *X* plate and the screen l_X,
2. distance between the edge of *Y* plate and the screen l_Y,
3. distance of the solenoid *D*,
4. length of the winding *L*,
5. number of turns *N*.

(*b*)

S. No.	Using *X*-Plates			Using *Y*-Plates		
	Direction of Current	A. C. voltage	Current in amperes	Direction of current	A. C. voltage	Current
1.	On Reverse	650 volts	0·535	On Reverse	650 volts	0·62
2.	—	—	0·535	—	—	0·62
3.	—	—	—	—	—	—
	—	—	—	—	—	—
	—	—	—	—	—	—
	—	—	—	—	—	—

Calculations :

(*a*) Using *X* plates

$$\left(\frac{e}{m}\right)_X = \left[\, 5\times 10^9 \left(\frac{L}{Nl_X \cos\theta}\right)^2\right] \frac{V}{I^2} \text{ e.m.u./gm.}$$

(*b*) Using *Y* plates,

$$\left(\frac{e}{m}\right)_Y = \left[\, 5\times 10^9 \left(\frac{L}{Nl_Y \cos\theta}\right)^2\right] \frac{V}{I^2} \text{ e.m.u./gm.}$$

The mean of these two values gives the value of *e*/m.

Expt. No. 7. To determine e/m for cathode rays by Thomson method.

Apparatus : Cathode ray tube, power supply, magnetic needle, stop watch, two bar magnets of the same strength.

Theory : Figure shows the deflected path of the electron in the tube.

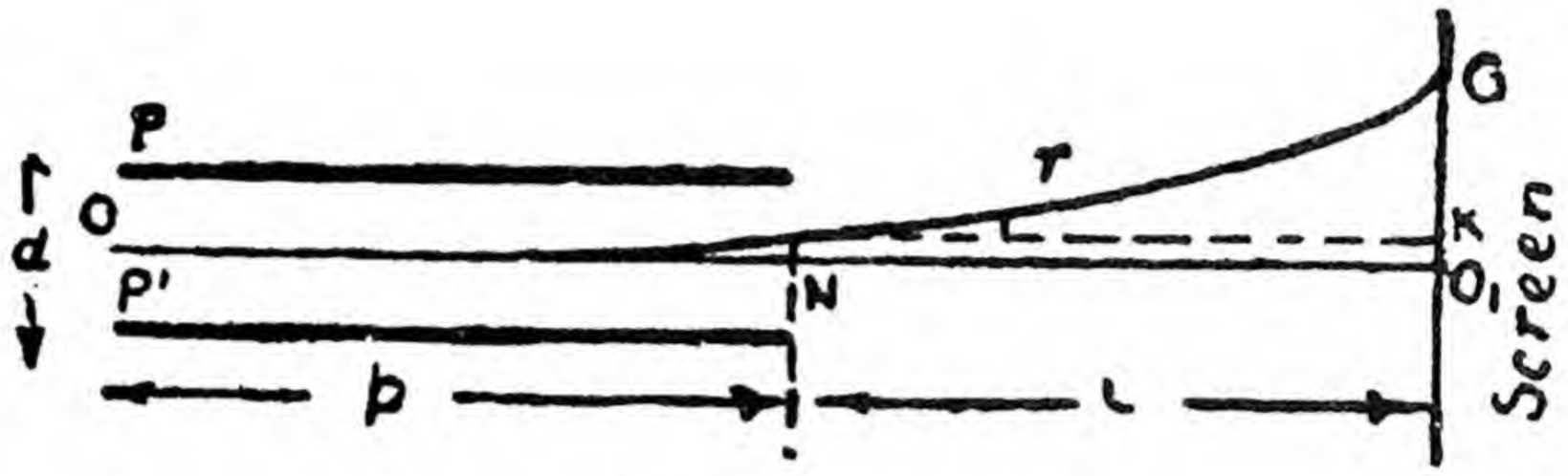


Fig. 84.

Here p is the length of the deflecting plate, d the spacing between deflecting plates and L, the distance of the screen from the edge of the plates.

If V is the potential applied between the plates and H is the intensity of the applied field, then the velocity of the electron is

$$v=\frac{(p/2+L)\,pV.10^8}{d\int\left[\int H\,dx\right]dx} \qquad \ldots(i)$$

The total deflection on the screen is $O_1G=y$, then

$$\frac{e}{m}=\frac{v_y}{\int\left[\int H\,dx\right]dx} \qquad \ldots(ii)$$

$$\text{Hence } \frac{e}{m}=\frac{(p/2+L)\,pVy\;10^8}{d\left[\int\left\{\int H\,dx\right\}dx\right]^2} \text{ coulombs/gm.} \qquad \ldots(iii)$$

Procedure :

(*i*) Place the cathode ray tube on a line facing north and

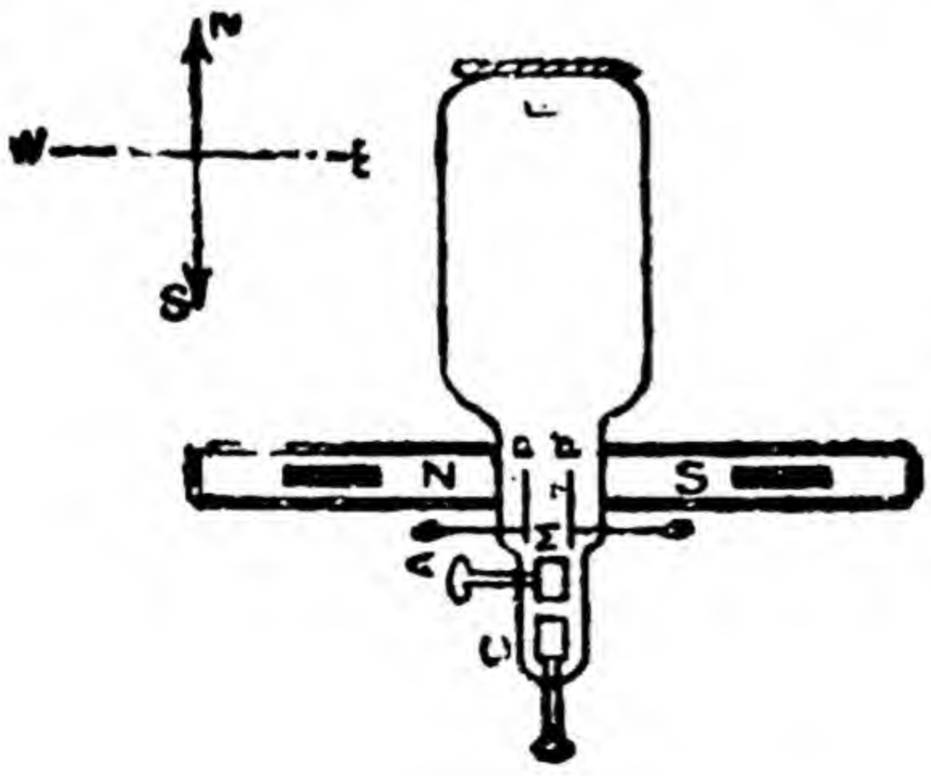


Fig. 85

PP′ → Parallel pl
C → plane-faced hode
A → ring anode
M → small hole

south. Draw a line perpendicular to it and place the wooden stand along this line.

(*ii*) Switch on the power supply and adjust the intensity so that there is a small and clear spot on the screen.

(*iii*) Read the initial reading of the spot, say it is —0·3, and give a deflection upwards such that the final reading is +0·7 *i.e.*, the total deflection is 1·0. Note the potential applied to the plates of the tube.

(*iv*) Arrange the poles *N* and *S* symmetrically and at such a distance apart that the field is of sufficient strength to reduce the deflection to its initial stage as shown in figure 85.

(*v*) Switch off the power supply and remove the cathode ray tube. Place the magnetic needle in the wooden stand and note down time period at different stages.

(*vi*) Note the time period when the magnetic field is removed.

Observations :

(*a*) Initial position of the spot on the scale,
Final position of the spot,
Total deflection of the spot *y*,
Voltage applied to the plate, *V*.

(*b*) Length of the deflecting plate, *p*,
Spacing between deflection plates *d*,
Distance of the screen from the edge of the plates *L*,
Distance *MF*, *l*.

(*c*) It will be apparent that the magnetic field used is not the uniform one; therefore the value $\iint\left[\int H\,dx\right]dx$ of equation (iii) must be found graphically. The integral can be expressed as

$$F=\int_0^x H\,dx, \text{ so that } \frac{dF}{dx}=H.$$

The value of $\int_0^l F\,dx$ is required.

Integrating by parts,

$$\int_0^l F\,dx=\left[\,Fx\,\right]_0^l-\int_0^l x\,\frac{dF}{dx}.dx=l\int_0^l H\,dx-\int_0^l xH\,dx$$

$$=\int_0^l (l-x)\,H\,dx.$$

Thus, a table of values of *H* and $(l-x)$ is required from which product $(l-x)\,H$ is plotted against *x* and the value of the integral is derived from the area of graph.

S. No.	Distance of the needle from M x (cm.)	Distance of needle from screen $(l-x)$	Time for 10 vibrations	Time for one vibration T (secs.)	θ	$\sin\theta$	$H=\frac{T_0^2}{T^2}H_0\sin\theta$	$H(l-x)$
1	0	17	...	...	...	...	...	...
2	1	16	...	...	...	...	...	...
3	2	15	...	...	...	...	...	...
4	3	...	...	...	...	...	...	...
5	4	...	...	...	...	...	...	...
6	5	...	...	...	...	...	...	...
7	6	...	...	...	...	...	...	...

If T_0 is the time period when the magnetic field is removed, then

$$H=\frac{T_0^2}{T^2}H_0\sin\theta.$$

Calculation : The graph between $H(l-x)$ and x is of the form given in fig. 86.

Find the area of the curve which gives the value

$$\int\left[\int H\,dx\right]dx$$

and applying the equation (iii), find the value of e/m.

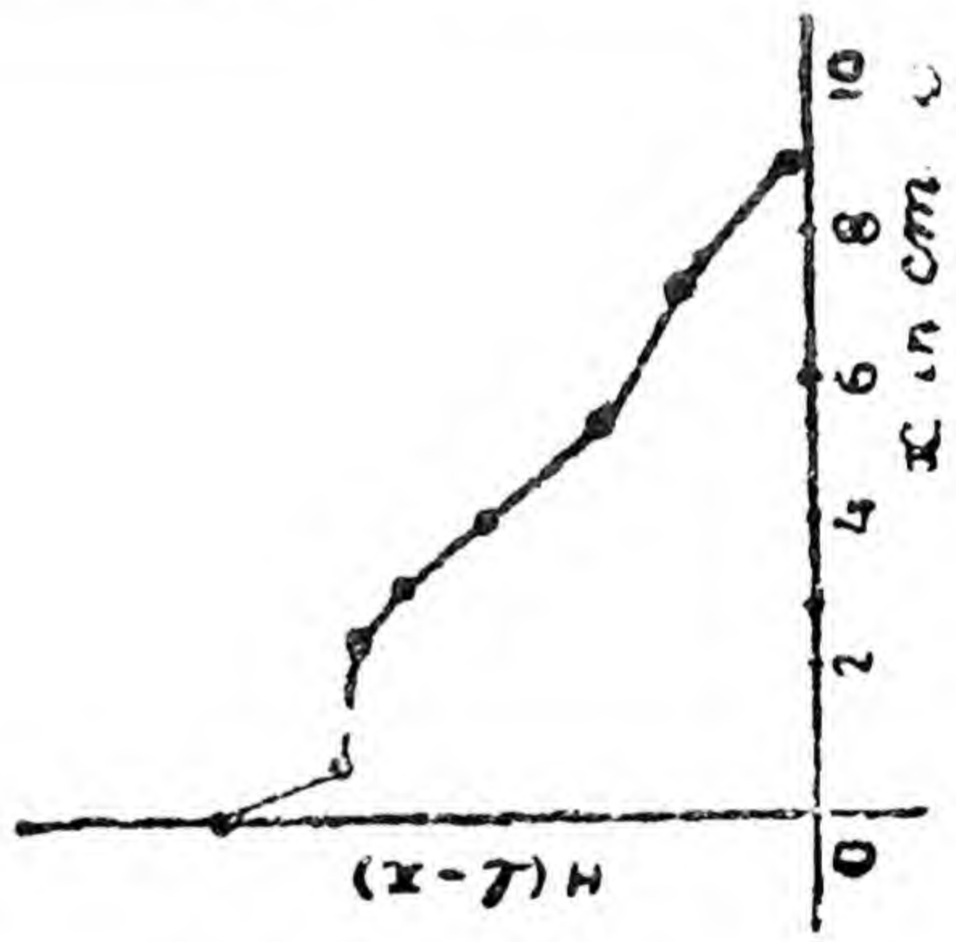


Fig. 86.

Expt. No 8. Measurement of e/m by Magnetron Method.

Apparatus. Magnetron valve (*RCA* 41 or 42), solenoid in which the valve can be placed, rectified power supply for plate voltage, ammeter, voltmeter, battery to heat up the filament.

Theory : In the theory of magnetron (cylindrical electrodes), with some simplifying assumption, we will deduce that

$$B_c=\frac{\sqrt{(8V_a)}}{\sqrt{\left\{\frac{e}{m}\left(1-\frac{r_k^2}{r_a^2}\right)\right\}}\,r_a}$$

If the cathode radius r_k is less than one tenth of the anode radius r_a, then the error in B_o due to neglecting r_k is less than one percent; hence

$$B_o=\frac{\sqrt{(2V_a)}}{r_a\sqrt{\left(\frac{e}{m}\right)}}$$

where B_c is the field intensity and V_a is the anode potential. In this formula B_c is in oersteds and V_a in e. m. units.

$$1 \text{ volt} = 10^8 \text{ e. m. u.}$$

The control grid and screen grid potentials of 41 and 42 pentodes do not affect the result because it is independent of the potential distribution.

Procedure :

(*i*) Make the connections as shown in the diagram 87.

(*ii*) Place the magnetron valve in the centre of solenoid to avoid the edge effects. Send a suitable current through the filament so that it becomes hot.

(*iii*) Apply a suitable potential difference between plate and filament. This results in a current recorded in the ammeter of the plate circuit.

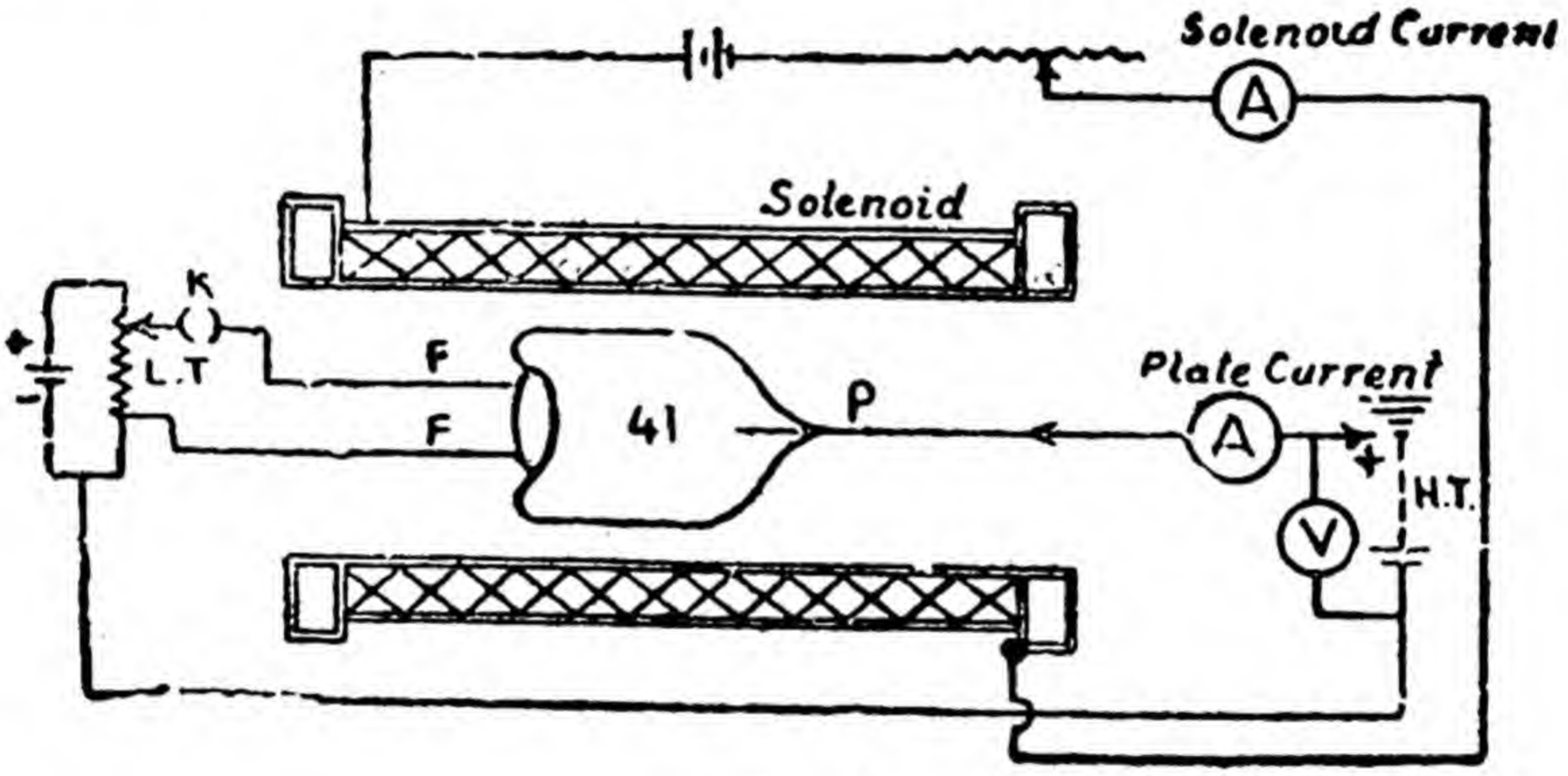


Fig. 87.

(*iv*) When this has been achieved, apply a magnetic field by means of solenoid. Increase the magnetic field from zero to a certain value at which the plate current suddenly suffers a great

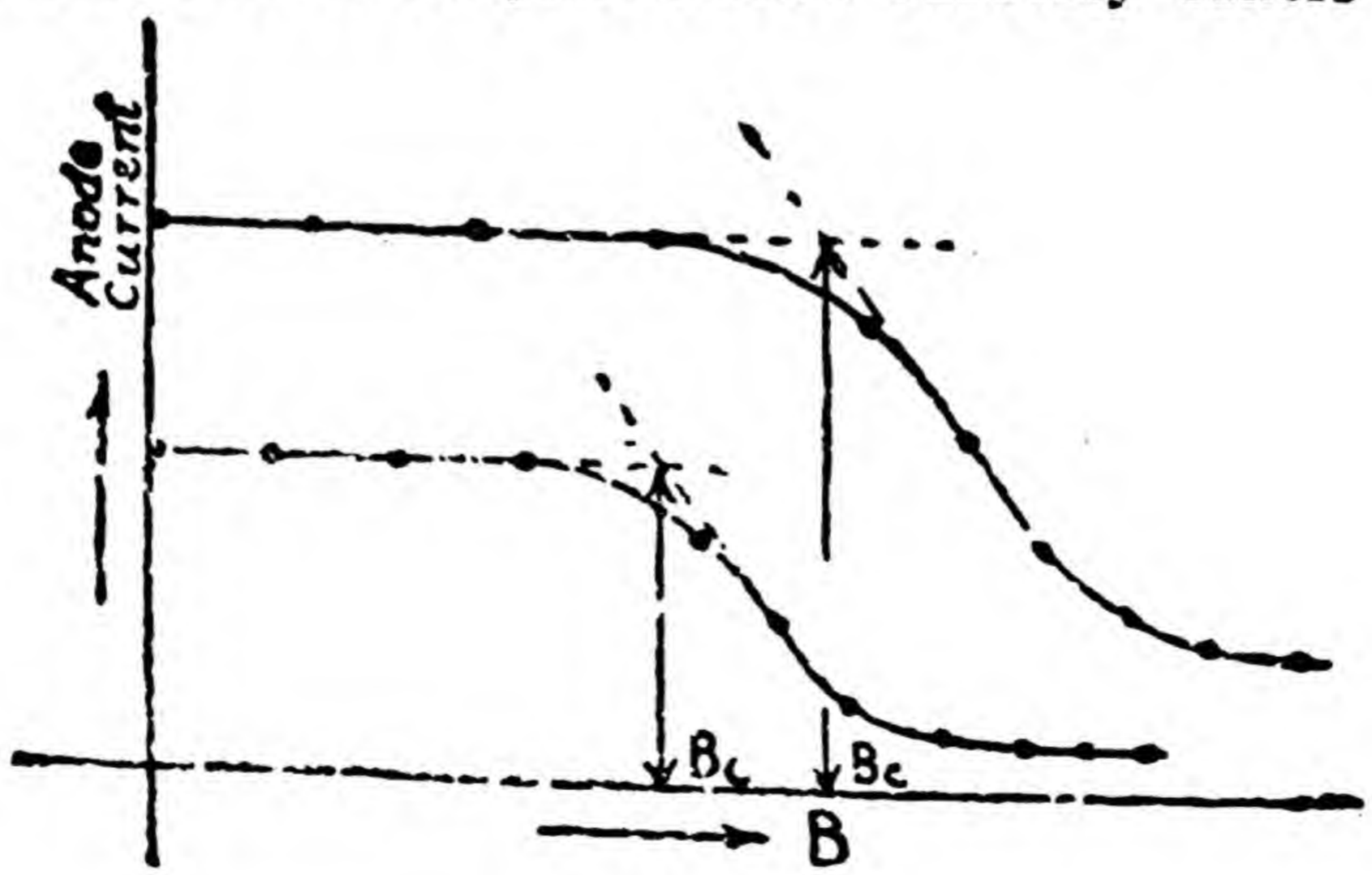


Fig. 18

variation. At this time, the electrons emitted by the cathode will be unable to reach the anode. The magnetic field corresponding to this current is B_c, the cut off value of the field.

(*v*) Draw a graph between the magnetic field and the anode current for a fixed value of anode potential.

If the horizontal and falling parts of the curve are produced, their intersection may be taken to define the knee of the characteristic. The corresponding value of the field gives the value of B_o. So with the help of this graph, we find the cut-off value of the magnetic field.

(*vi*) Record the anode voltage.

Calculation : For different solenoid currents, find the value of magnetic field using the formula

$$B=\frac{4\pi Ni}{10};$$

then
$$\frac{e}{m}=\frac{8\times 10^8\times V_a^2}{r_a^2\times B_c^2} \text{ e. m.u /gm.}$$

NUMERICAL EXAMPLES

1. *If the transverse magnetic field in a C. R. O. is 1·30 milli wb/m² over a distance of 6 cm , find the value of the accelerating voltage if the electron beam is to be deflected through 15°.*

We know that

$$F=BQ_e v=\frac{mv^2}{r}.$$

$$\therefore \qquad v=\frac{BQ_e r}{m} \qquad \text{...(i)}$$

The deflection on *C. R. O.* is given by

$$\sin\theta=\frac{l}{r},$$

so that
$$r=\frac{l}{\sin\theta}=\frac{0\cdot 06}{\sin 15^\circ}=0\cdot 232 \text{ metre.} \qquad \text{...(ii)}$$

Substituting the value of r, from equation (ii) in equation (i), we have

$$v=1\cdot 30\times 10^{-3}\times 1\cdot 76\times 10^{11}\times 0\cdot 232$$

where
$$\frac{Q_e}{m}=1\cdot 76\times 10^{11}$$

$$=5\cdot 32\times 10^7 \text{ metre/sec.}$$

and
$$V=\frac{v^2 m}{2Q_e}=\frac{(5\cdot 32\times 10^7)^2}{2\times 1\cdot 76\times 10^{11}}$$

$$=\mathbf{8040 \text{ volts.}}$$

2. *A potential difference of 300 volts is applied to two parallel plates 6 cm. long and 2 cm. apart. An electron beam projected is deflected by 5·37 mm. while passing between the plates. If simulta-*

neously a magnetic field of 0·5 m. wb/m² is applied perpendicular to the electric field, no deflection occurs. Determine (i) the velocity with which the electron enters the field, and (ii) the ratio of Q_e/m of the electron.

(i) If v_k is the velocity with which the electron enters the field and V_d is the potential between two plates, then

$$\frac{Q_e V_d}{d} = Q_e B v_k \text{ or } v_d \frac{V_d}{dB} = \frac{300}{0{\cdot}02 \times 0{\cdot}5 \times 10^{-3}}$$

$$= 3 \times 10^7 \text{ metres/sec.}$$

(ii) We know that $\tan\theta = \frac{d}{(l/2)} = \frac{0{\cdot}00537}{0{\cdot}03} = 0{\cdot}179$

$$\therefore \quad \theta = 10{\cdot}15°.$$

Now $$r = \frac{l}{\sin\theta} = \frac{0{\cdot}06}{\sin 10{\cdot}15°} = 0{\cdot}34 \text{ metre.}$$

Again $$r = \frac{mv_k}{BQ_e} \text{ or } \frac{Q_e}{m} = \frac{v_k}{Br} = \frac{3 \times 10^7}{0{\cdot}5 \times 10^{-3} \times 0{\cdot}34}$$

$$= 1{\cdot}7 \times 10^{11} \text{ Coul./Kg.}$$

3. *In a C.R. tube a pair of deflecting plates are 4 cm. long and are spaced 0·5 cm. apart. The distance between the flourescent screen and the deflecting plate is 16 cm. The find anode voltage is 200 volts. Calculate (a) displacement produced by a deflecting voltage of 20 volts. (b) angle of beam with the tube axis on emerging from the field.*

(*a*) Total deflection $d = \frac{V_d l}{2V_a S}(\frac{1}{2}l + D)$

$$= \frac{20 \times 4}{2 \times 2000 \times 0{\cdot}5}(\tfrac{1}{2} \times 4 + 16),$$

$$d = \frac{18}{25} = 0{\cdot}72 \text{ cm.}$$

(*b*) $\tan\phi = \frac{d}{\frac{1}{2}l + D} = \frac{V_d l}{2V_a S}$

$$= \frac{20 \times 4}{2 \times 2000 \times 0{\cdot}5}$$

$$= 0{\cdot}04$$

$$\phi = 2° \ 18'.$$

4. *In a cathode ray tube a pair of deflecting plates are 2·0 cm. long and are spaced 0·5 cm. apart. The distance from the centre of the plate to the screen is 24 cm. The final anode voltage is 1000 volts. Calculate (a) the displacement produced by deflecting voltage of 30 volts (b) the angle which the beam makes with the axis of the tube on emerging from the field, and (c) velocity of the beam on emerging from the field.*

(*a*) Approximately the deflection produced

$$d=\frac{V_d/D}{2.SV_a}$$

$$=\frac{30\times2\times24}{2\times0\cdot5\times1000}\text{ cm.,}$$

$$d=\mathbf{1\cdot44\text{ cm.}}$$

(*b*) $$\tan\phi=\frac{d}{D}=\frac{V_dl}{2.S.V_a}$$

$$=\frac{30\times2}{2\times0\cdot5\times1000}=0\cdot06,$$

$$\phi=\tan^{-1}0\cdot06=\mathbf{3^\circ\ 26'}.$$

(*c*) $$v=\sqrt{\left(\frac{2eV_a}{m}\right)}$$

$$=5\cdot94\times10^6\sqrt{(V_a)}\text{ metres/sec.}$$

$$=5\cdot94\times10^5\sqrt{(1000)},$$

$$v=\mathbf{18\cdot78\times10^6}\text{ metres /sec.}$$

5. *The spot of C. R. tube is displaced vertically by magnetic field of flux density 10^{-3} webers per square meter. The space between the y deflection plates is 0·4 cm. The final anode voltage is 600 volts. The length of the magnetic field is 3 cm. while that of the deflecting plates is also 3 cms. Determine the voltage to be applied to y-deflecting plates to return the spot to the centre of the screen.*

Take $m=9\cdot0\times10^{-31}$ kgm. and $e=1\cdot6\times10^{-19}$ Coulombs.
For magnetic deflection,

$$d=\frac{DlB\sqrt{\left(\frac{e}{m}\right)}}{\sqrt{(2V_a)}}.$$

For electrostatic deflection,

$$d=\frac{V_dlD}{2V_aS}$$

If magnetic deflection=electrostatic deflection

then $$\frac{DlB}{\sqrt{(2V_a)}}\sqrt{\left(\frac{e}{m}\right)}=\frac{V_d}{V_a}.\frac{lD}{2S}.$$

$$V_d=\sqrt{(2V_a)}.BS\sqrt{\left(\frac{e}{m}\right)}$$

$$=\sqrt{(2\times600)}\ 10^{-3}\times0\cdot4\times10^{-2}\times\sqrt{\left(\frac{1\cdot6\times10^{-19}}{9\times10^{-31}}\right)}$$

$$V_a=\mathbf{58\cdot4\text{ volts.}}$$

6. *In a cathode ray tube employing magnetic deflection, the distance from the screen to the centre of the coil is 24 cms. The length of the field along the axis is 4 cms. Calculate the magnetic flux density required to reduce a deflection of 1·2 cm. on the screen for final anode voltage (a) 400 and (b) 600 volts.*

$$d = \frac{D.l.B\sqrt{\left(\frac{e}{m}\right)}}{\sqrt{(2V_a)}}$$

(*a*) For $V_a = 400$ volts,

$$1{\cdot}2\times10^{-2} = \frac{B\times24\times10^{-2}\times4\times10^{-2}(1{\cdot}7\times10^{11})^{1/2}}{\sqrt{(2\times400)}}$$

$$= B\times96\sqrt{(2{\cdot}2)},$$

$$B = \frac{1{\cdot}2\times10^{-2}}{96\times\sqrt{(2{\cdot}2)}}$$

$$B = \mathbf{8{\cdot}428\times10^{-5}\ weber/sq.\ metre.}$$

(*b*) For $V_a = 600$ volts,

$$1{\cdot}2\times10^{-2} = \frac{B\times24\times10^{-2}\times4\times10^{-2}\ (1{\cdot}7\times10^{11})^{1/2}}{\sqrt{(2\times600)}}$$

$$= B\times96\sqrt{\left(\frac{17{\cdot}6}{12}\right)}$$

$$B = \frac{1{\cdot}2\times10^{-2}\sqrt{(12)}}{26\times\sqrt{(17{\cdot}6)}}$$

$$B = \mathbf{10{\cdot}35\times10^{-5}\ webers/sq.\ metre.}$$

7. *A sinusoidal voltage 20 sin 100πt is applied to the y deflection plates of a C. R. tube. Time base voltage is applied on the x deflection plates. What should be the frequency of the time base if two complete sine waves are required to be seen on the screen. If the time base voltage is required to be obtained from a neon time base, calculate the value of charging to be used. It is given that charging condenser = 0·01 μf. Battery voltage $E_{d.c.} = 220$ volts, striking voltage = 150 volts, extinguishing voltage = 120 volts.*

Periodic time of sine wave $= \frac{1}{50}$ sec.

To get two sine waves, periodic time base

$$= 2\times\frac{1}{50} = \frac{1}{25}\ \text{sec.}$$

∴ Frequency of time base = 25 c/s.

For neon time base, $f = \dfrac{1}{CR\log_e\dfrac{E_{d.c.}-V_e}{E_{d.c.}-V_s}}$.

For this case

$$50 = \frac{1}{R\times\ 01\times10^{-6}\times2{\cdot}306\log_{10}\dfrac{220-120}{220-150}}$$

or

$$R = \frac{1}{50\times{\cdot}01\times10^{-6}\times2{\cdot}306\log_{10}\dfrac{100}{70}}$$

$$= \mathbf{2{\cdot}25\times10^{5}\ ohms.}$$

8. *The following table gives plate currents for different values of plate voltage for grid voltages zero and −2 volts, using a triode valve :*

Plate voltage in volts	80	100	120	140
Plate current in mA, for grid voltage$=2V$	8	10·9	13·8	—
Plate current in mA, for grid voltage$=-2V$	2·2	5·1	8	10·8

Calculate values of amplification factor μ, *dynamic plate resistance,* r_p *and mutual conductance,* g_m, *for the tube at plate voltage of 100 volts.*

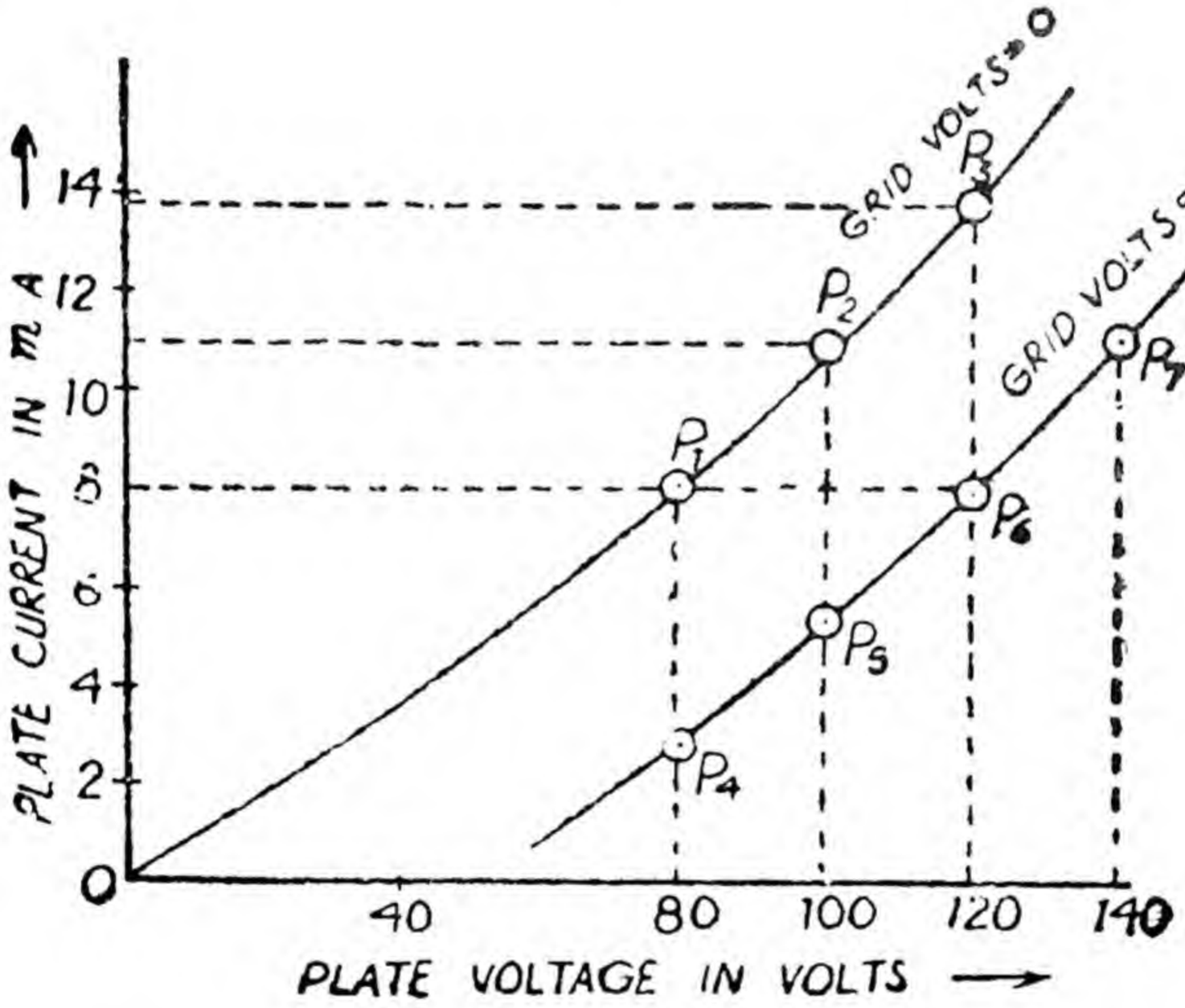


Fig. 89.

From figure,

$e_{c_1}=0$ between points P_1 and P_2

$\Delta e_b=(120-80)=40$ volts

and $\Delta i_b=(13\cdot8-8\cdot0)=5\cdot8$ mA.

Hence $$r_p=\left|\frac{\Delta e_b}{\Delta i_b}\right|_{e_p}$$

$$=\frac{40}{5\cdot8\times10^{-3}}$$

$\mathbf{r_p=6\cdot9}$ **kilo-ohms.**

At $e_b=100$ volts, between point P_2 and P_5.

$\Delta i_b=(10\cdot9-5\cdot1)=5\cdot4$ mA.

$\Delta e_c=2$ volts.

Hence $g_m = \left|\frac{\Delta i_b}{\Delta e_o}\right|_{e_b}$

$$= \frac{5{\cdot}4 \times 10^{-3}}{2} \text{ mho.}$$

$g_m = 2{\cdot}7$ **milli-mho.**

Now $\mu = r_p \times g_m$

$= 6{\cdot}9 \times 10^3 \times 2{\cdot}7 \times 10^{-3}$.

$\mu = 18{\cdot}63$.

9. ***The following table gives two amplification characteristic curves of a triode. Calculate the coefficients of the triode at a grid voltage of zero volt.***

Grid voltage in volts	—30	—20	—10	0	+10	+20
Anode voltage for anode current of 5 mA.	450	370	300	220	160	110
Anode voltage for anode current of 10 mA	495	415	340	260	197	145

Considering points P_1 and P_2. [Fig. 90]

$\Delta e_o = 260 - 220 = 40$ volts,.

$\Delta i_b = 10 - 5 = 5$ mA.

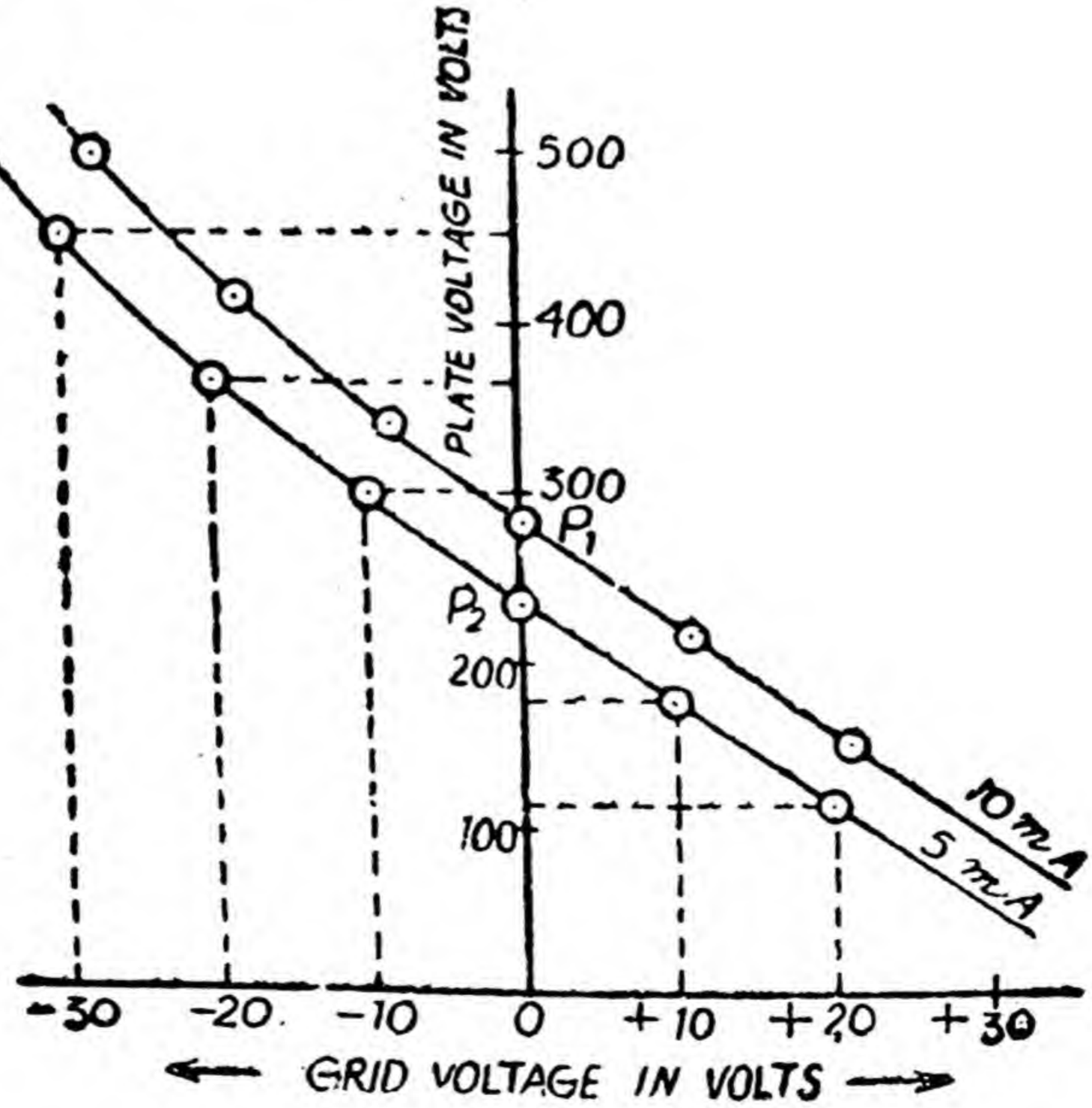


Fig. 90;

Hence plate resistance

$$r_p=\left|\frac{\Delta e_o}{\Delta i_b}\right|_{e_b}$$

$$=\frac{4 \text{ volts}}{\text{mA.}}$$

$$=8000\Omega.$$

Considering points P_1 and P_3,

$$\Delta e_c=0-(-10)=10 \text{ volts},$$

$$\Delta e_b=340-260=80 \text{ volts}.$$

Hence $\mu=\left|\frac{\Delta e_b}{\Delta e_c}\right|_{i_b}=\frac{80}{10}=8.$

$$\therefore \ \mu=8.$$

10. *An electron is travelling with a velocity of $2{\cdot}25\times10^6$ m/sec. Find whether it possesses sufficient energy to*

(*i*) *ionize a normal atom of neon,*
(*ii*) *ionize a normal atom of nitrogen,*
(*iii*) *excite an atom of helium.*

The ionization energies for neon and nitrogen are 21·5 and 14·5 ev. respectively and excitation energy for helium is 19·7 ev.

We know that the energy of the electron $=\frac{1}{2}mv^2$

$=\frac{1}{2}\times9{\cdot}1\times10^{-31}\times(2{\cdot}25)^2\times10^{12}=28{\cdot}45\times10^{-19}$ joules

$=\frac{28{\cdot}45\times10^{-19}}{1{\cdot}6\times10^{-19}}=$ **17·8 ev.**

Therefore,

(i) neon will not be ionized, (ii) nitrogen will be ionized, (iii) helium will not be excited.

11. *A sodium atom is excited and electron from the outermost orbit goes to next higher permitted orbit by absorbing 2·09 ev of energy. Calculate the frequency and the wavelength of radiation when the atom comes back to the normal state.*

We know that $f=\frac{\Delta W}{h}=\frac{2{\cdot}09 \text{ e.v.}}{6{\cdot}624\times10^{-34}}=\frac{2{\cdot}09\times1{\cdot}6\times10^{-19}}{6{\cdot}24\times10^{-34}}$

$$=\frac{3{\cdot}344\times10^{-19} \text{ joules}}{6{\cdot}624\times10^{-34}}$$

$$=5{\cdot}05\times10^{14} \text{ Hz}$$

and $\lambda=\frac{c}{f}=\frac{3\times10^8}{5{\cdot}05\times10^{14}}=5{\cdot}94\times10^{-7}$ metres

$$=5{,}940 \text{ Å}.$$

12. *The emitter of a photo-tube is illuminated by light of wavelength* 4800 Å. *The emitter surface has a threshold wavelength of* 6400 Å. *Determine the maximum velocity of electrons emitted from the photo emitter.*

We know that $h\nu = h\nu_t + \frac{1}{2}m_e v^2$
where ν_t = Threshold frequency
or $\frac{1}{2}m_e v^2 = h\nu - h\nu_t = h(\nu - \nu_t)$

$$= 6{\cdot}62 \times 10^{-34}\left(\frac{c}{\lambda} - \frac{c}{\lambda_t}\right)$$

$$= 6{\cdot}62 \times 10^{-34}\left(\frac{3\times10^8}{4800} - \frac{3\times10^8}{6400}\right)\frac{1}{10^{-10}}$$

$$= \frac{19{\cdot}86\times10^{-18}}{192} \qquad [1\text{Å} = 10^{-10}\ m]$$

Solving we have $v = 477\times10^3$ **m/sec.**

13. *The light flux on the surface of photocathode is* 1/4800 *lumen. The sensitivity of cathode is* 15 μA *per lumen. The multiplier stages of the photo multiplier are* 8, *load resistance is* 5 *meghom and output voltage is* 100 *volts. Find the multiplication per stage.*

The emission current $= \frac{1}{4800}\times 15 = \frac{1}{320}\ \mu A.$

Let the multiplication per stage $= x$

The output current $= \frac{1}{320}\times x^8\ \mu A.$

Output voltage $= \frac{x^8}{320\times10^6}\times 5\times10^6$ volts $= 100$

$$x^8 = 64\times100$$

or $$x = 2{\cdot}99.$$

EXERCISES AND PROBLEMS

1. Sketch the construction and explain the operation of typical diode valve. Discuss the meaning of 'space charge' and 'saturation current' as applied to a thermionic diode.
2. Explain the constructional feature of a triode. Describe how its characteristics may be experimentally studied and show how the constants may be evaluated from the curve obtained.
3. Describe an experimental arrangement for obtaining the characteristic curves of a triode valve. How are the amplification factor, mutual conductance and internal resistance calculated from these curves ? Show that

$$\mu = r_p g_m.$$

What do these constants signify ?

4. What are the principal advantage of tetrode valve as compared to triode valve ?
5. Trace the development of screen grid and pentode vacuum tubes as amplifying valves, giving reason for the introduction of the auxiliary grids, and explanation of their functions.
6. Explain the construction and use of a variable mu tube.
7. Draw the plate characteristic curves of a tetrode and indicate the region of negative resistance. Explain why the plate current reduces with the increase of plate voltage in this region. How is this phenomenon eliminated by the use of a suppressor grid at low potential ?

8. **Discuss how secondary emission effects are eiiminated in a beam power tube.** Draw the plate characteristic and discuss the advantages and limitations of a beam power tetrode and compare it with a pentode.
2. Using labelled diagram describe an experiment to determine the characteristic curve of a triode tube of plate current variation as the grid voltage is varied.
10. Why should the fixed grid bias of a tube be always greater than the maximum value of the input signal in amplifying circuit ?
11. What is meant by interelectrode capacitance ? Why is it necessary to keep these capacities at a minimum ?
12. What is meant by secondary emission ?
13. What effect does secondary emission have upon the operating characteristic of a tetrode ?
14. How does the addition of a suppressor grid reduce the effect of secondary emission ? Or
    Explain the necessity of having a suppressor grid in addition to screen grid. Point out the essential features of the pentode.
15. How do the characteristic curves for a pentode compare with those for a tetrode ?
16. Describe the construction of C.R. tube. Discuss clearly the function of each part.
17. (i) Derive an expression for the electrostatic deflection sensitivity of a cathode ray tube.
    (ii) Derive an expression for the magnetostatic deflection sensitivity of a cathode ray tube.
18. Point out the importance of 'time base' circuits in the functioning of a cathode ray oscilloscope. Discuss a simple form of the time base circuit and compare its performance with the ideally expected performance.
19. Make the correct answer :
    (*a*) In an oscilloscope, the volage to be viewed is applied to the
    (i) horizontal input terminals,
    (ii) vertical input terminals,
    (iii) vertical or horizontal input terminals, depending on its frequency.
    (iv) Synchronising amplifier terminal.
    (*b*) A pattern on the oscilloscope screen can be made to stand still by carefully advancing the
    (i) Sync-amplitude control, (ii) focus control,
    (iii) horizootal-positioning control,
    (iv) Coarse-frequency control

    [Ans : (a) (ii), (b) (i)]
20. Make the true (*T*) or false (*F*) answers in the following :
    (*a*) The effect of ionization in a gas is to increase the electric space charge.

(*b*) Photo tube is evacuated to a high degree so that ionisation is almost absent.

(*c*) The sensitivity of photo-tube as compared to gas photo tube is large.

[Ans. (a) *F*, (b) *T*, (c) *F*]

21. Make the correct answer :

(*a*) Gas phototuble is filled with.
(i) inert gas at low pressure.
(ii) inert gas at high pressure.

(*b*) In photo-conductive cell the electrical resistance
(i) decreases with increase of illumination,
(ii) increases with increase of illumination,
(iii) neither increases nor decreases with increase of illumination.

[Ans. (a) (i), (b) (i)]

22. A vacuum tube has an amplification factor of 7. If the grid voltage is changed from —12 to —3 volts, what will be the effect on the plate current ? Must the plate voltage be increased to return the plate current to the original value ? What is the actual change in plate voltage required to do this ?

23. A tube has an amplification factor 4·5. What grid voltage will produce cut off if the following plate voltages are applied : (*a*) 500 *V*. (*b*) 700 *V*. (*c*) 250 *V* ?

24. The static mutual characteristic of a triode for an anode voltage of 60*V* can be represented by a straight line from zero anode current at a voltage of —1·2 to anode current of 1·2 mA, at +36*V*. The amplification factor is 1·2. Plot the characteristic for 90*V* on the anode. Find the slope resistance and the mutual conductance.

26. The plate resistance of a 6*J*5 triode is 7·7 kilo-ohm, and the transconductance is 2·6 milli-mhos. If only the plate voltage is increased by 50 volts, what is the increase in plate current ? What change in grid voltage will now bring the plate current to its former value ?

26. The following values of anode current were obtained with a triode .

Anode voltage	25	50	75	100	volts
Anode current with grid voltage=0	0·4	2·8	6		mA.
Anode current with grid voltage=—4·4	0	0·6	3·6	5·0	mA.
Anode current with grid voltage=—8·3	0	0	1·0	2·9	mA.

Plot the anode current anode voltage characteristic curves. If a battery of 150*V* and a resistance of 15,000Ω are connected in series with the anode and cathode, what will be the anode current at the above three values of grid voltage ?

27. A cathode ray tube has final anode voltage of 1800 volts and the distance from the anode to screen is 20 cm. The intensity of the earth's field is ·3 *AT*/cm. and the angle of dip is 60°. Calculate the deflection of the spot on the screen from the axis of C.R. tube when this axis is horizontal and (*a*) in the plane of magnetic field (*b*) at right angle to the meridian.

28. A voltage of 10 volts is applied to the *Y-Y* deflection plates, placed 0·8 cm. apart. Calculate the flux density of magnetic flux required to return the spot to the centre of the screen The final anode voltage is 800 volts. The length of the magnetic field is the same as that of the electrostatic field and the magnetic field is horizontal. The distance from the screen to the centre of the deflection field is same.

29. Calculate the theoretical deflection sensitivity of C. R. tube having an accelerating voltage of 3,000*V*, in which the deflector plates are 1 cm., apart and have an effective length of 3 cm. the screen being 30 cm., distant from the mid-point of the deflections.

30. The flux density of a magnetic field in C.R. tube is 10^{-3} weber/sq. cm. for a distance of 2 cm. along the tube axis. The distance of the screen from the field is 10 cm. The final anode voltage is 1000 volts. Calculate (*a*) the deflection of electrons in centimeters. (*b*) displacement of a charged particle of charge twice that of an electron and mass 700 times as large.

Assume $m = 9 \times 10^{-31}$ kgm. and $e = 1{\cdot}6 \times 10^{-19}$ coulombs

31. What necessitated the invention of the tetrode after the triode? Compare the performance of the two kinds of the tube.

Draw plate voltage-plate current characteristics of the tetrode for screen volts and show how the tube can have negative resistance.

How can the defect of tetrode be removed by the use of a pentode?

6 SEMICONDUCTORS AND TRANSISTORS

There are a large number of materials which have resistivities lying between those of an insulator and a conductor. Such materials are known as 'semiconductors'. At absolute zero, pure and perfect crystals of the semiconductors are nonconducting, their resistivity approaching to the resistivity of an insulator. They can be made conducting by adding impurities, and due to thermal agitation, lattice defects etc. Resistivity of a semiconductor depends upon the temperature, and it decreases with the rise in temperature; consequently, a semiconductor crystal becomes conducting even at room temperature. At room temperature, their resistivity lies in the range 10^2 to 10^9 ohm. cm. and is thus intermediate between the resistivity of good conductor (10^{-6} ohm. cm.) and insulator (10^{14} to 10^{22} ohm. cm.)

From the study of their electrial properties (say in case of germanium) it is noted that ;

(*i*) Semiconductors have negative temperature coefficient of resistance *i.e.*, their resistivity decreases or conductivity increases rapidly with the increase of temperature. In other words the number of current carrier increases rapidly with the temperature, especially at higher temperatures. Conductors. quite contrary to the above, have positive temperature coefficient of resistance.

(*ii*) Their electrical conductivity is very much affected by even a very minute amount of impurity added to it.

Based on semiconductors is an important device in electronics called transistor, which is basically used in amplifying electrical impulses, applied at its input terminals.

The merits of transistors over vacuum tubes are :

(*i*) Smaller in size, (*ii*) lighter in weight, (*iii*) greater mechanical strength, (*iv*) almost unlimited life time (the life time of vacuum tube depends upon the life of cathode which deteriorates with use), (*v*) freedom from microphonic noise, (*vi*) no warm up time (as no filament is required), (*vii*) small power dissipation (operate with relatively low voltages), (*viii*) higher efficiency.

Following are the demerits of transistors as compared to vacuum tubes :

(*i*) Higher noise level (except the microphonic), (*ii*) highly sensitive to temperature (transistor characteristics vary markedly with temperature), (*iii*) relatively inferior high frequency response.

(*iv*) relatively low output power (maximum output power is 100W, while in case of tube it is 300 kW.)

In order to study the action of transistor, we shall first explain the conduction of electricity by pure and impure semiconductors like germanium and silicon.

6·1. CHEMICAL BONDS IN SEMICONDUCTORS LIKE GERMANIUM AND SILICON :

A pure germanium or silicon is known as intrinsic semiconductor. A pure germanium atom has 32 electrons. Out of these 32 electrons, 28 are tightly bound to the nucleus, whereas the remaining four revolve in the outermost level and are called as valence electrons. Nucleus with 28 tightly bound electrons forms the positive 'core' of the atom.

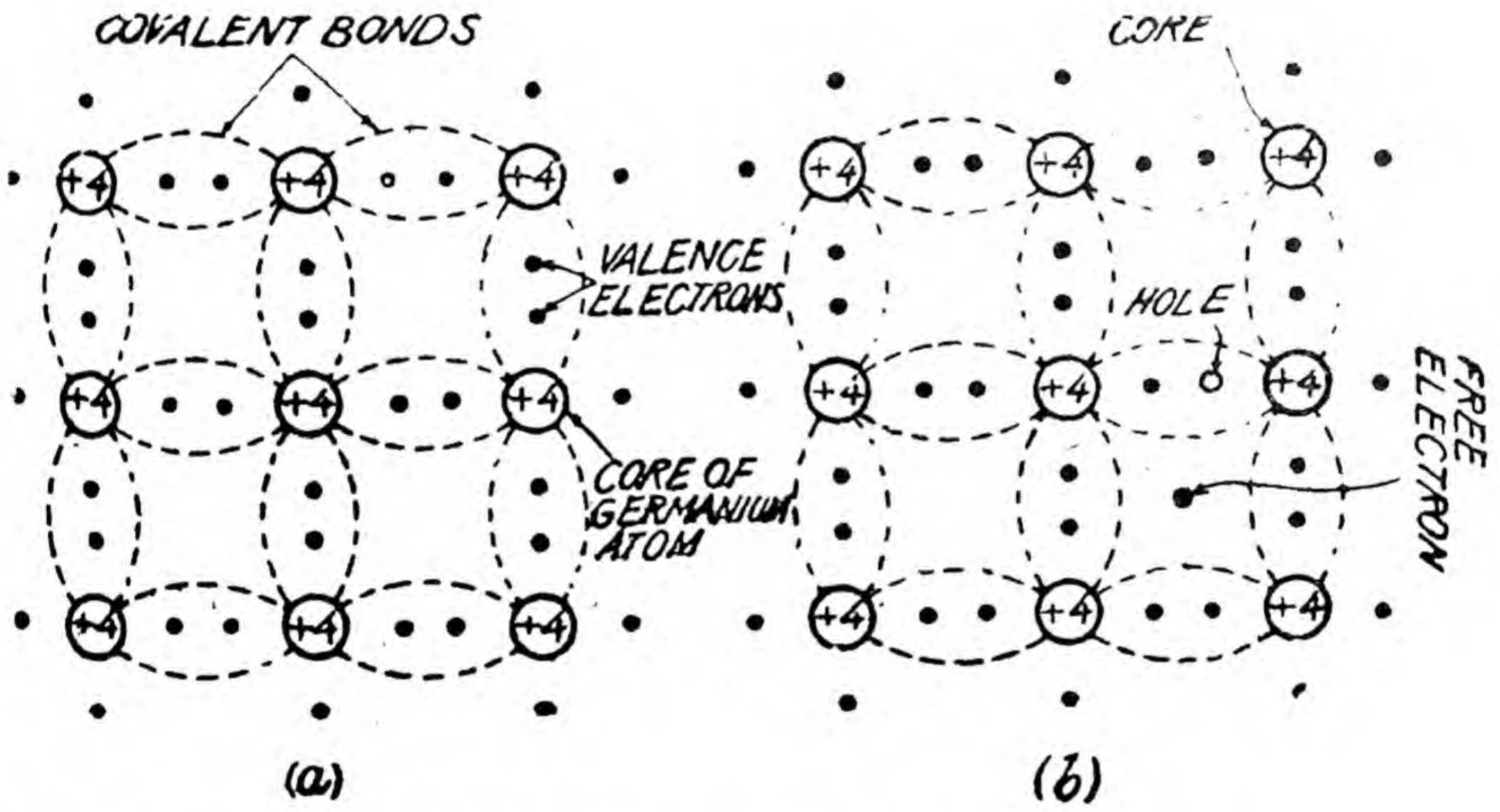


Fig. 1. (a) Germanium crystal structure showing covalent bond.
(*b*) Germanium crystal structure with an electron hole pair.

When the two atoms of germanium are brought close to each other then positive core of one atom interacts with one of the valance electrons of the other atom. Each core will attract the electron of the other atom and the two electron will be shared by two atoms. Two such electrons form an 'electron pair' as shown in figure 1(*a*). Equilibrium state will reach when the attractive force is balanced by the repulsive force between two positive cores and covalent bond is thus formed. These covalent bonds are the binding forces between the adjacent atoms of the crystal and the two electrons contributed by the two atoms still not available for conduction. It means electron pair is bound to the cores of the atoms. For their disruption, considerable amount of energy is required. Because of these covalent bonds, a crystal of pure germanium at 0^0K is a non-conductor of electricity as there are no free mobile carriers of charge

At ordinary temperature, because of thermal energy, crystal lattice is in continuous random agitation. As a result, an individual electron of a covalent bond acquires sufficient energy even at room temperature to break the bond and becomes free. These escaped electrons which are neither repelled nor attracted by the core of the atom move freely about the crystal in a random way like gas molecules. But when an electric field is applied, this random motion experiences a drift towards the positive electrode representing the flow of current.

When electron escapes from covalent bond to become free then empty space left behind is called a *hole*. The tendency of one remaining electron in covalent bond is to form again an electron pair for which it steals one electron from an adjacent pair due to favourable situation arising out of thermal agitation. Thus the previous hole is now shifted to another place from which electron is stolen. Thus when a hole is once created, it moves about in the crystal in a random way in the same manner as do free electrons. In the presence of an external electric field, a steady drift towards the negative electrode superimposed upon the random motion of these holes and thus represents a circuit flow which is transported by the electrons.

A free electron moving to and fro in a crystal lattice may collide with one of the broken covalent bond and as a result of this collision it may combine with a hole. This free electron on combining with hole will destruct an electron-hole pair and a covalent bond will be formed again *i.e.* a free electron will be converted into bound electron.

The concentration of holes and electrons in a pure semiconductor rises to a level such that rate of combination of holes and electrons equal their rate of production which increases as the temperature is raised.

6·2. BAND GAP : FORBIDDEN ENERGY GAP, VALENCE AND CONDUCTION BANDS :

We know that for the transition of an electron from the inner orbit to outer orbit of the atom some energy is required; therefore it should be obvious that different orbits correspond to different levels of potential energy. The electrons in an atom try to occupy the inner orbit but according to Pauli's exclusion principle no more than two electrons having opposite spins in an atom can exist in the same energy level. In normal atom, therefore, the electrons occupy all the lower energy levels, while the higher energy levels remain unoccupied.

When two similar atoms are brought close together, there is an interaction or coupling between the orbits of their electrons that causes splitting of each individual energy level into two slightly different levels. In a solid, there is a large number of atoms which are very close together; therefore the energy levels

produced after splitting due to interaction or coupling between the various orbits of different electrons will be so numerous and so close together as to form a band that it can be regarded as essentially continuous. If an energy band consists of as much electrons as permitted by Pauli's Exclusion Principle, then it is said to be completely filled band and thus in such a band there will be no free electrons for conduction of electricity, while for a partially filled band conduction is possible.

In fig. 2 theoretical curve of the energy band of the diamond lattice as a function of atomic spacing is drawn. When atomic spacing is sufficiently wide, interaction between atoms is negligible and energy levels are those to be found in isolated atoms shown at position *A* in the figure 2). When atoms are brought further close to each other, their interaction increases and the two energy levels split into bands (position *B*). At position *C* in the figure 2, the two bands overlap and when atomic spacing is further reduced (at point *D* in fig. 2), the bands divide and spread widely. The actual lattice constant of the diamond crysta the spacing which atoms assume) is indicated at *D* where the two distinct and widely separated bands exist. The lower band is completely filled called valence band, whereas the upper band is normally empty called conduction band.

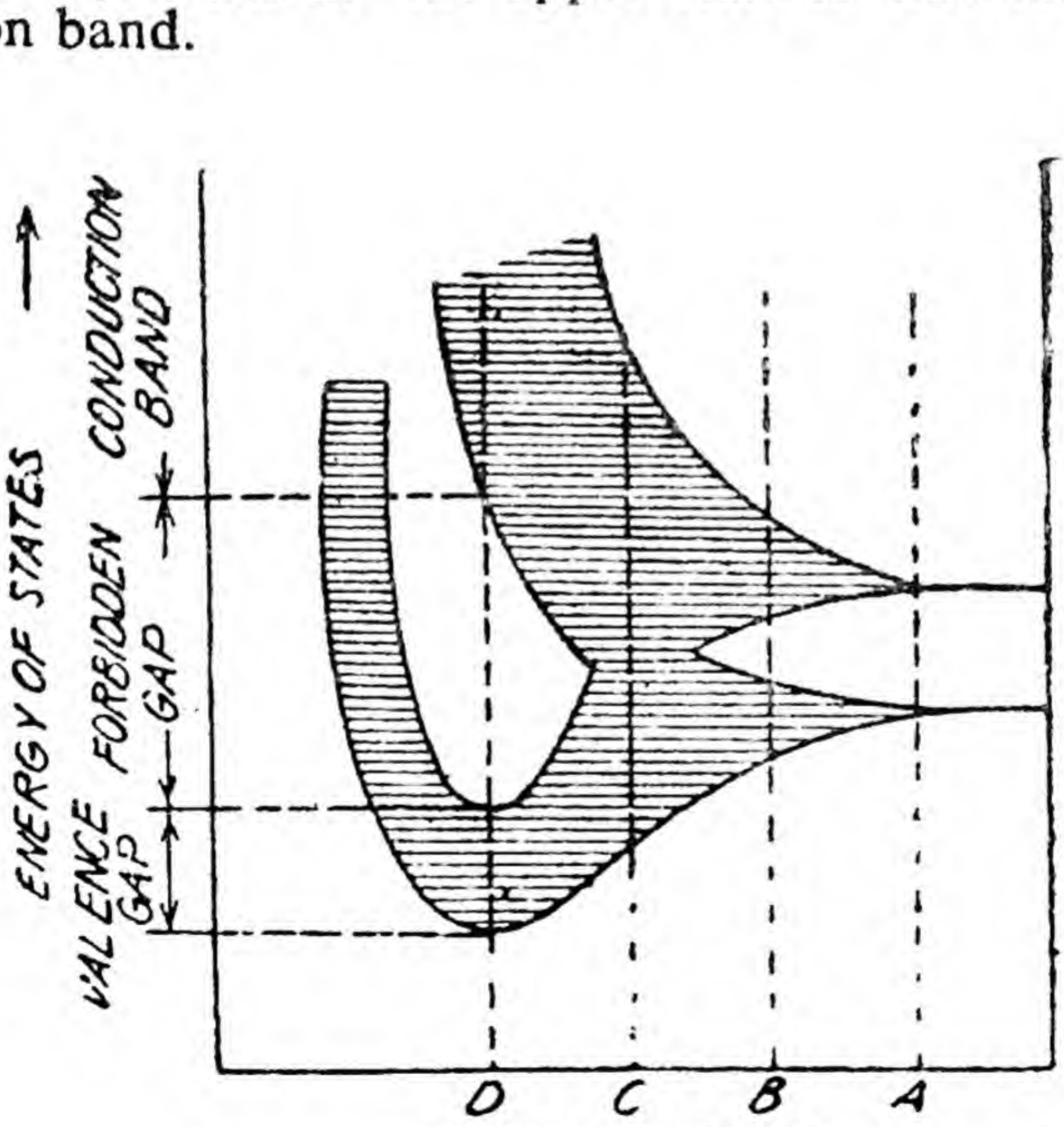


Fig. 2. Energy band of diamond lattice as a function of atomic spacing.

In an insulator and pure semi-conductor (at °K) lower band is completely filled while the upper band is co npletely em oty. The gap between these two bands is called forbidden band and is a measure of energy E_g. Thus E_g is the amount of energy that should be imparted to the electron in valence band to jump to

conduction band. Electron thus jumped to conduction band leaves its vacancy in valence band and is called a hole. This hole freely moves in the band.

In insulator E_g is of the order of 5 to 10 volts and it is practically impossible to impart this much amount of energy to the electrons in valence band so as to jump to conduction band. Consequently, conduction is not possible in the case of insulators. On the other hand E_g is of the order of 1 or 0·75 volt in the case of semi-conductor and this amount of energy can be imparted by thermal agitation of the crystal lattice. With the rise in temperature, more electrons from valence band jump to conduction band and this result in the increase of electron hole pairs. *The band gap E_g is the energy required to break an electron out of the covalent bonds so as to make it free for conduction.*

In an intrinsic semiconductor, the number of electrons going to conduction band as a result of imparted energy through thermal agitation of crystal lattice leave equal number of vacancies in the valence band is shown in fig. 3 (*d*). *Thus number of holes in valence band equals the number of electrons in conduction band.* It is also obvious from fig. 3 (*d*) that all valence electrons have an energy equal to or less than E_v, while all conduction band electrons have energy equal to or greater than E_c.

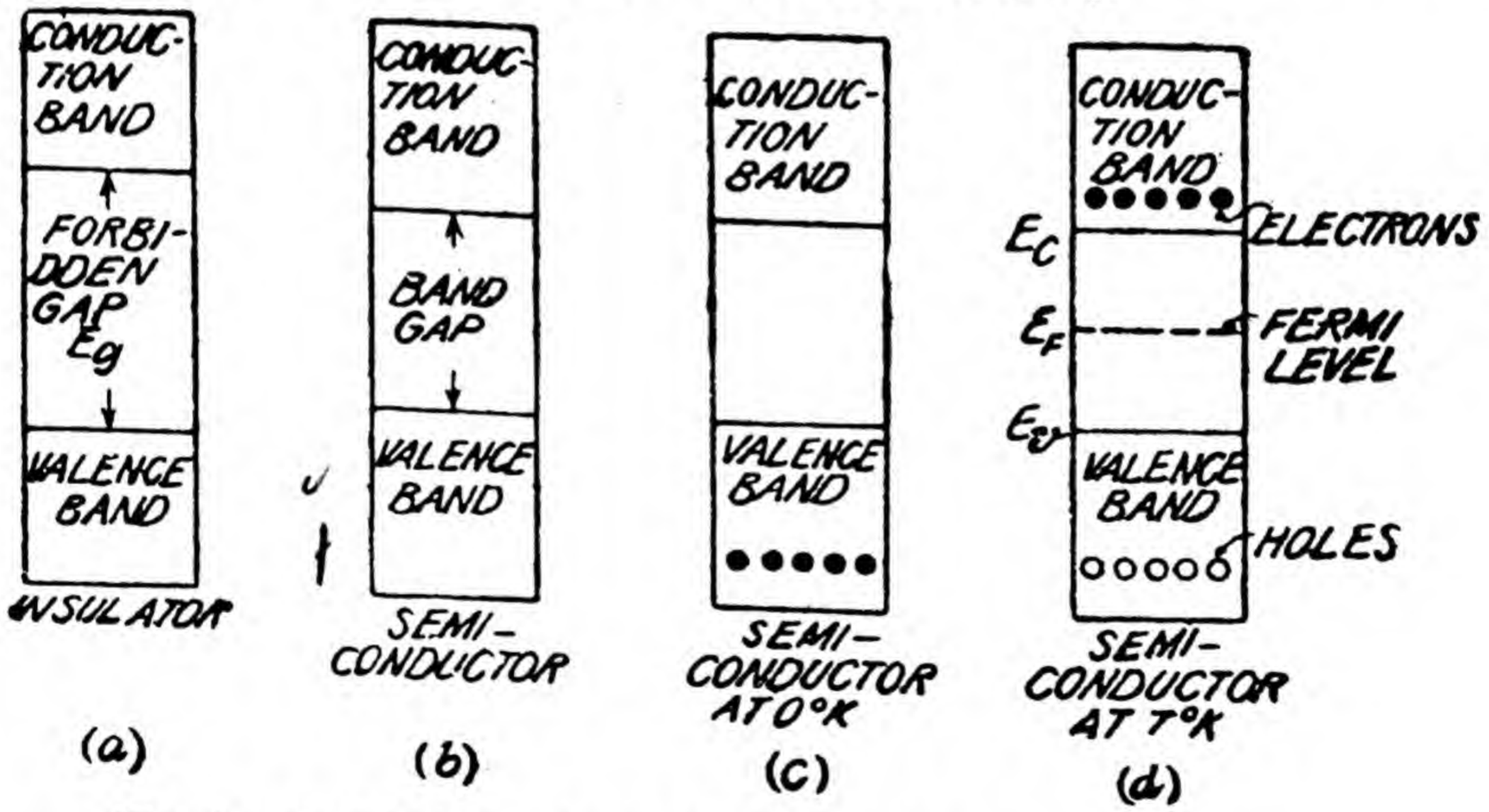


Fig. 3. Forbidden valence and conduction bands in an insulator and semi-conductor.

Now we introduce here the concept of Fermi level which is the reference energy level from which all other energies are conveniently measured. The probabiiity $F(E)$ of a state corresponding to energy E being occupied by an electron at temperature T^0K is given by

$$F(E) = \frac{1}{1+\exp.\left(\frac{E-F_F}{K_B T}\right)}$$

where K_B is Boltzmann constant.

At $T=0°K$, if $E > E_F$ then $F(E)=0$, *i.e.* energy state is empty,

$E < E_F$ $F(E)=1$ energy state is occupied by an electron,

At $T\neq 0°K$, and $E=F_F$ $F(E)=\frac{1}{2}$ energy state is 50% occupied.

Therefore, we find that at $T=0°K$, all the energy state above E_F are empty, whereas all those below E_F are filled with electrons. With the rise in temperature, states above E_F no more remain empty but they are then occupied by the electrons to some extent. In other words, when the temperature is raised there is a greater possibility of electrons being found above the Fermi level with an equal possibility of finding an electron vacancy (hole) below Fermi level. Therefore, for an instrinsic semiconductor in which number of electrons in conduction band equals the number of holes in valence band, Fermi level should be situated in the middle of the forbidden gap (fig. 3*d*) between conduction and valence band; and the position (of Fermi level) is independent of temperature.

6·3. PURE OR INTRINSIC SEMICONDUCTOR AND IMPURITY OR EXTRINSIC SEMICONDUCTOR.

Now we proceed to classify the semiconductors :

(i) Pure or intrinsic semiconductor :—

Germanium and silicon, in the pure form are known as intrinsic semiconductors. They have crystalline structure. In metals, the atoms are very close to each other and the electron in the outer shell of an atom is as much associated with its parent atom as with its neighbour. Due to this fact, the electron has almost zero affinity with an individual atom. Under the action of applied electric field. these electrons are free to move within the body of the metal. In semiconductor, at low temperature. the electrons are strongly bound to the atoms and the conduction is not possible. When the temperature is increased, more and more electrons, become available for conduction. Germanium has four outer electrons and each of the four electrons, forms covalent bond with one valence electron of the adjacent atom. The covalent bond provides force which is responsible to bind the adjacent atoms Thus every electron is tightly bound and an appreciable energy (0·76 eV) is required to release an electron from its bond. The energy may be provided by incidence of photons in the visible range or by bombardment of semiconductor sample by α particles, electrons and other particles. Once a covalent bond is broken, the electron is released and moves through the lattice in same manner as electron moves in a metal. When an electron is removed from a covolent bond, an empty space is left behind which is known as a hole and has a unit positive charge. The holes move

through the crystal lattice in a random fashion as liberated electrons and contribute to current when an electric field is applied. A semiconductor in which electrons and holes are created solely by thermal excitation across the energy gap is called an intrinsic semiconductor. The electrons reaching conduction band due to thermal excitation leave equal number of vacancies or holes in valence band *i.e.* in intrinsic semiconductor concentration of holes and electrons must always be the same.

(ii) Impurity or Extrinsic Semiconductors :—

The intrinsic semiconductors are of little importance because of negligible conductivity and moreover the conductivity has little flexibility. These limitations are overcome by adding a small and measured amount of chemical impurity to the intrinsic semiconductor. The resulting semiconductor is known as extrinsic semiconductor. The added impurity is very small of the order of 1 atom per million atoms of the pure semiconductor. Thus when impurity atoms are incorporated in a pure semiconductor crystal, the electric properties are altered and most of charge carriers originate from impurity atoms ; such an impurity semiconductor is known as extrinsic semiconductor.

6·4. MOBILITY, DRIFT VELOCITY AND CONDUCTIVITY OF INTRINSIC SEMICONDUCTOR :

Due to thermal agitation and lattice vibration, some of the covalent bonds in semiconductor are broken and some electrons become free. When an electric field is applied, the electrons drift towards the positive electrode and hence constitute the current. Similarly, the absence of an electron from a band known as hole constitutes a net positive charge. This hole moves towards the negative electrode constituting the current.

Thus due to the application of potential difference between both ends of a crystal, charge carriers attain a drift velocity v_d which is proportional to the field strength F and is different for different semiconductor materials and for different type of charge carriers. Therefore,

$$v_d \propto F$$
$$= \mu F,$$

where μ is called mobility of a charge carrier defined as the velocity of a charge carrier produced due to unit field strength.

It should be noted here that in a pure semiconductor under the influence of an electric field electrons move in a nearly empty conduction band, whereas holes move in a nearly filled band*.

*Properties of an electron in a crystal is different from those of a free electron. Forces of attraction and repulsion act upon an electron because of the close proximity of the atomic nuclei and of other electrons.

Therefore, an electron in conduction band is subjected to different conditions compared to a hole in valence band. Thus properties assigned to a hole *e.g.* mass mobility etc. will not be the same as those of an electron in the conduction band.

Since the type of drift of electrons and holes under the action of an applied force is different, μ, the electron mobility, at any temperature will be different (greater) than the mobility, μ_p, of holes. In table 1, electron and hole mobilities at 330°K and at any temperature $T°K$ are listed.

TABLE-1

ELECTRON AND HOLE MOBILITIES

Material	Resistivity in ohm cm. at 290°K	Electron Mobility μ_n		Hole Mobility μ_p	
		300°K	at $T°K$	at 300°K	at $T°K$
Silicon	$3 \cdot 10^5$	1350	$2 \cdot 1 \times 10^9 T^{-2 \cdot 5}$	480	$2 \cdot 3 \times 10^9 T^{-2 \cdot 7}$
Germanium	48	3900	$4 \cdot 9 \times 10^7 T^{-1 \cdot 65}$	1900	$1 \cdot 05 \times 10^9 T^{-2 \cdot 3}$

If n is the density of the electrons in conduction band of semiconductor, then net charge per unit volume available for the conduction of electric current will be

$$= ne$$

The drift current density J_n due to electrons is defined as the charge flowing across unit area of cross-section per unit time due to their drift under the influence of applied electric field F and is given by

$$J_n = ne v_{dn}, \qquad \ldots(1)$$

where v_{dn} is the drift velocity of electrons and is expressed as

$$v_{dn} = \mu_n F. \qquad \ldots(2)$$

If σ_n is the conductivity of semiconductor material due to free electrons, then current density J_n due to drift electrons is also given by

$$J_n = \sigma_n F. \qquad \ldots(3)$$

From equations (1) and (3),

$$\sigma_n = \frac{ne\, v_{dn}}{F}$$

$$= \frac{ne\, \mu_n F}{F} \text{ (from eq 2)}$$

$$= ne\mu_n. \qquad \ldots(4)$$

Similarly, in valence band, if p be the density of holes, then σ_p, the conductivity due to holes is

$$\sigma_p = pe\mu_p$$

where μ_p is the mobility of holes.

Therefore, total conductivity is

$$\sigma = \sigma_p + \sigma_n$$
$$= e(p\mu_p + n\mu_n).$$

For intrinsic semiconductor density of holes in valence band is equal the density of electrons in conduction band, *i.e.*

$$n = p = n_i,$$

so that

$$\sigma = en_i\ (\mu_n + \mu_p).$$

where n_i is the number of electrons or holes per unit volume in the material and is given by

$$n_i = \int Z(E)\ F(E)\ dE,$$

the limit of integration is from the bottom of lowest energy band to the top of the highest band concerned. For free electrons

$$Z(E)\ dE = CE^{1/2}\ dE,$$

where

$$C = \frac{4\pi}{h^2}\ (2m)^{3/2},$$

and $Z\ E)$ is the energy density of states*.

6·5. CARRIER CONCENTRATION IN INTRINSIC SEMICONDUCTORS :

In calculating the carrier concentration in intrinsic semiconductor we shall proceed in two ways :

(1) In the first way we assume that width of conduction and valence bands are small as compared with forbidden gap so that we can take that all the conduction electrons have energy equal to E_c, whereas all valence electrons have energy equal to E_v. That is the energy of valence band can be presented by a single energy E_v and that of conduction band by E_c. This assumption is not accurate.

(2) As stated in point (1), it is not justified to take up a single value of energy for a complete band and we take up the widths of allowed energy band as comparable to forbidden gap. We take that the electrons in conduction band may have energy lying between E_c to ∞ while the electrons in valence band of energy lying from $-\infty$ to E_v.

First Way : Let each band contain Z possible state per unit volume ($Z = 10^{22}$ per cm^3). At temperature greater than 0°K the density of electrons in the conduction band is given by

---

*The number of occupied states between energy E and $E+dE$ is given by

$$N(E)\ dE = Z(E)\ F(E)\ dE.$$

Obviously, $Z(E)\ dE$ gives the total number of available states, in the range E and $E+dE$. $Z\ (E)\ F\ (E)\ dE$ is the probability of occupation

†At 0°K, valence band and all the lower bands are completely filled and conduction band is completely empty. No conduction is possible and material behaves as an insulator.

$$\left.\begin{aligned} n_c &= \frac{Z}{\exp.\{(E_c - E_F)/K_B T\}+1} \\ \text{and} \quad n_v &= \frac{Z}{\exp.\{(E_v - E_F)/K_B T\}+1} \end{aligned}\right\} \quad \ldots(1)$$

When the forbidden gap is large enough (of the order of several eV) then electrons in conduction band must have come only from valence band and we can ignore the presence of other bands below that of valence band. Then we may write

$$n_c + n_v = Z$$

$$\frac{Z}{1+\exp\{(E_c - E_F)/K_B T\}} + \frac{Z}{1+\exp\{(E_v - E_F)/K_B T\}} = Z$$

or $[1+\exp\{(E_c - E_F)/K_B T\}]+[1+\exp\{(E_v - E_F)/K_B T\}]$

$$= [1+\exp\{(E_v - E_F)/K_B T\}]\,[1+\exp\{(E_c - E_F)/K_B T\}].$$

Expanding exponential upto the second term,

$$\text{or} \quad 1+1+\left(\frac{E_c - F_F}{K_B T}\right)+1+1+\left(\frac{E_v - E_F}{K_B T}\right)$$

$$= \left[\,2+\frac{(E_c - F_F)}{K_B T}\right]\left[\,2+\frac{(E_v - E_F)}{K_B T}\right]$$

$$\text{or} \quad (E_c + E_v - E_F) = 0$$

$$\text{or} \quad E_F = \frac{E_c + E_v}{2} \quad \ldots(2)$$

Fermi level is thus located exactly in the middle of the valence and conduction bands. Its position is independent of temperature.

Suppose $(E_c - E_F) >$ a few $K_B T$ then we can neglect 1 in comparison to exponential term in the denominator of n_c and write,

$$n_c = \frac{Z}{\exp\{(E_c - E_F)/K_B T\}}$$

$$= Z \exp\left\{-\left(E_c - \frac{E_c + E_v}{2}\right)\Big/ K_B T\right\}.$$

$$= Z \exp(-E_g/2K_B T)$$

$$= 10^{22} \exp(-E_g/2K_B T), \quad \ldots(3)$$

where E_g is forbidden energy gap and is obviously equal to $(E_c - E_v)$. This expression gives the density of electrons in the conduction band.

Conductivity of material is given by

$$\sigma = n_c\, e\,(\mu_n + \mu_p)$$

$$= 10^{22} \exp(-E_g/2K_B T)\, e\,(\mu_n + \mu_p). \quad \ldots(4)$$

From eqs. (3) and (4), we note that if a graph of log n_c or log σ against $1/T$ is plotted then a straight line with a slope $-E_g/2K_B$ is obtained as shown in figure 4.

From graph we infer that conductivity of semiconductors increases with the rise in temperature.

Second Way : The derivation in first way is not accurate since in an intrinsic semiconductor neither all the electrons in valence band have energy equal to E_v, nor all the electrons in conduction band possess energy equal to E_c. Electrons in conduction band may have, energies lying from E_c to ∞ while the electrons in valence band may have energies lying from $-\infty$ to E_v.

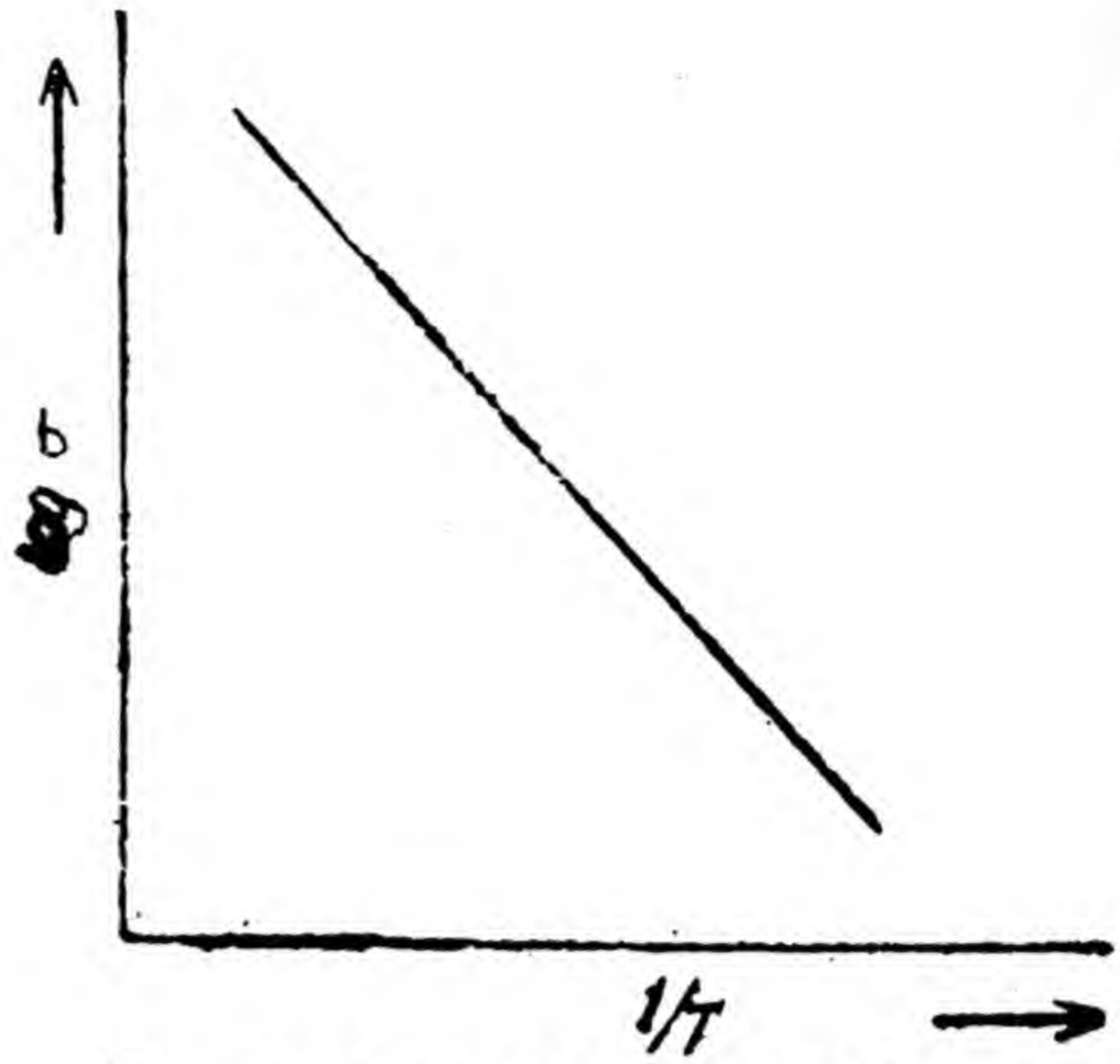


Fig. 4. Plot of conductivity as a function of temperature.

(*a*) *Density of electrons is conduction band* : Therefore the density of electrons in conduction band will be given by

$$n_e = \int_{E_c}^{\infty} Z(E)\ F(E)\ dE. \qquad \ldots(5)$$

The upper limit of the integral is put as infinity for convenience in the integration. The limit will certainly include all the electrons in the conduction band (E_c is the energy at the bottom of the conduction band). Further function $F(E)$ decreases as we move up in the conduction band because for $E >> E_g$, $F(E)=0$ and hence we can take $E=\infty$ as the upper limit for this band.

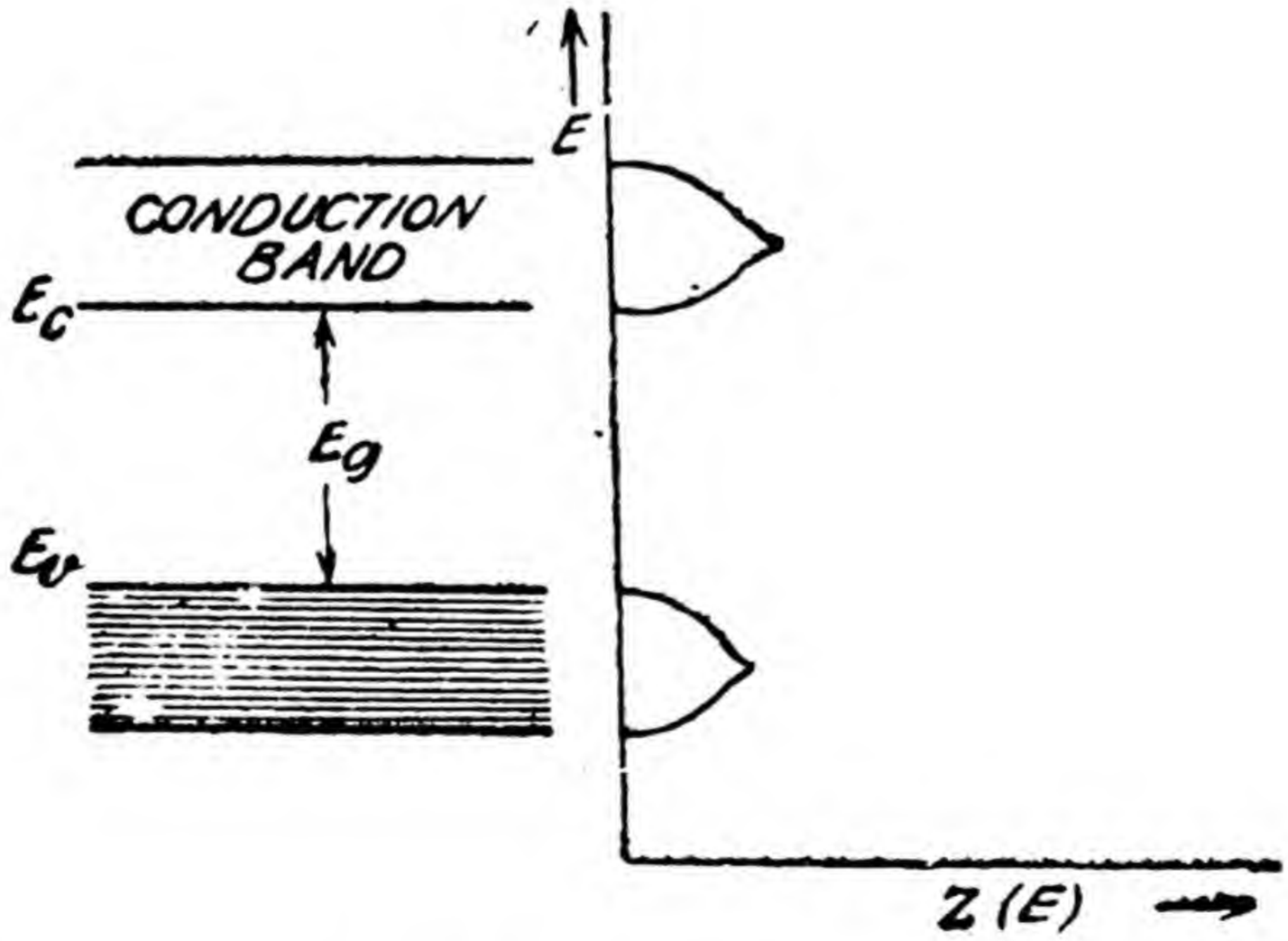


Fig. 5. Density of states in an intrinsic semiconductor, Near lower edge of conduction band $Z(E) \propto (E-E_0)^{1/2}$ and near top of valence band $Z(E) \propto (E_0-E)^{1/2}$

The energy density of states at the bottom of the conduction band of a semiconductor is given by

$$Z(E)=\frac{4\pi}{h^3}(2m_e^*)^{3/2}(E-E_c)^{1/2}, \qquad ...(6)$$

where m_e^* is the effective mass of the electron in the band. It is necessary to introduce the concept of effective mass to take into account the influence of the neighbouring atoms and charges on the motion of an electron in the band. Thus an electron will respond to an applied force as if it had a mass m_e^* (not the free electronic mass m) when it is in conduction band. In the valence band the effective mass of the hole is written as m_h^*.

$$F(E)=\frac{1}{\exp.\{(E-E_F)/K_BT\}+1}$$

and applying equation (6), we get this density of electrons in conduction band as

$$n_c=\frac{4\pi}{h^3}(2m_e^*)^{3/2}\int_{E_c}^{\infty}\frac{(E-E_c)^{1/2}\,dE}{\exp.\{(E-E_F)/K_BT\}+1}$$

If $(E_c-E_F) >> 4K_BT$ then term 1 in the denominator of the integrand can be neglected so that

$$n_0=\frac{4\pi}{h^3}(2m_e^*)^{3/2}\int_{E_c}^{\infty}(E-E_c)^{1/2}\exp\{(E_F-E)/K_BT\}\,dE$$

$$=\frac{4\pi}{h^3}(2m_e^*)^{3/2}\times\exp.\left(\frac{E_F-E_c}{K_BT}\right)\int_{E_c}^{\infty}(E-E_c)^{1/2}\exp.\left(\frac{E_c-E}{K_BT}\right)dE.$$

Putting

$$\frac{E-E_c}{K_BT}=x \text{ (giving } dE=K_BT\,dx),$$

we get

$$n_0=\frac{4\pi}{h^3}(2m_e^*)^{3/2}\times\exp.\left(\frac{E_F-E_c}{K_BT}\right)\int_0^{\infty}x^{1/2}(K_BT)^{1/2}e^{-x}K_BT\,dx$$

$$=\frac{4\pi}{h^3}(2m_e^*)^{3/2}\times(K_BT)^{3/2}\exp.\left(\frac{E_F-E_c}{K_BT}\right)\int_0^{\infty}x^{1/2}e^{-x}\,dx$$

$$=\frac{4\pi}{h^3}(2m_e^*)^{3/2}\times(K_BT)^{3/2}\exp\left(\frac{E_F-E_c}{K_BT}\right),\frac{\pi^{1/2}}{2},$$

where for integrand, $\frac{\pi^{1/2}}{2}$ has been substituted.

Thus

$$n_0=2\left(\frac{2\pi m_0^*K_BT}{h^2}\right)^{3/2}\exp\left(\frac{E_F-E_0}{K_BT}\right). \qquad ...(7)$$

(*b*) *Density of holes in valence band* : To calculate the density of holes, n_h in valence band, we shall use $[-F(E)]$ instead of $F(E)$ as this represents the probability for a state of energy E to be unoccupied.† Therefore.

†$F(E)$ is the probability for a state of energy E to be occupied. Therefore probability for a state of energy E to be unoccupied will be $[1-F(E)]$.

$$n_h = \int_{-\infty}^{E_v} Z(E)\,[1-F(E)]\,dE, \qquad \ldots (8)$$

where the lower limit has been taken $-\infty$ for conveinence and for it will include certainly all the holes in valence band. With $(E_F - E) > >$ a few K_BT,

$$1-F(E) = 1-[\exp\{(E-E_F)/K_BT\}+1]^{-1}$$
$$= 1-[1-\exp\{(E-E_F)/K_BT\}]$$
$$= \exp\{(E-E_F)/K_BT\},$$

from which we infer that since in valence band $E < F_F$, function $[1-F(E)]$ decreases exponentially. In other words, in going down below the top of valence band probability of finding holes decreases. This implies that holes reside near the top of valence band. The value of $Z(E)$ near the top of valence band is

$$Z(E) = \frac{4\pi}{h^3}(2m_h^*)^{3/2}\,(E_v-E)^{1/2},$$

where m_h^* is the effective mass of hole near the top of valence band. Therefore from eq (8),

$$n_h = \frac{4\pi}{h^3}(2m_h^*)^{3/2}\int_{-\infty}^{E_v}(E_v-E)^{1/2}\exp\left(\frac{E-E_F}{K_BT}\right)dE$$
$$= \frac{4\pi}{h^2}\cdot(2m_h^*)^{3/2}\exp\left(\frac{E_v-E_F}{K_BT}\right)\int_{-\infty}^{E_v}(E_v-E)^{1/2}\exp\left(\frac{E-E_v}{K_BT}\right)dE.$$

Let $$\frac{E_v-E}{K_BT}=x.$$

and $$dE = -xK_BT, \text{ so that}$$

$$n_h = \frac{4\pi}{h^3}(2m_h^*)^{3/2}\exp\left(\frac{E_v-E_F}{K_BT}\right)\int_{\infty}^{0}(x\,K_BT)^{1/2}\,e^{-x}\,(-K_BT)\,dx$$
$$= \frac{4\pi}{h^3}(2m_h^*)^{3/2}.\exp\left(\frac{E_v-E_F}{K_BT}\right)(K_BT)^{3/2}\int_0^{\infty}x^{1/2}\,e^{-x}\,dx$$
$$= 2\left[\frac{2\pi m_h^*\,K_BT}{h^2}\right]^{3/2}\exp\left(\frac{E_v-E_F}{K_BT}\right), \qquad \ldots (9)$$

after using $\frac{\pi^{1/2}}{2}$ for integrand.

(c) *Fermi Level* : In an intrinsic semiconductor, we know that

$$n_e = n_h.$$

After putting values of n_e and n_h from equations (7) and (9), we get

$$(m_e^*)^{3/2}\exp\left(\frac{E_F-E_c}{K_BT}\right) = (m_h^*)^{3/2}\exp\left(\frac{E_v-E_F}{K_BT}\right)$$

or

$$\exp\left(\frac{2E_F-E_c-E_v}{K_BT}\right) = \left(\frac{m_h^*}{m_e^*}\right)^{3/2}.$$

Taking logarithm,

$$\frac{2E_F - E_c - E_v}{K_B T} = \frac{3}{2} \log \left(\frac{m_h^*}{m_e^*} \right)$$

or
$$E_F = \left(\frac{E_c + E_v}{2} \right) + \frac{3}{4} K_B T \log \left(\frac{m_h^*}{m_e^*} \right). \qquad \ldots(10)$$

When $m_h^* = m_e^*$, we find

$$E_F = \frac{E_c + E_v}{2},$$

since log 1 is zero. This means Fermi level lies exactly half way between the top of valence band and bottom of conduction band. But in actual case m_h^* is greater than m_e^* and Fermi level is raised slightly as T increases.

Density of electrons in conduction band, n_e, and density of holes in valence band, n_h, interms of band gap E_g :

From equation (10), we obtain

$$(E_F - E_c) = \left(\frac{E_v - E_c}{2} \right) + \frac{3}{4} K_B T \log \left(\frac{m_h^*}{m_e^*} \right)$$

$$= -\frac{E_g}{2} + \frac{3}{4} K_B T \log \left(\frac{m_h^*}{m_e^*} \right).$$

Similarly,

$$(E_v - E_F) = \frac{(E_v - E_c)}{2} - \frac{3}{4} K_B T \log \frac{m_h^*}{m_e^*} = -\frac{E_g}{2} - \frac{3}{4} K_B T \log \frac{m_h^*}{m_e^*}.$$

Substituting these values in equations (7) and (9) respectively.

$$n_e = 2 \left(\frac{2\pi m_e^* K_B T}{h^2} \right)^{3/2} \exp \left[\frac{-E_g}{2K_B T} + \frac{3}{4} \log \frac{m_h^*}{m_e^*} \right]$$

$$= 2 \left(\frac{2\pi m_e^* K_B T}{h^2} \right)^{3/2} \exp \left(\frac{-E_g}{2K_B T} \right) \exp \left[\log \left(\frac{m_h^*}{m_e^*} \right)^{3/4} \right]$$

$$= 2 \left(\frac{2\pi m_e^* K_B T}{h^2} \right)^{3/2} \left(\frac{m_h^*}{m_e^*} \right)^{3/4} \exp \left(\frac{-E_g}{2K_B T} \right)$$

$$= 2 \left(\frac{2\pi K_B T}{h^2} \right)^{3/2} (m_h^* m_e^*)^{3/4} \exp \left(\frac{-E_g}{2K_B T} \right) \qquad \ldots(11)$$

Similarly,

$$n_h = 2 \left(\frac{2\pi m_h^* K_B T}{h^2} \right)^{3/2} \exp \left[\frac{-E_g}{2K_B T} - \frac{3}{4} \log \frac{m_h^*}{m_e^*} \right]$$

$$= 2 \left(\frac{2\pi m_h^* K_B T}{h^2} \right)^{3/2} \left(\frac{m_h^*}{m_e^*} \right)^{-3/4} \exp \left[\frac{-E_g}{2K_B T} \right]$$

$$= 2 \left(\frac{2\pi m_h^* K_B T}{h^2} \right)^{3/2} \left(\frac{m_e^*}{m_h^*} \right)^{3/4} \exp \left[\frac{-E_g}{2K_B T} \right]$$

$$= 2 \left(\frac{2\pi K_B T}{h^2} \right)^{3/2} (m_h^* m_e^*)^{3/4} \exp \left[\frac{-E_g}{2K_B T} \right] \qquad \ldots(12)$$

From equations (11) and (12), we infer that

$$n_e = n_h = n_i$$

so that conductivity of intrinsic semiconducting material is given by

$$\sigma = n_i e (\mu_n + \mu_p)$$

$$=2\left(\frac{2\pi K_B T}{h^2}\right)^{3/2}(m_h^* m_c^*)^{3/4}.\ \exp.\left[\frac{-E_g}{2K_B T}\right] e(\mu_n+\mu_p). \quad \ldots (13)$$

From equations (11), (12) and (13) we arrive at following conclusions :

(*i*) If a graph is plotted in log n_c or log n_h against $1/T$, then a straight line with slope $-E_g/2K_B$ is obtained. The graph gives the temperature dependence of carrier concentration—it rises with the rise in temperature (fig. 6).

(*ii*) In an intrinsic semiconductor density of electrons in conduction band equals the density of holes in valence band.

(*iii*) The slope of curve log σ against $1/T$ determines the value of E_g which agrees fairly well with the results of other methods.

It is to be noted that the important assumption on which these results are deduced is that the Fermi level is more than a few K_BT away from the bottom of conduction band, *i.e.*,

$(E_c - E_F) >$ a few K_BT

and from the top of valence band, *i.e.*

$(E_F - E_v) >$ a few K_BT.



Fig. 6 Graph representing the dependence of electron density on temperature.

6·6. LAW OF MASS ACTION :

From equations (11) and (12) of art. 6·5, we find that

$$n_e \times n_h = 4\left(\frac{2\pi K_B T}{h^2}\right)^3 (m_e^* m_h^*)^{3/2} \exp.\left[\frac{-E_g}{K_B T}\right].$$

We infer that at a given temperature the product of hole and electron densities is constant and is independent of the Fermi level. If n and p be the total concentrations of electrons and holes in conduction and valence bands respectively, then by using the fact that

$n_e = n =$ number of electrons in conduction band,

and $n_h = p =$ number of holes in valence band,

we write

$$n_e n_h = np = 8\left(\frac{2\pi K_B T}{h^2}\right)^3 (m_e^* m_h^*)^{3/2} \exp.\left(\frac{-E_g}{K_B T}\right)$$

$=$ a constant at a constant temperature.

Thus the product of electron and hole concentrations, for a given material, is constant at a given temperature. If impurity is added to increase n, there will be corresponding decrease in p, so that product np remains constant. This is known as law of mass action.

Therefore in an intrinsic semiconductor

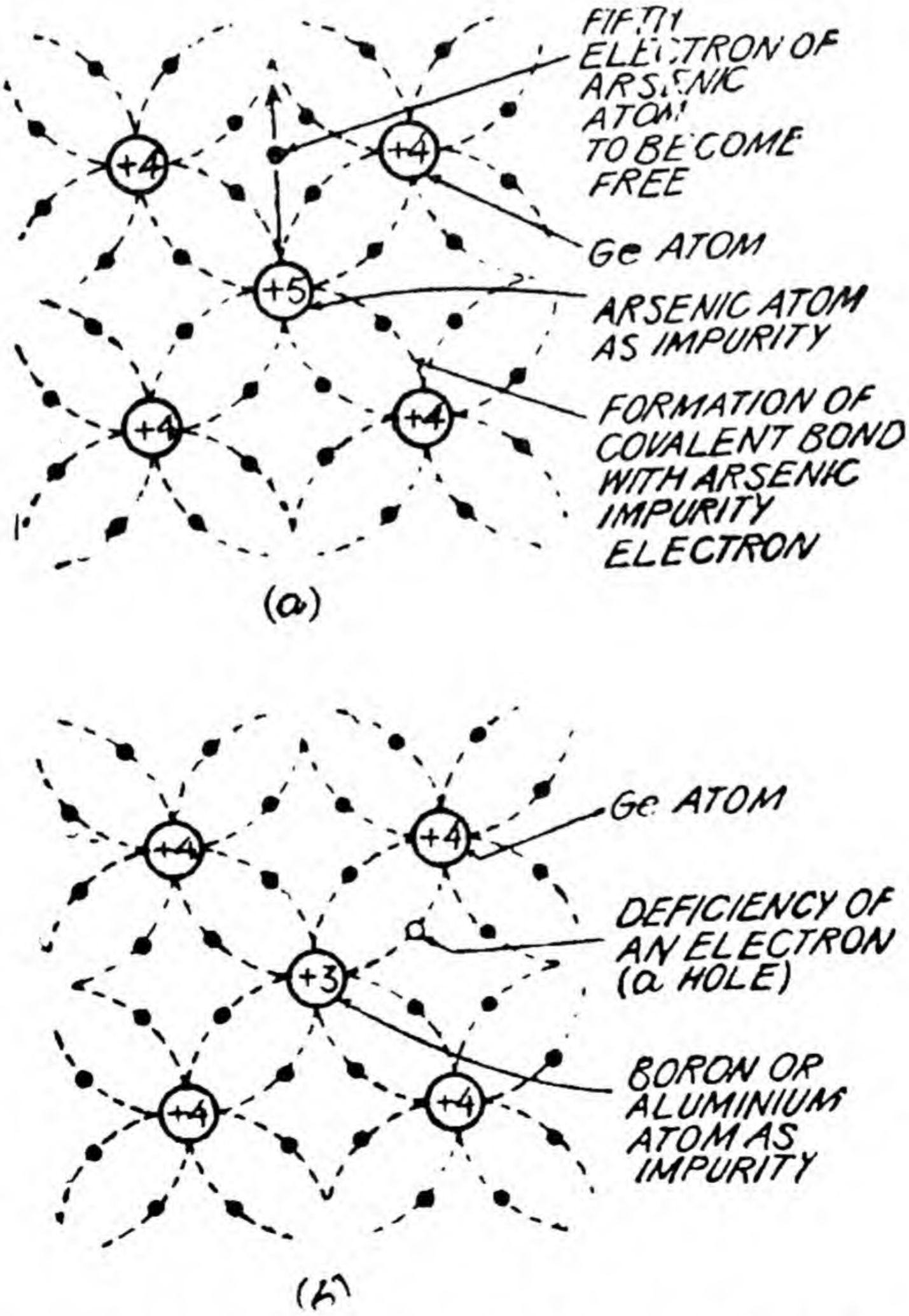
$$np = n_i p_i = n_i^2,$$

since $n = p = n_i$ for such a semiconductor. The product np remains constant for a semiconductor irrespective of its being extrinsic or intrinsic : n_i is called the intrinsic density of either carrier.

6·7. IMPURITY SEMICONDUCTORS : THERMAL IONISATION OF IMPURITIES :

The conductivity of semiconductors can be increased by adding some impurity atoms. The conductivity of the resulting crystal will depend upon the type and amount of impurity added. The added impurity may be either donor or acceptor type as discussed below for germanium crystal :

Donor Impurities. If a pentavalent impurity like arsenic or antimony is added to pure germanium crystal then resulting crystal is called *n* type or donor type.



(a) Arsenic impurity added to germanium crystal. Donor type impurity.
(b) Acceptor type impurity.

Fig. 7. n and p type semiconductor bands.

As shown in fig. 7 (*a*), four out of five valence electrons of arsenic atom form covalent bonds with the valance electrons of germanium atoms. The fifth valence electron of arsenic atom is not covalently bonded but it is not completely detached from the arsenic atom itself. A small but finite amount of energy is required to detach it from the nucleus of arsenic atom. This amount of energy is known as ionisation energy of the arsenic atom. In germanium, ionisation energy of arsenic, antimony, phosphorous atoms is found to be approximately equal to 0·01 eV. It is observed that ionisation energy decreases slightly with the increase of impurity contents. Since ionisation energy of an arsenic atom in germanium, almost all arsenic atoms will be ionised at room temperature so that fifth valence electron in each arsenic atom is detached completely and moves freely in a germanium crystal.

More are the impurity atoms, greater will be the free electrons. Thus the conductivity of pure germanium increases. Since the current carriers are negatively charged particles, this types of semiconductor, is called *n* type semiconductor. As the impurity atom donates electrons to the crystal conductivity, it is called *donor*.

Let us consider the case when *n* type semiconductor is placed between two electrodes and an electric field is applied between them. The free electrons will drift towards the positive electrode while the immobile ionized donor atoms are left in the vicinity of negative electrode. The donor atoms, now pull electrons from the metal of negative electrode and the number of pulled electrons is the same as the number of free electrons to the positive electrode. In this way the number of free electrons remains equal to the number of donor atoms and thereby the semiconductor remains electrically neutral. The voltage difference applied to the electrode appears as a voltage drop in the semiconductor. The holes do not contribute much to the current flow and hence they are called as *minority carriers* while the free electrons are known as *majority carriers*.

Acceptor Impurities : If trivalent impurity atom (like boron or aluminium) is added to a germaninm crystal then its three valence electrons form covalent bond with the valence electrons of three germanium atoms. There is a deficiency of one electrons and this deficiency is called a hole. The remaining fourth atom of the germanium also tries to form a covalent bond. As a result, it seeks to acquire fourth electron and it does so by taking the advantage of any thermal motion that fortunately brings an electron from neighbouring germanium atom into a favourable condition to be captured. Due to this stealing action, a hole is created in the adjacent atom that moves about in a random way due to thermal effects..For such an stealing action in which other electrons jump into the holes, some energy is required and is approximately equal to the ionisation energy of donor levels in the same material.

Impurity atoms that contribute holes in this manner are termed as *acceptors* because they accept electrons from germanium atoms. Since current carriers are positively charged particles (holes), this type of semiconductor is called *p*-type semiconductor.

The *p*-type semiconductors has a large number holes (majority carriers) with equal number of immobile ionized acceptor atoms; of course, the number of electrons (minority carriers) is negligibly small. Under the influence of electric field, mobile holes drift towards the negative electrode. When the holes reach the negative electrode, they pull free electrons from negative electrode and combine with them. Hence the holes are neutralized. At the same time the drifting away of holes from the vicinity of the positive electrode, leaves the immobile ionized acceptor atoms unneutralized. This results in an electric field which detaches the loosely bound electrons from the ionized acceptor atoms. The free electrons then strike the positive electrode and enter the metal of positive electrode. Every acceptor atom, after having lost one electron, steals another electron from the adjoining atom, resulting in a new hole. The hole so created drifts towards negative electrode, while the electron captured by the acceptor atom is detached and lost to the positive electrode. The semiconductor remains electrically neutral as the number of mobile holes under all conditions remain equal to the number of acceptors. The voltage difference appears as the voltage drop in the semiconductor.

6·8. IMPURITY STATES : ENERGY BAND DIAGRAM AND THE FERMI LEVEL :

In impurity semiconductor, impurity atoms, however, reduce the energy gap E_g. Donors represent isolated energy levels located so close to unfilled band (0·01 eV below the lower edge of conduction band) that very little energy is required to lift an electron from the donor level into the unfilled band (conduction band), where it is available for the conduction of electricity. Similar case

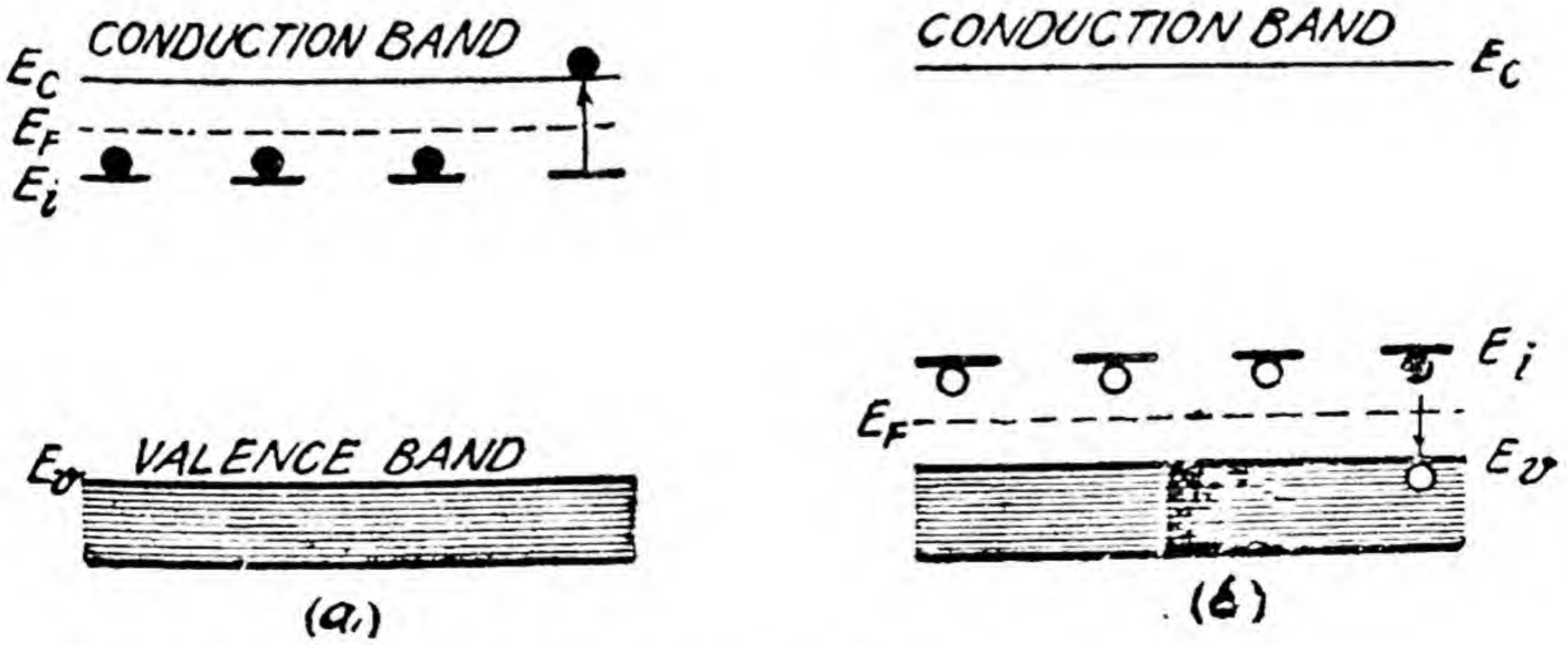


Fig. 8. (a) Indicating donor levels.
(b) Indicating acceptor levels.

is for acceptors representing isolated energy levels close to the filled band. The energy levels of the impurity atoms are shown as isolated dots, no as a band beca ıse these atoms are isolated from each other, (*i*. . there is no interaction between impurity atoms).

n-type semiconductor : Refer to fig. 8 (*a*). Below conduction band there are N_d donor levels per cm³ of energy E_i. At low temperatures, small fraction of donors will be ionised and practically donor levels will be filled with electrons. Let us assume that

$$(E_c - E_F) > 4K_BT,$$

then in that case density of electrons in conduction band will be given by equation (7) of art. 6·5, *i.e.*

$$n_c = 2\left(\frac{2\pi m_o^* K_B T}{h^2}\right)^{3/2} \exp.\ \{(E_F - E_c)/K_B T\}. \qquad \ldots(1)$$

If we assume that E_F lies more than a few K_BT above the donor level then the density of empty donors is given by

$$N_d\,[1 - F(E_i)] \approx N_d \exp\{(E_i - E_F)/K_BT\}. \qquad \ldots(2)$$

But the density of empty donors should be the same as the density of electrons in the conduction band. Then from equations (1) and (2), we get

$$2\left[\frac{2\pi m_o^* K_B T}{h^2}\right]^{3/2} \exp\{(E_F - E_c/K_BT\} = N_d \exp\{(E_i - E_F)/K_BT\}$$

Taking logarithm,

$$\frac{(E_F - E_c)}{K_BT} - \frac{(E_i - E_F)}{K_BT} = \log N_d - \log 2\left[\frac{2\pi m_o^* K_BT}{h^2}\right]^{3/2}$$

or
$$E_F = \frac{E_i + E_c}{2} + \frac{K_BT}{2} \log\left[\frac{N_d}{2\left(\frac{2\pi m_o^* K_BT}{h^2}\right)^{3/2}}\right] \qquad \ldots(3)$$

At $T = 0°K$

$$E_F = \frac{E_i + E_c}{2},$$

which shows that Fermi level lies exactly half wave between the donor levels and bottom of conduction band. As T increases, Fermi level drops. This is shown in fig. 9 for the case $(E_F - E_c) = 0{\cdot}2$ eV for three different values of N_d.

Putting the value of E_F from eqn. (3) into exp $(E_F - E_c)$, we get

$$\exp\{(E_F - E_c)/K_BT\} = \exp\left[\frac{(E_i - E_c)}{2K_BT} + \tfrac{1}{2} \log\left\{\frac{N_d}{2\left(\frac{2\pi m_o^* K_BT}{h^2}\right)^{3/2}}\right\}\right]$$

$$= \exp\left(\frac{E_i - E_c}{2K_BT}\right) \cdot \frac{(N_d)^{1/2}}{\left\{2\left(\frac{2\pi m_o^* K_BT}{h^2}\right)^{3/2}\right\}^{1/2}}.$$

which when substituted in equation (1) gives

$$n_c = 2\left[\frac{2\pi m_e^* K_B T}{h^2}\right]^{3/2} \cdot \frac{(N_d)^{1/2}}{\left[2\left(\frac{2\pi m_e^* K_B T}{h^2}\right)^{3/2}\right]^{1/2}} \cdot \exp\{(E_l - E_c)/2K_B T\}$$

$$= (2N_d)^{1/2}\left[\frac{2\pi m_e^* K_B T}{h^2}\right]^{3/4} \cdot \exp\{(E_l - E_c)/2K_B T\}.$$

$$= (2N_d)^{1/2}\left[\frac{2\pi m_e^* K_B T}{h^2}\right]^{3/4} \exp(-\Delta E/2K_B T). \quad \ldots(2)$$

where $\Delta E = (E_c - E_l)$ is called ionisation energy of donors. We note that :

(i) density of electrons in the conduction band is proportional to the square root of the donor concentration.

(ii) as the temperature increases, the Fermi level falls below the donor level and it approaches the centre of forbidden gap which makes the substance an intrinsic semiconductor. The conductivity of intrinsic semiconductor is smaller than n or p type semiconductors.

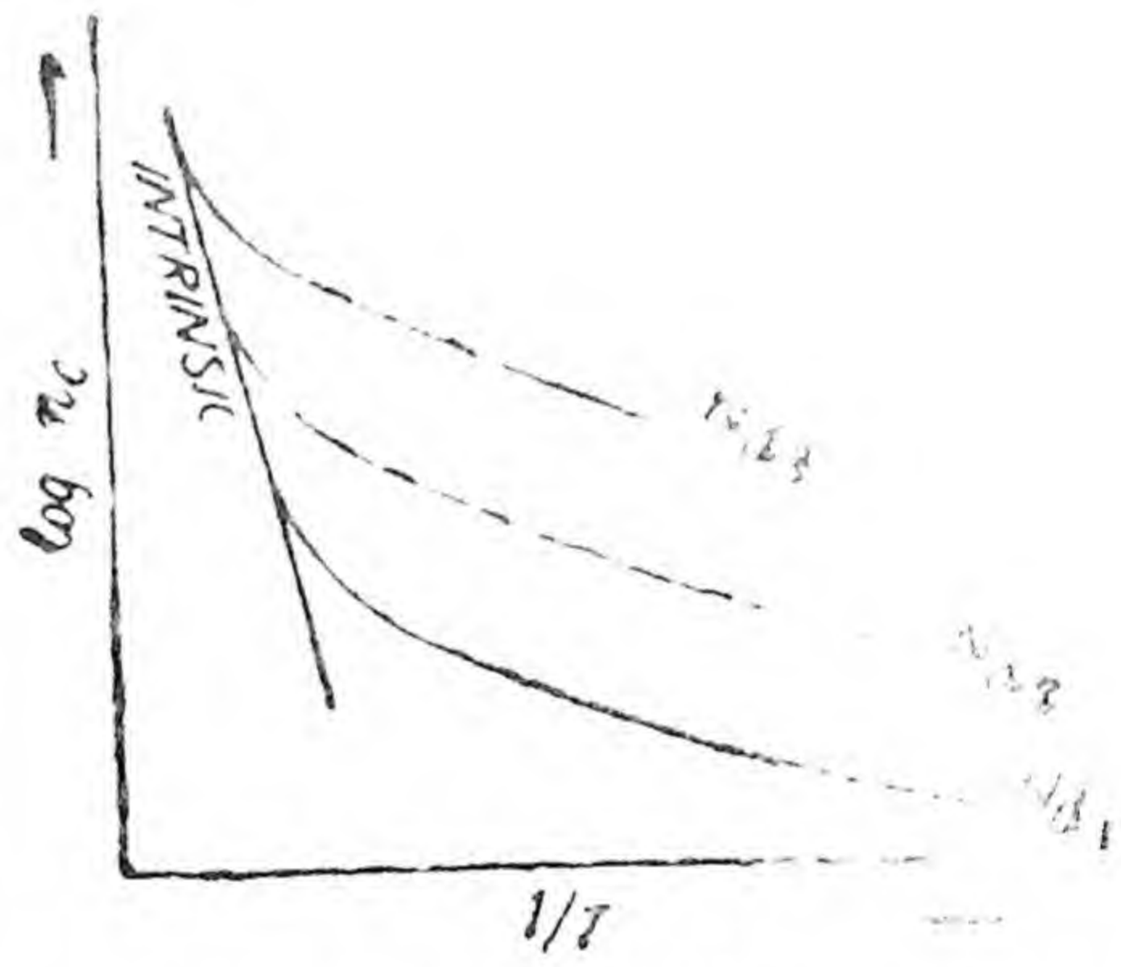


Fig. 9. Variation of conduction electron density with temperature

In fig. 9, a graph is plotted in log n_c and $1/T$ which gives a straight line with a slope $-\Delta E/2K_B$. If T becomes sufficiently high so as to excite electrons from valence band to reach directly to conduction band (intrinsic excitation) then the slope changes to $-E_g/2K_B$ as mentioned in second way of Art. 6·5.

p-type semiconductor : Refer to fig. 8 (b). In this case acceptor levels lie above the valence band The density of holes in the valence band can be obtained exactly in the same way as in the case of n_c We can show that

$$n_h = (2N_a)^{1/2}\left(\frac{2\pi m_h^* K_B T}{h^2}\right)^{3/4} \exp(-\Delta E/2K_B T).$$

where $\Delta E = E_l - E_v$.

In this case Fermi level lies half way between the acceptor levels and top of valence band at $T = 0°K$.

The above discussion holds true when density of available levels Z_l with energy E_l are occupied by density of donor electrons N_d at $T = 0°K$ as in the case of germanium and silicon containing trivalent and pentavalent impurities. But if Z_l is much greater than N_d, as happens in the case of alkali halides containing excess metal,

then only a fraction of available levels will be occupied. In this case, at any temperature T, number of filled impurity levels is equal to

$$N_d = \frac{Z_i}{\exp\{(E_i - E_F)/K_B T\} + 1}.$$

When $(E_i - E_F) >$ a few $K_B T$, then

$$N_d = Z_i \exp\{(E_F - E_i)/K_B T\}. \quad \ldots(6)$$

From equation (1),

$$n_c = 2\left(\frac{2\pi m_c^* K_B T}{h^2}\right)^{3/2} \exp\left(\frac{E_F - E_c}{K_B T}\right)$$

$$= \tfrac{1}{2}\left(\frac{2\pi m_c^* K_B T}{h^2}\right)^{3/2} \exp\left(\frac{E_F - E_i}{K_B T}\right) \exp\left(\frac{E_i - E_c}{K_B T}\right)$$

$$= 2\left(\frac{2\pi m_c^* K_B T}{h^2}\right)^{3/2} \left(\frac{N_d}{Z_i}\right) \exp\left(\frac{E_i - E_c}{K_B T}\right).$$

on using equation (6).

Further putting

$$\Delta E = (E_c - E_i)$$

we find that the density of electrons in conduction band is given by

$$n_c = 2\left(\frac{N_d}{Z_i}\right)\left(\frac{2\pi m_c^* K_B T}{h^2}\right)^{3/2} \exp\left(\frac{-\Delta E}{K_B T}\right),$$

in which exponential term consists of ΔE instead of $\Delta E/2$ as compared to equation (4). n_c is proportional to N_d.

6·9. p-n JUNCTION :

When a piece of p-type material is in contact (say, used) with n-type material then the interface between the different regions is called p-n junction. A typical width of such a junction is 10^{-4} cm. The Fermi level in p type material is located close to the top of valence band whereas in n-type material it lies to the bottom of conduction band.

The situation is shown in fig. 10 (*b*). When the contact between two types of materials is made then to bring Fermi level in a line, electrons in conduction band on n-type side (supplied by pentavalent impurity atoms) travel across the junction and leave the positively ionised impurity atoms unneutralized. Consequently, there is a positively charged region adjacent to the junction in n-type material. On p type side, the electrons which have traversed the boundary recombine with positive holes in the valence band. Near to the junction on p-type side, there is a layer of unneutralised negatively ionised trivalent impurity atoms which form a negatively charged region fig. 10 (*c*). These unneutralised impurity atoms are termed as ions. These ions being bound in the crystal lattice are immobile and no charges are available for conduction close to the junction. As shown in fig. 10 (*c*), this region around the junction is called charge depletion region or space charge region. Due to this space charge, an internal potential barrier V_B

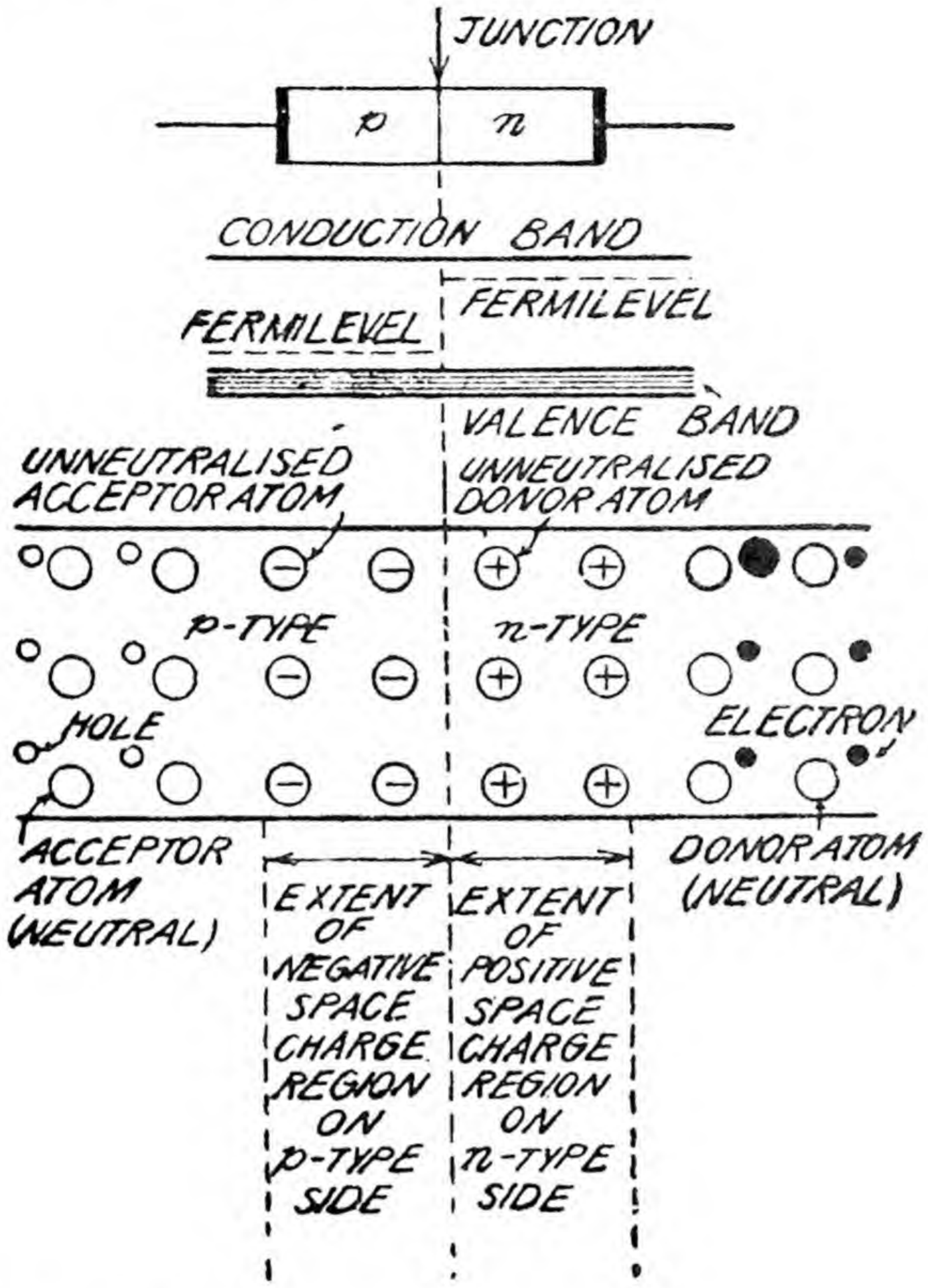


Fig. 10. (a) a *p-n* junction.
(b) non-equilibrium energy band scheme.
(c) development of space charge near junction.

exists between *n* and *p*-regions. The band edges in the two samples shifts themselves to make the alignment of Fermi level possible and the energy band diagram remains no more of the shape shown

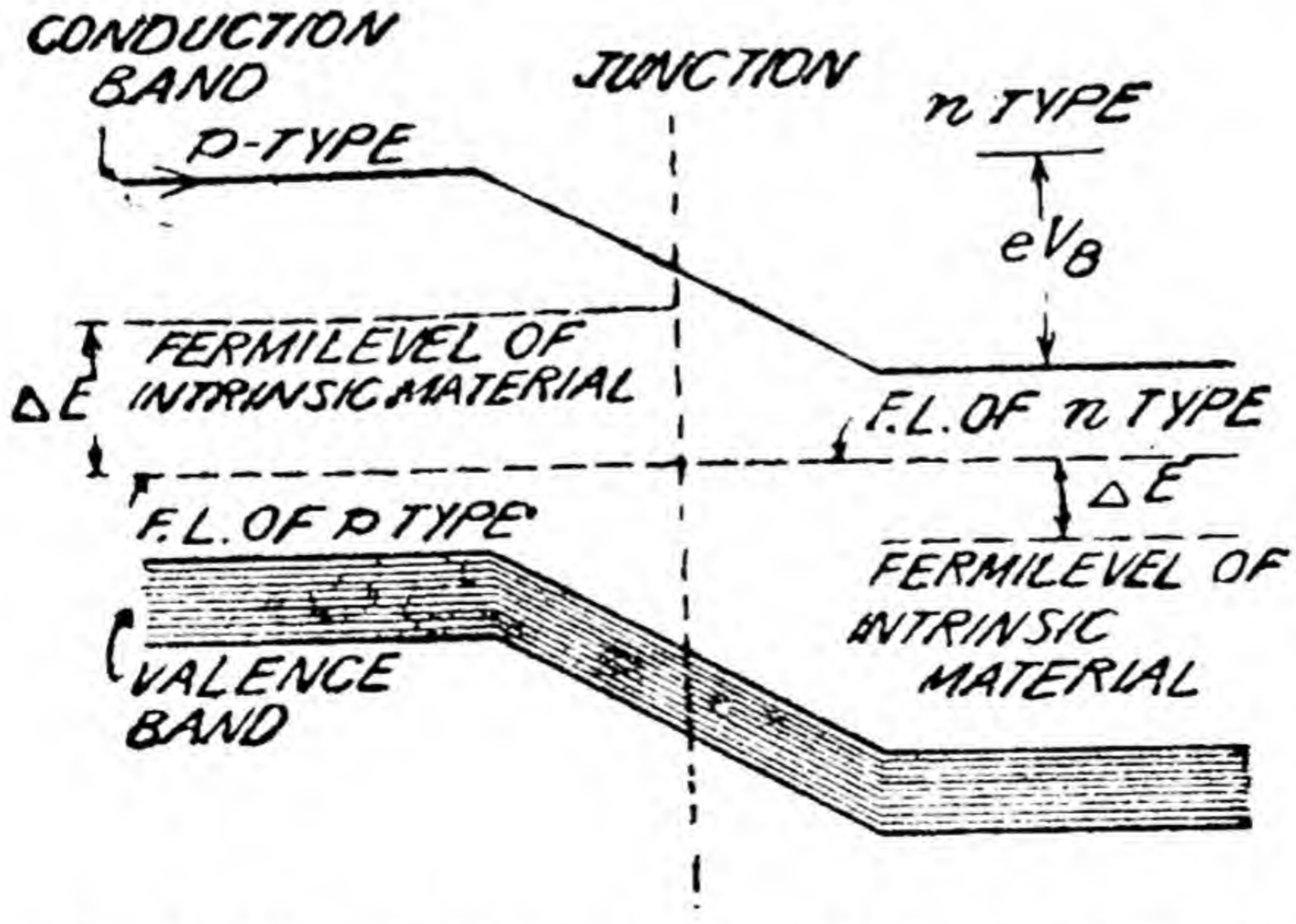


Fig. 10 (d). Shifting of the bands of the two samples.

in fig. 10 (*b*), but assumes a shape shown in fig. 10 (*d*). The conduction band of *p* type is shifted upwards by eV_B over the conduction band of *n* type where V_B is potential barrier across the junction arising due to already mentioned space charge.

Obviously, minority electrons in the conduction band of *p*-type crystal is at a higher energy than the majority electrons in the conduction band of *n*-type crystal. The electrons crossing the junction from *p* region side will not encounter the potential barrier while the electrons crossing the junction from *n*-region side will face this barrier.

6·10. BARRIER VOLTAGE OR CONTACT POTENTIAL ACROSS THE P-N JUNCTION :

As mentioned in previous article that due to the positively charged region on *n*-type side and negatively charged region on *p*-type side, there exists a space charge near the *pn* junction. The variation of potential arising out of space charge is shown in fig. 11. It is shown that barrier potential is $V_B = |V_1| + |V_2|$ and the total width of space charge region

$$X = X_1 + X_2,$$

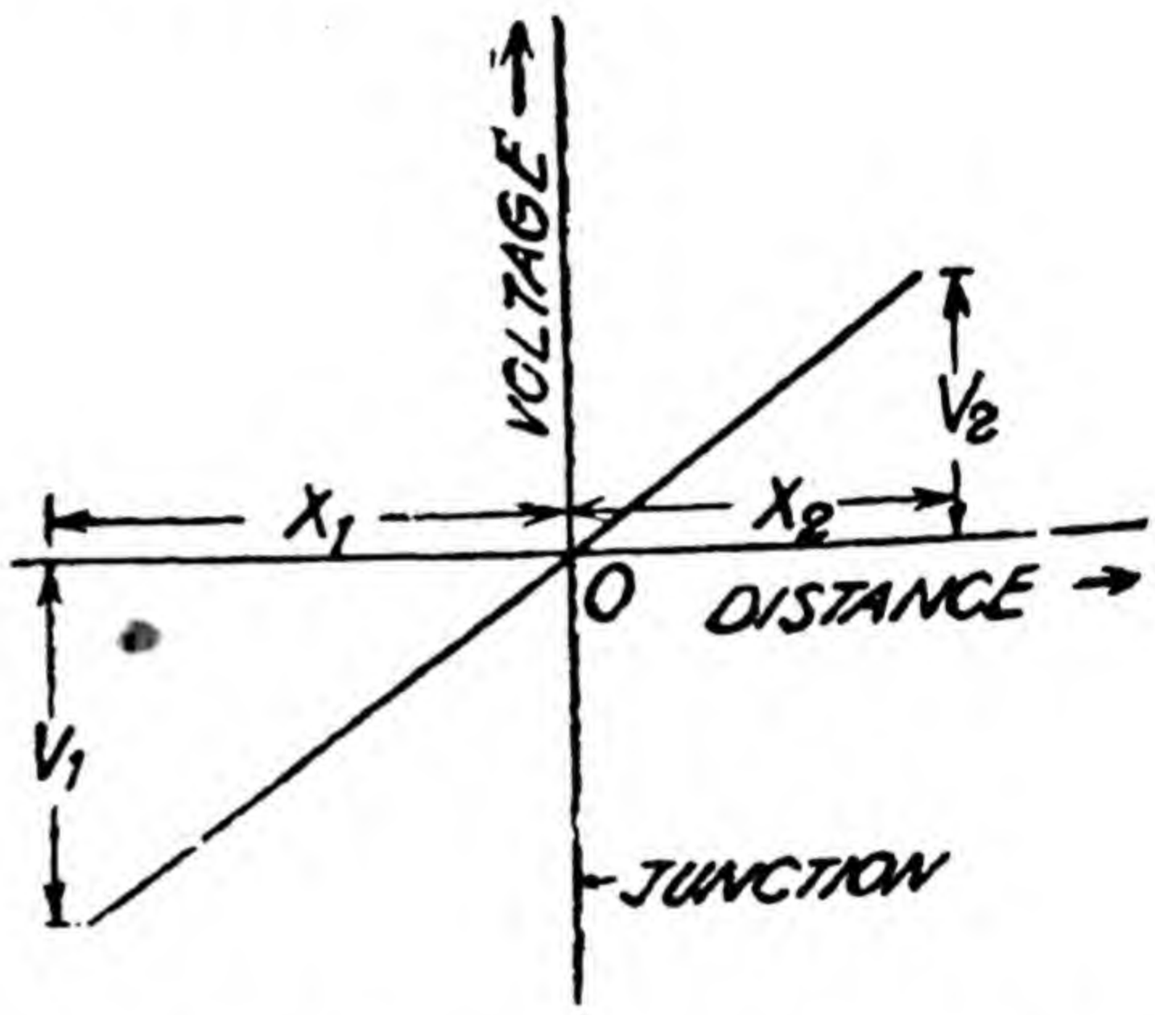


Fig. 11. Space charge region around the junction.

where X_1 and X_2 are the widths of space charge region in the *p* and *n* sides respectively.

To find the distribution of barrier potential in the space charge region, we proceed with Poisson's equation

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\epsilon}, \qquad \ldots (1)$$

in one dimension, ρ is the volume density of charge and ϵ is the permittivity of the medium.

The charge density in *p* side of the space charge region is given by

$$\rho = -eN_a$$

where N_a is the density of negatively charged (ionised) acceptor atoms. The potential variation in *p* side region is given by

$$\frac{d^2V}{dx^2} = \frac{eN_a}{\epsilon}$$

Integrating,

$$\frac{dV}{dx}=\frac{eN_ax}{\epsilon}+C_1, \qquad \ldots(2)$$

where C_1 is a constant of integration.

Applying the boundary condition :

$$\text{at } x=-X_1, \qquad \frac{dV}{dx}=0,$$

to equation (2), we arrive at

$$C_1=\frac{eN_aX_1}{\epsilon},$$

so that

$$\frac{dV}{dx}=\frac{eN_aX}{\epsilon}+\frac{eN_aX_1}{\epsilon}. \qquad \ldots(3)$$

Integrating the above equation, we find that

$$V=\frac{eN_ax^2}{2\epsilon}+\frac{eN_aX_1x}{\epsilon}+C_2, \qquad \ldots(4)$$

where C_2 is another constant of integration. Applying the boundary condition :

$$\text{at } x=0,\ V=0,$$

to equation (4), we find that

$$C_2=0,$$

so that

$$V=\frac{eN_ax^2}{2\epsilon}+\frac{eN_aX_1x}{\epsilon}, \qquad \ldots(5)$$

At $x=X_1,\ V=V_1$

giving

$$V_1=-\frac{eN_aX_1^2}{2\epsilon}. \qquad \ldots(6)$$

Similarly, for the potential distribution in space charge region on *n* side, we apply the equation

$$\frac{d^2V}{dx^2}=-\frac{eN_d}{\epsilon},$$

where N_d is the density of the positively charged (ionised) donor atoms.

The solution of above equation, on applying the boundary condition is

$$\text{at } x=0,\ V=0$$

and

$$x=X_2,\ \frac{dV}{dx}=0,$$

we get

$$V=-\frac{eN_dx^2}{2\epsilon}+\frac{eN_dX_2x}{\epsilon}.$$

But at

$$x=X_2,\ V=V_2,$$

so that $$V_2=\frac{eN_dX_2^2}{2\epsilon}. \qquad ...(7)$$

Therefore height of potential barrier across the junction is given

$$V_B=|V_1|+|V_2|$$

$$=e\left(\frac{N_aX_1^2+N_dX_2^2}{2\epsilon}\right) \qquad ...(8)$$

Since the crystal as a whole is electrically neutral, we get the charge neutrality condition

$$N_aX_1=N_dX_2 \qquad ...(9)$$

Putting for X_2 from equation (9) in equation (8), we get

$$V_B=\frac{eN_a}{2\epsilon}X_1^2\left(1+\frac{N_a}{N_d}\right), \qquad ...(10)$$

giving $$X_1=\left[\frac{2\epsilon V_B}{2N_d\left(1+\frac{N_a}{N_d}\right)}\right]^{1/2} \qquad ...(11)$$

Similarly, we can obtain

$$X_2=\left[\frac{2\epsilon V_B}{N_d\left(1+\frac{N_d}{N_a}\right)}\right]^{1/2} \qquad ...(12)$$

The space charge region has a width equal to

$$X=X_1+X_2$$

$$=\left(\frac{2\epsilon V_B}{e}\right)^{1/2}\left[\left\{\frac{N_d/N_a}{(N_a+N_d)}\right\}^{1/2}+\left\{\frac{N_a/N_d}{(N_a+N_d)}\right\}^{1/2}\right]$$

$$=\left[\frac{2\epsilon V_B}{e(N_a+N_d)}\right]^{1/2}\left[\left(\frac{N_d}{N_a}\right)^{1/2}+\left(\frac{N_a}{N_d}\right)^{1/2}\right] \qquad ...(13)$$

If $N_d >> N_a$, then we can neglect N_a in the sum (N_a+N_d) and also the term N_a/N_d can be approximated to zero. Then equation (13) gives

$$X=\left(\frac{4\epsilon V_B}{eN_d}\right)^{1/2}\left(\frac{N_d}{N_a}\right)^{1/2}$$

$$=\left(\frac{2\epsilon V_B}{eN_a}\right)^{1/2}. \qquad ...(14)$$

which shows that width of space charge region decreases as the impurity concentration increases.

Further, we find the ratio

$$\frac{V_1}{V_2}=\left(\frac{X_1}{X_2}\right)^2\left(\frac{N_a}{N_d}\right)$$

and $$\frac{X_1}{X_2}=\frac{N_d}{N_a}.$$

From above two relations we find that

$$\frac{V_1}{V_2}=\frac{N_d}{N_a} \qquad ...(15)$$

We note from equation (15) that when $N_d >> N_a$, $V_1 >> V_2$ which implies that the potential charge is confined to the region which has impurity in small amounts (lightly doped).

As mentioned in art. 6·9, the potential barrier is created due to the alignment of Fermi levels of the two sides. Refer to fig. 11, Fermi level on p side has come down by $\Delta E'$ and has gone up by ΔE on n side, thus

$$\Delta E' + \Delta E = eV_B. \qquad \ldots(16)$$

We have seen in equations (10) and (11) of art. 6·5 that electron density in an intrinsic semiconductor depends on the Fermi-level and addition of donor impurities simply increase the Fermi factor by exp $(\Delta E/K_BT)$ where ΔE is the shift in Fermi-level.

We know that density of electrons in conduction band in an intrinsic semiconductor is given by equation (7) of art. 6·5which is

$$n_i = 2\left(\frac{2\pi m_e^* K_B T}{h^2}\right)^{3/2} \exp\left(\frac{E_F - E_c}{K_B T}\right).$$

Addition of donor impurities raises the Fermi level, say, by an amount ΔE so that new Fermi level is

$$E'_F = E_F + \Delta E$$

Therefore density of majority carriers, *i.e.* electron density on n side will be

$$(n_o)_{n\ side} = 2\left(\frac{2\pi m_e^* K_B T}{h^2}\right)^{3/2} \exp\left(\frac{E'_F - E_F}{K_B T}\right)$$

$$= 2\left(\frac{2\pi m_e^* K_B T}{h^2}\right)^{3/2} \exp\left(\frac{E_F + \Delta E - E_c}{K_B T}\right)$$

$$= n_i \exp(\Delta E/K_B T). \qquad \ldots(17)$$

The density of minority carriers, *i.e.* electron density on p side will be

$$(n_c)_{p\ side} = \frac{n_i^2}{n_h},$$

where n_h is the hole density on p side and given by

$$n_h = n_i \exp(\Delta E'/K_B T).$$

Therefore $(n_c)_{p\ side} = n_i \exp(-\Delta E'/K_B T)$. $\qquad \ldots(18)$

From equations (17) and (18), we get

$$\frac{(n_c)_n}{(n_c)_p} = \exp\left(\frac{\Delta E + \Delta E'}{K_B T}\right)$$

$$= \exp(eV_B/K_B T)$$

or

$$V_B = \frac{K_B T}{e} \log\left[\frac{(n_c)_n}{(n_c)_p}\right]. \qquad \ldots(19)$$

We can calculate the minority carrier density from the relation

$$n_c \times n_h = n_i^2,$$

for a semiconductor or crystal.

In a crystal containing N_d completely ionised donor atoms, we can write

$$(n_o' + N_d)\, n_h' = n_i^2,$$

where n_o' and n_h' are thermally generated electron-hole pairs. When doping is such that $N_d >> n_o'$ and is true for n type crystal, then

$$N_d\, n_h' = n_i^2,$$

in which n_h' should give the hole concentration in n type crystal or in other words give minority carrier concentration. Thus

$$(n_h')_n = \frac{n_i^2}{N_d} \qquad \ldots(20)$$

Similarly, in crystal containing N_a completely ionised acceptor atoms, we can write

$$n_c' \,(N_a + n_h') = n_i^2;$$

when doping is such that $N_a >> n_h'$ and is true for p type crystal then

$$n_c' N_a = n_i^2.$$

$$(n_c')_p = n_i^2 / N_a. \qquad \ldots(21)$$

gives minority carriers (electrons) concentration in p type crystal.

From equations (21) and (19), we get

$$V_B = \frac{K_B T}{e} \log \left[\frac{(n_c)_n\, N_a}{n_i^2} \right], \qquad \ldots(22)$$

assuming that minority carriers on p side are thermally generated *i.e.* taking

$$(n_c')_p = (n_o)_p.$$

Further majority carriers in n type will be largely contributed by donor atoms and we can take

$$(n_o)_n = N_d,$$

so that equation (22) becomes

$$V_B = \frac{K_B T}{e} \log \left(\frac{N_d N_a}{n_i^2} \right) \qquad \ldots(23)$$

6·11. BIASSING OF P-N JUNCTION (*quantitative*) :

We shall discuss biassing of pn junction shown in fig. 12 in three steps :

(A) p-n junction in equilibrium (no bias) :

Refer to fig. 13(a). When pn junction is in equilibrium, the number of carriers diffusing from p side to n side is equal to the number of carriers diffusing from n side to p side. Consequently, there is no current across the junction; this can also be concluded in the following way ·

The electron current from p to n is proportional to the product of electron density in p region and the probability factor. The probability of an electron sweeping from p to n region is unity because any electron which comes to the edge of transition region

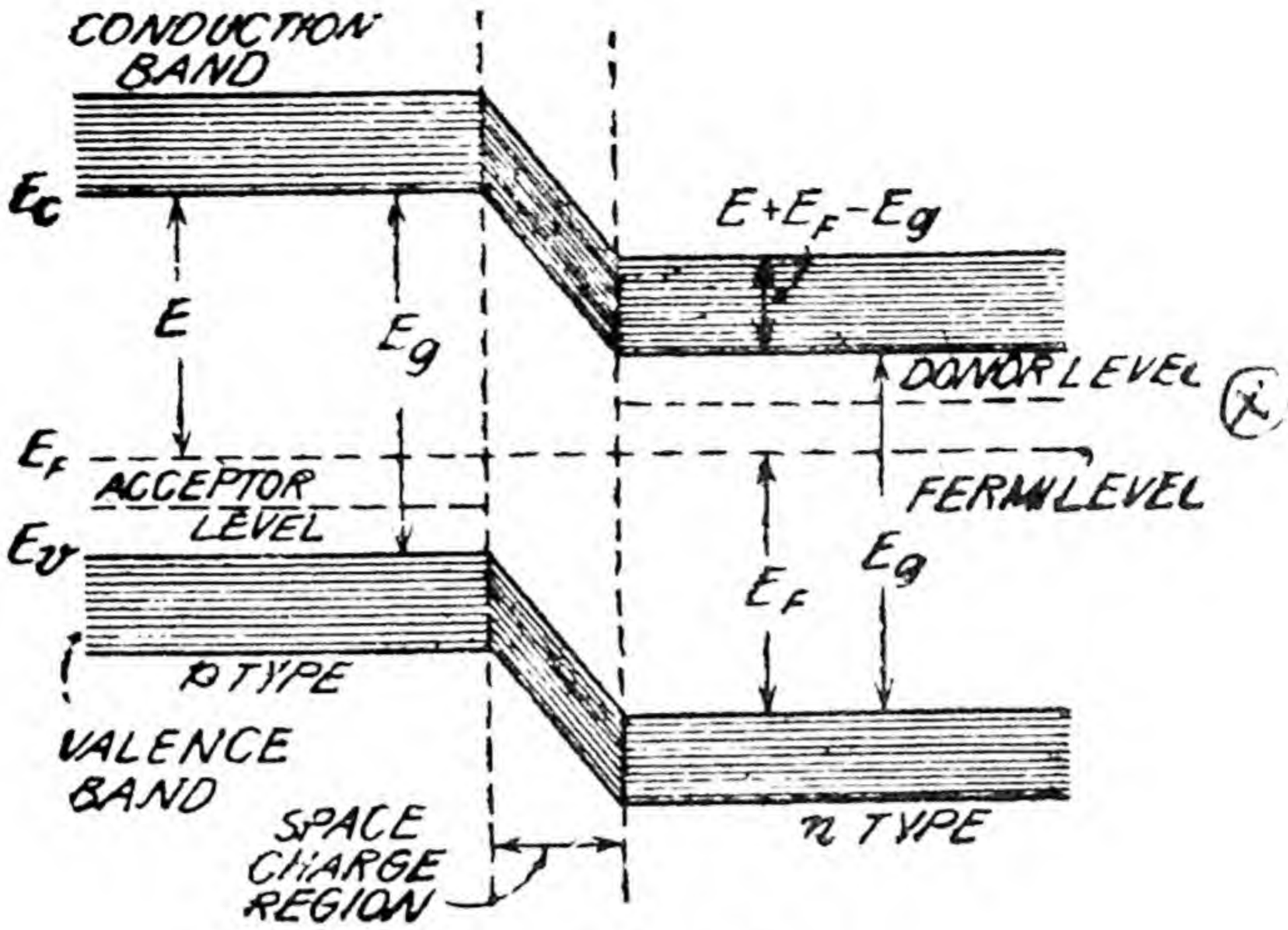


Fig. 12. *p-n* junction.

on *p* side will definitely sweep to *n* side for the former is a higher potential side. The density of electrons in the conduction band of *p* type is proportional to the Boltzmann factor $\exp[(E_F - E_c)/K_BT]$; therefore electron current from *p* to *n* is

$$I_{p\to n} = \text{probability} \times \text{density of electrons in conduction band of } p \text{ type,}$$

$$= 1 \times 2\left(\frac{2\pi m_c^* K_B T}{h^2}\right)^{3/2} \exp\{(E_F - E_c)/K_B T\}$$

$$= C \exp(-E/K_B T), \qquad \ldots(1)$$

where C is the constant of proportionality and E is $(E_c - E_F)$.

To calculate the electron current from *n* side to *p* side, we note that

$$I_{n\to p} = \text{probability} \times \text{density of electron in conduction band of } n \text{ type.}$$

The probability of an electron to climb up the potential barrier of height $(E+E_F-E_g)$ is proportional to the Boltzmann factor $\exp[-(E+E_F-E_g)/K_BT]$. The density of electrons in the conduction band of *n* type is proportional to $\exp[-(E_g-E_F)/K_BT]$. Therefore,

$$I_{n\to p} = \exp[-(E+E_F-E_g)/K_BT] \times C \exp[-(E_g-E_F)/K_BT]$$

$$= C \exp[(-E-E_F+E_g-E_g+E_F)/K_BT]$$

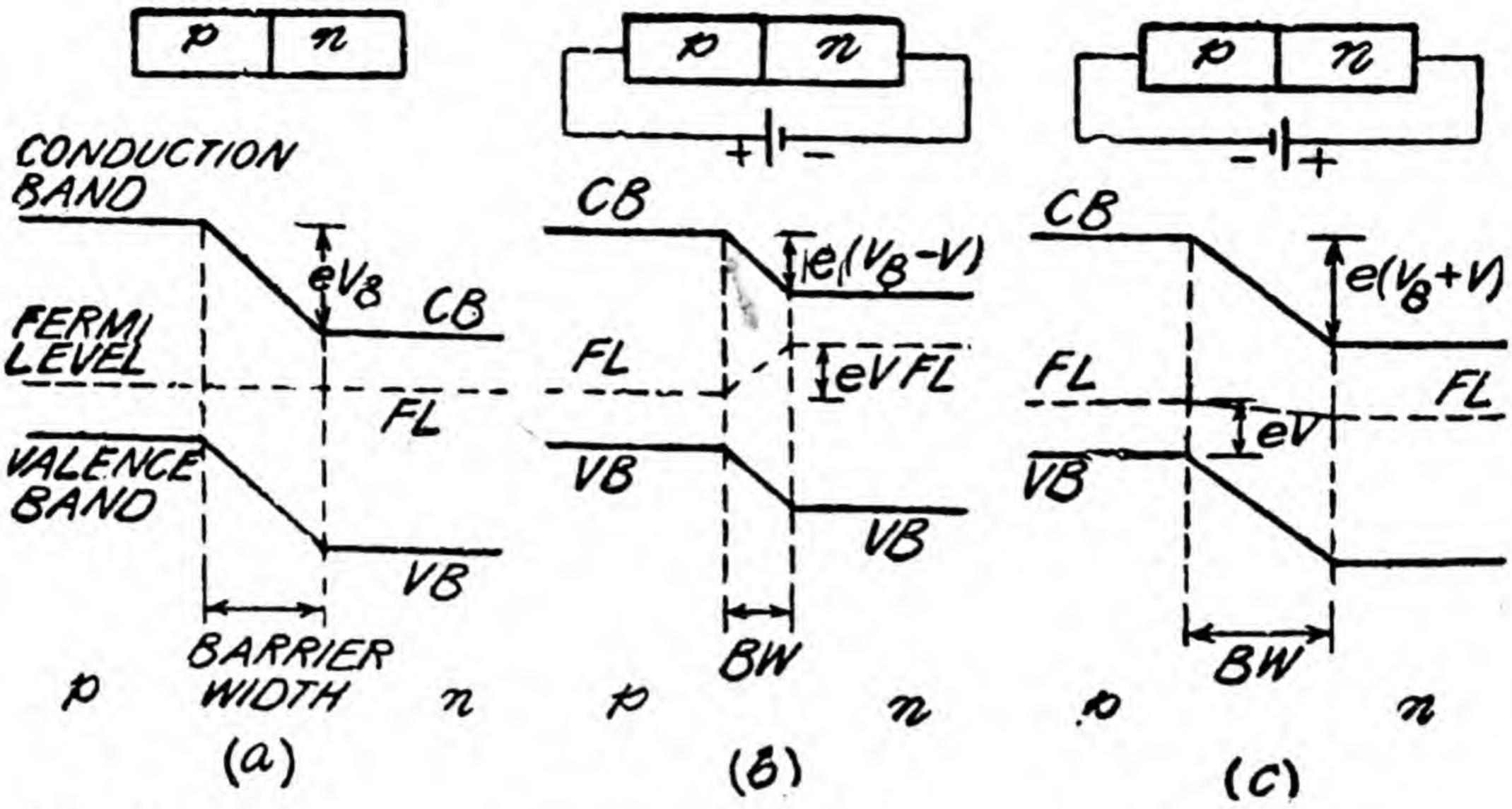
$$= C \exp(-E/K_BT). \qquad \ldots(2)$$

From equations (1) and (2), we note that

$$I_{p\to n} = I_{n\to p}$$

and therefore, there is no net current across *p–n* junction. The same argument holds good for holes too.

(B) p-n junction with forward bias : In fig. 13 (*b*) forward



(*a*) No bias (*b*) Forward bias (*c*) Reverse bias

Fig. 13. Biassing of *p-n* junction.

bias connections are shown. *p* side is connected to the positive of the battery. Due to the external bias, equilibrium conditions are disturbed and therefore energy bands and the Fermi level are altered. Since in this case, negative of the battery is connected to *n* type side, energy of the electrons in *n* type region increases by an amount eV, where V is voltage applied by the battery. Consequently, as shown in fig. 13 (*b*) Fermi level rises by eV and the energy bands adjust their positions so as to suit the elevation of Fermi level. Due to the increase in energy on *n* type side, potential barrier is reduced to $e(V_B - V)$. According to equation (14) of art 6·10, space charge region width or barrier width or depletion width is also reduced. The net result is that the electrons crossing the junction from *n* side will now face a low potential barrier and small widths of space charge region. Consequently, they can easily cross the junction. Since electrons in *n* type side are in majority, current across the junction increases considerably as shown in fig. 14. The variation of junction current with different forward biassing voltages is shown in fig. 14.

(C) Reverse Bias : Connections are shown in fig. 13 (*c*), *n* side is connected to the positive of battery. This lowers the Fermi level on *n* side by an amount eV raising the barrier height to $e(V_B + V)$ and increasing the depletion width too. The electrons that are in majority in *n* side will now face, in crossing the junction, a greater potential barrier and larger width of space charge region. Therefore, number of electrons crossing the junction from *n* side to *p* side decreases. Consequently, current is very much reduced in reverse biasing of the *p-n* junction as shown in fig. 14. *p-n* junction thus acts as a rectifier like vacuum diode.

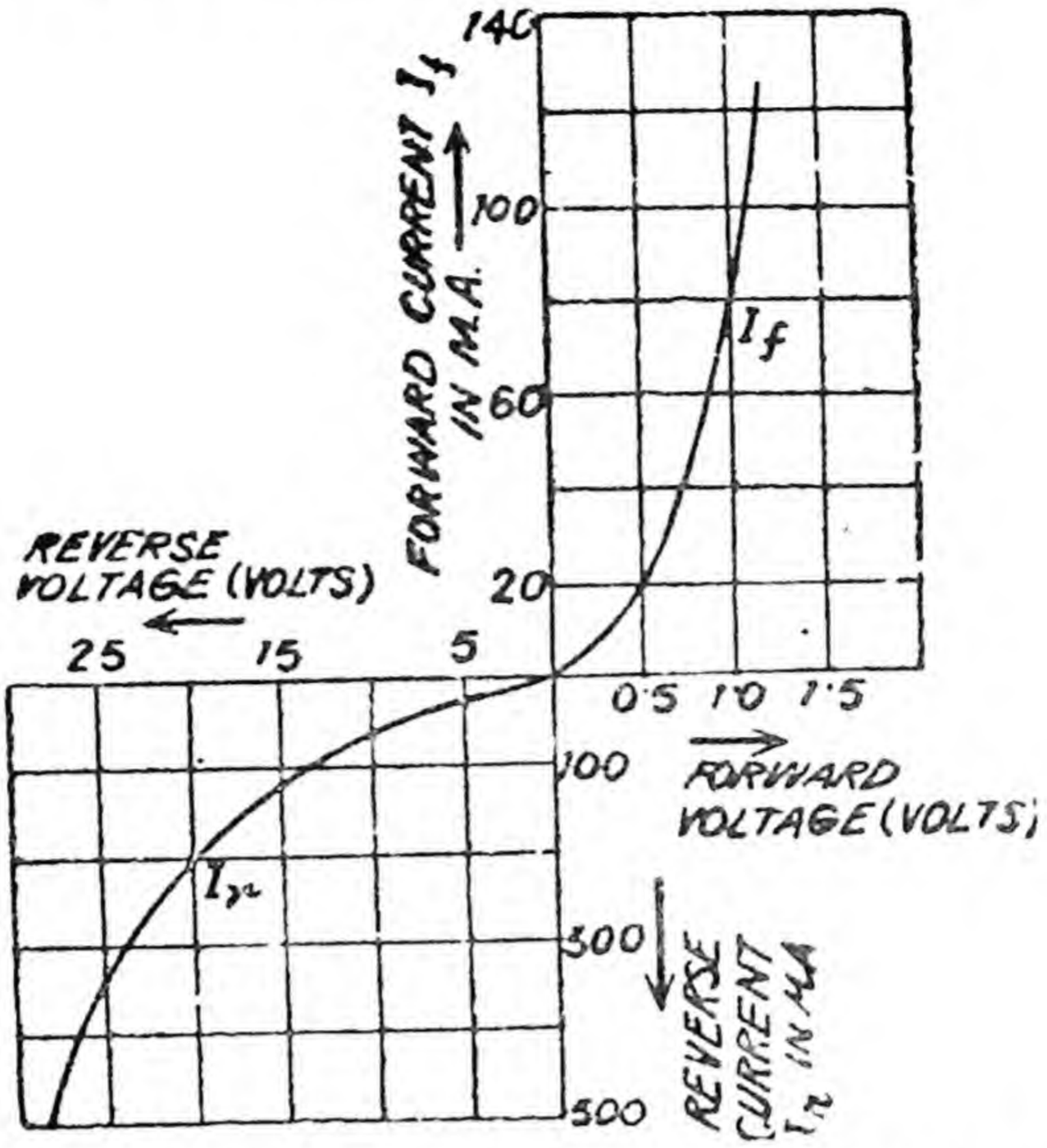


Fig. 14. Variation of current with different biassing voltages.

Breakdown of p-n junction : When a large reverse bias is applied, *p-n* junction breaks down. This break down occurs due to two main reasons :

(*i*) *Avalanche breakdown* : In this breakdown, conduction band electrons make ionising collisions with the atom of the crystal resulting in the production of electrons which further, in their turn, collide with other atoms. Thus process of collisions is stepped up. The cummulative effect of increased number of electrons is to break the *p-n* junction.

(*ii*) *Zener breakdown* : This breakdown occurs when applied electric field is sufficiently high to push out the electrons of valence band in large numbers to conduction band. However, the voltage at which this breakdown occurs can be controlled by adjusting the impurity concentrations of the junction materials.

6·12. CURRENT ACROSS P-N JUNCTION (*qualitative*) :

When a *P* type material is intimately joined to *N* type, *P-N* junction is formed. Simply by putting the two junctions in contact, *P-N* junction can not be produced because the lattice remains discontinuous at the junction.

In fig. 15 unneutralised acceptor atoms with negative symbols in *P* section are encircled. The negative symbol is used because of the presence of the additional electron which was obtained from an adjacent electron pair bond. The hole left by the removed electron is represented by a small positive symbol. In *N* section, unneutralised donor atoms with positive symbols are encircled. The positive symbol is used because the atom is left with positive charge when the fifth electron has been removed. The free electrons are represented by small negative symbols. The circled sign represent charges on impurity atoms and are fixed in position

in the crystal, while the uncircled sign represent the charges of free electrons and holes released by the impurities. These electrons and holes are known as majority carrier. Due to thermal dissociations of covalent bonds, there are few holes in *N* region and few electrons in *P* region. These are referred to as minority carriers.

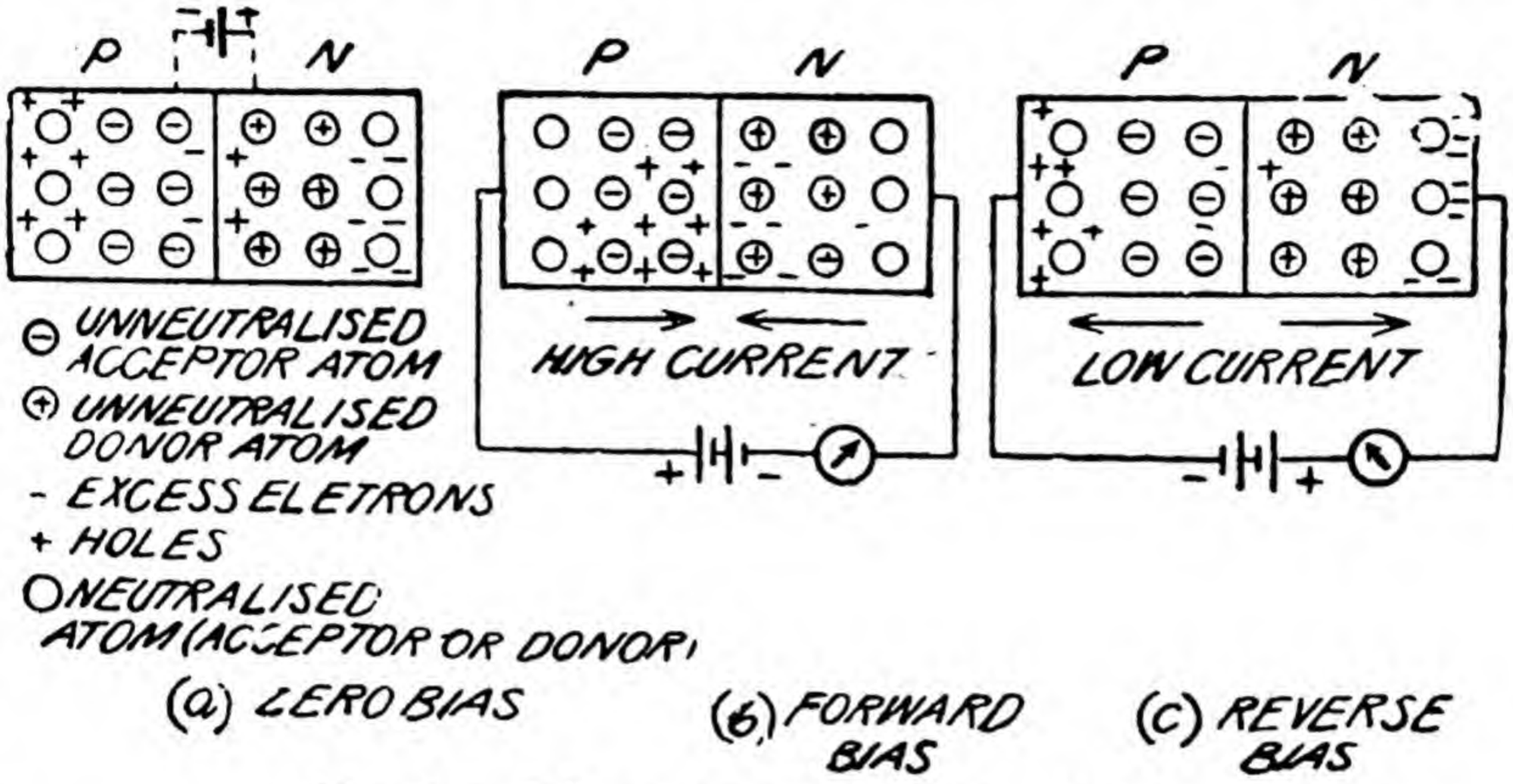
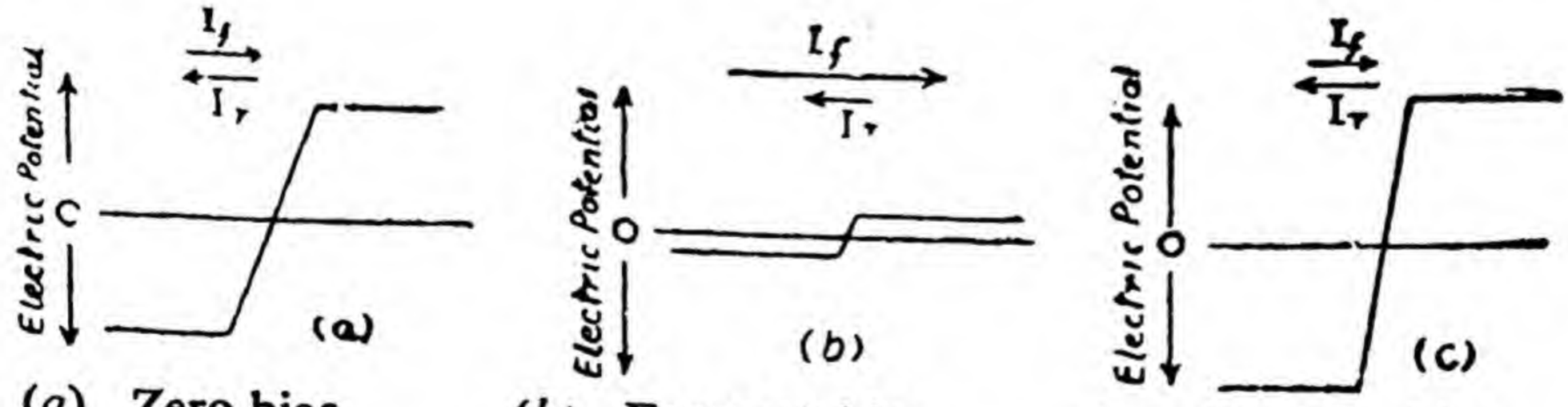


Fig. 15. Effect of junction bias voltage.

In such a crystal, near the junction due to diffusion, free electrons of the *N* section will combine with the holes of *P* section leaving unneutralised positive donor atoms and negative acceptor atoms in *N* and *P* sections respectively. This produces an electric field across the junction as if a fictitious battery, shown in fig. 15 (*a*), connected with its positive terminal to *N* and negative terminal to *P* section. This will prevent the electrons and holes from crossing the junction against the electrostatic forces developed. Due to thermal agitation, only few electrons and holes can cross the junction as ordinarily their energy will be insufficient to overcome this repelling force. Variation of potential at *P-N* junction is shown in fig. 16.



(*a*) Zero bias, (*b*) Forward bias, (*c*) Reverse bias.

Fig. 16. Variation of potential at *P-N* junction.

Forward bias : In order to obtain current through the junction, it is necessary to neutralise this repelling force across the junction. This is done by applying forward bias across the crystal, that is, a battery is connected in the opposite direction to fictitious battery, making *P* section positive and *N* section negative. The free electrons in *N* section will be repelled by the negative force set up by the power source and will move towards the junction. At the

same time, the holes in the *P* section will be repelled by the positive force set up by the power source and will also move towards the junction. The applied voltage should be high to impart sufficient energy to these carriers to overcome the potential barrier at the junction and enable them to cross through it. In this way there is a flow of a number of free electrons from *N* region to *P* region and a good number of holes from *P* region to *N* region through the junction. As the applied voltage is increased further, the electrons and holes having lower energy will be able to cross the potential barrier and the current will increase further.

The current flow can be understood as follows : For each hole in *P* section that combines with an electron from *N* section, an electron from an electron pair bond leaves the crystal and enters the positive terminal of battery. Due to this action, a new hole is created which is again forced by the power source to move towards the junction. For each electron in *N* section that combines with a hole from *P* section, an electron enters the crystal from the negative terminal of the battery. The constant movement of electrons towards the positive terminal and the holes towards the negative terminal produces a high forward current I_f. The minority carriers in the two sections (electrons in *P* section and holes in *N* section) which are newly created produce a small reverse current I_r. The net current $(I_f - I_r)$ increases as the applied power of source is increased but the maximum value is always limited in order to avoid the damage of crystal junction. The variation of I_f with applied voltage is shown in fig. 14.

Reverse bias : If the crystal is reverse biased then the battery increases the restraining force and the current through junction is exceedingly small. The cause of this small current is the conduction by the minority electrons and holes obtained due to light and thermal agitation. The electrons in *N* section and holes in *P* section are attracted away from the junction under the action of applied power source and thus causing minute current in opposite direction. The reverse current variation with applied voltage is shown in fig. 14. The curves are known as static characteristics. The maximum reverse voltage that can be applied to the semiconductor should not produce excessive reverse current.

6·13. RECTIFIER EQUATION :

The net current density across the *p-n* junction has following four components : Refer to fig. 13.

(1) The current I_1 due to the flow of minority electrons in the conduction band of *p* type to the conduction band of *n*-type. The direction of I_1 will be shown from *n* to *p* opposite to the direction of electron flow from *p* to *n* thus I_1 flows from right to left.

(2) The current I_2 due to the flow of majority electrons in the conduction band of *n* type to the conduction band of *p* type. The direction of I_2 will be from *p* to *n* *i.e.* from left to right.

(3) The current I_3 due to the flow of majority holes in the valence band of p type to the valence band of n type. I_3 flows from p to n *i.e.* from left to right (the direction of current being the direction of flow of holes).

(4) The current I_4 due to the flow of minority holes in the valence band of n type to the valence band of p-type. The direction of I_4 is from n to p *i.e.*, from right to left.

Therefore net current density, I, flowing from left to right is

$$I=(I_2+I_3)-(I_1+I_4). \qquad ...(1)$$

The minority current of electrons from p side to n side is not affected by the application of forward or reverse bias because whatever be the bias, potential barrier is always down to n side and therefore any electron reaching the junction boundary from p side will easily sweep to n side. It means for I_1 no barrier exists and is thus constant. For this constancy I_1 is called saturation current density and is independent of applied voltage V.

The majority current of electrons from n side to p side is very sensitive to the applied voltage since electrons in the n type material face a barrier which may be raised or lowered by the applied voltage. Thus, for forward bias, current density for majority electron flow is

$$I_2=C_1 \text{ exp. } [-e\,(V_B-V)/K_BT], \qquad ...(2)$$

where C_1 is a constant.

By similar arguments I_4 is the saturation current density which does not depend on the applied voltage. I_3 arises due to majority holes flow from p side to n side and is given by

$$I_3=C_2 \text{ exp. } [-e\,(V_B-V)/K_BT], \qquad ...(3)$$

where C_2 is another constant.

We know that in equilibrium when no voltage is applied across the junction, there is no net current across the junction. This implies that when applied voltage is zero, net current density in each band is separately zero *. That is

with $V=0$

$$\left.\begin{aligned} I_2&=I_1 \\ I_3&=I_4 \end{aligned}\right\} \qquad ...(4)$$

Applying condition (4) to equation (2), we get

$$I_1=C_1 \text{ exp. } (-eV_B/K_BT)$$

$$\therefore \quad C_1=I_1 \text{ exp. } (eV_B/K_BT).$$

Substituting this value of constant in equation (2), we find

$$I_2=I_1 \text{ exp. } (+eV/K_BT). \qquad ...(5)$$

Similary applying condition (4) to equation (3), we get

$$I_4=C_2 \text{ exp. } (-eV_B/K_BT)$$

$$\therefore \quad C_2=I_4 \text{ exp. } (eV_B/K_BT).$$

*It is to be mentioned here that current flow to one band is indepenpent of the flow in other band since there can not be any accumulation of charge in either band under equilibrium conditions.

With substitution of C_2, equation (3) becomes

$$I_3 = I_4 \text{ exp. } (+eV/K_BT) \qquad ...(6)$$

Putting equations (5) and (6) in equation (1), we get

$$I = (I_1 + I_4) \text{ [exp. } (eV/K_BT - 1)]. \qquad ...(7)$$

$(I_1 + I_4)$ is replaced by I_0 the total saturation current density giving

$$I = I_0 \text{ [exp. } (eV/K_BT) \doteq 1]. \qquad ...(8)$$

Equation 8) is known as the rectifier equation. V is positive for forward bias and negative for reverse bias.

Suppose a forward bias is applied such that

$$V \gg 4 \frac{K_BT}{e},$$

then

$$\text{exp. } \left(\frac{eV}{K_BT}\right) \gg 1,$$

and equation (8) becomes

$$I = I_0 \text{ exp. } (eV/K_BT),$$

which shows that in forward bias current increases exponentially as shown in fig. 14.

When a reverse bias is applied such that

$$-V \gg \frac{4K_BT}{e}$$

then

$$\text{exp. } \left(-\frac{eV}{K_BT}\right) \ll 1.$$

and equation (8) becomes

$$I = -I_0,$$

which shows that current in reverse bias remains constant at I_0, the saturation current until the junction breaks down as shown in Fig. 14.

6·14. DIODE RESISTANCE :

One of the most important properties of the diode is its resistance in the forward and reverse bias. An ideal diode must offer zero resistance in forward bias and infinitely large resistance in reverse bias. Here we consider the two resistances *i.e.*, one d.c. resistance and the other a.c. resistance or dynamic resistance. The equivalent circuit for the diode is shown in fig. 17 where R_p and R_n are the resistances of P and N materials respectively.

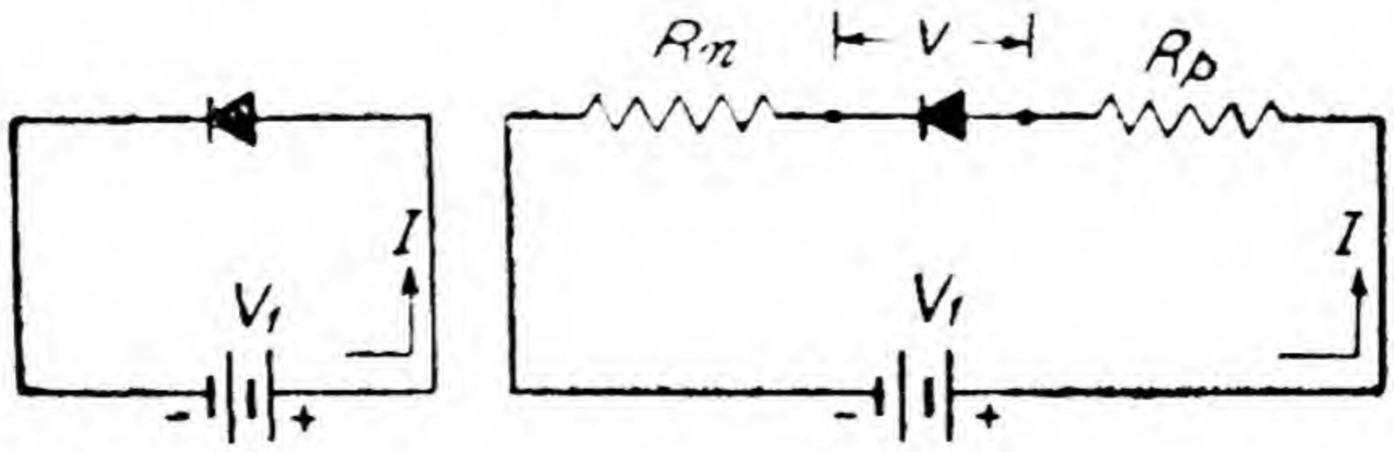


Fig. 17. The equivalent circuit of a diode.

The juction voltage is given by

$$V=V_1-I\,(R_n+R_p) \qquad \ldots(1)$$

The d.c. resistance is given by

$$\frac{V_1}{I}=\frac{V}{I}+R_n+R_p$$

or

$$R_{d.c.}=r_{d.c.}+R_n+R_p, \qquad \ldots(2)$$

where $r_{d.c.}$ represents the d.c. resistance of the ideal part of the diode. The value of $r_{d\,.c.}$ is given by

$$r_{d.c.}=\frac{V}{I}=\frac{V}{I_0\,[\text{exp. }(eV/K_BT)-1]} \qquad \ldots(3)$$

The a.c. resistance ($r_{a.c.}$) of the ideal diode is given by

$$r_{a.c.}=\frac{dV}{dI}=\frac{1}{dI/dV}$$

$$=\frac{1}{(I_0e/K_BT)\text{ exp. }(eV/K_BT)}$$

$$=\frac{K_BT}{e\,(I+I_0)} \qquad \ldots(4)$$

The a.c. resistance of a real diode is given by

$$R_{a.c.}=r_{a.c.}+R_n+R_p$$

or

$$R_{a.c.}=\frac{K_BT}{e\,(I+I_0)}+R_n+R_p \qquad \ldots(5)$$

At room temperature (300°K), the value of $r_{a.c.}$ is given by

$$r_{a.c.}\simeq\frac{0\cdot026}{I}\text{ ohms.}$$

where the current is in amperes.

6·15. DEPLETION REGION (SPACE CHARGE) CAPACITANCE :

The depletion region capacitance arises from the widening and narrowing of the depletion layers as a func tion of the junction voltage. The actual barrier width and the capacitance associated with the junction depend upon the exact junction structure. Here we shall consider the capacitance due to a gradual junction (as is the case with a grown junction). Here it is assumed that the junction is symmetrical and the conductivity of P and N materials is good so that there is negligible voltage drop in the regions outside the junction thickness. The different conditions are shown in fig.(18 The apparent junction is assumed to be located at O, where concentrations of donors, N_d and that of acceptors, N_a are equal. The variation of electron potential across the junction along X-axis is shown in fig. (18 b). The distribution of space charge is shown in fig. (18 c). The modified form of space charge is shown in fig. (18 d).

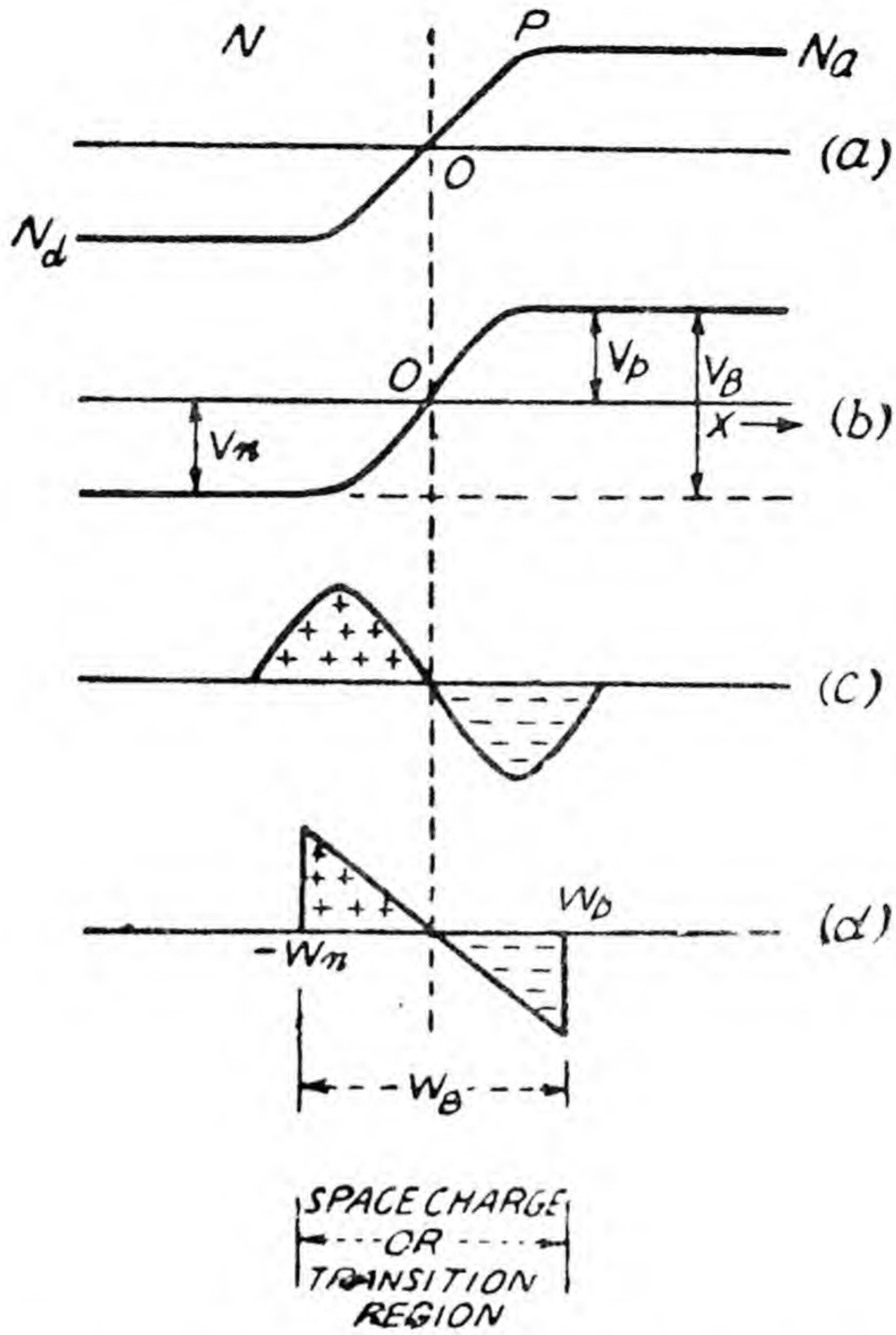


Fig. 18. Condition in a gradual junction.

The space charge density is assumed to vary linearly with distance along X-axis in space charge region and is given by

$$\rho(x) = ax \text{ for } -W_n < x < W_p. \qquad \ldots(1)$$

where $$a = \frac{e(N_d + N_a)}{W_B} = \frac{1}{W_B}\left(\frac{\sigma_n}{\mu_n} + \frac{\sigma_p}{\mu_p}\right).$$

The symbols have their usual meaning.

Considering P side of the junction and applying Poisson's equation, we have

$$\frac{d^2V}{dx^2} = \frac{-\rho(x)}{\epsilon} = -\frac{ax}{\epsilon} \qquad \ldots(2)$$

Applying the boundary conditions *i.e.*, $dV/dx = 0$ at $x = W_p$ and $V = 0$ at $x = 0$ and integrating twice the equation (2), we get

$$V = \frac{-ax}{2\epsilon}\left(\frac{x^2}{3} - W_p^2\right) \qquad \ldots(3)$$

The value of V_p, *i.e.*, the barrier potential in P region is given by

$$V_p = V(W_p) - V_{(0)} = \frac{aW_p^3}{3\epsilon} \qquad \ldots(4)$$

Again $W_n = W_p$

($\because$ junction is assumed to be symmetrical)

or $W_B = 2W_p$

Similarly, $V_n = V_p$ or $V_B = 2V_p$

$$\therefore \quad V_B = 2V_p = \tfrac{2}{3}\,\frac{aW_p^3}{\epsilon}$$

Thus $$W_B = 2W_p = 2\left[\frac{3\epsilon V_B}{2a}\right]^{1/3}. \qquad \ldots(5)$$

The total charge on either side of the junction can be found from the triangle of fig. (18.4*d*).

$$Q = \tfrac{1}{2}\,(\text{base})\,(\text{height}) = \tfrac{1}{2}\left(\frac{W_B}{2}\right)\left(\frac{aV_B}{2}\right)$$

$$= \frac{aW_B^2}{8}$$

$$\therefore \quad Q = \frac{a^{1/3}}{4}\left[\frac{3}{2}\,\epsilon V_B\right]^{2/3} \qquad \ldots(6)$$

Now $$C_T = -\frac{dQ}{dV_B} = \frac{a^{1/3}\,\epsilon^{2/3}}{4}\left(\frac{3}{2}\right)^{2/3} V_B^{-1/3}$$

$$= \left(\frac{3\epsilon}{16}\right)^{2/3}\left(\frac{a}{V_B}\right)^{1/3}$$

or $$C_T = \left(\frac{3\epsilon}{16}\right)^{2/3}\left\{\frac{(\sigma_n/\mu_n)+(\sigma_p/\mu_p)}{W_B V_B}\right\}^{1/3}. \qquad \ldots(7)$$

This shows that capacitance varies inversely as the cube root of applied voltage.

SOLVED PROBLEMS :

1. *Find the mobility of electrons in copper assuming that each atom contributes one free electron for conduction. Resistivity of copper is* $1{\cdot}7 \times 10^{-6}$ *ohm-cm. Atomic weight* 63·54. *Density* 8·96 *gms./c.c. Avogadro's number* $6{\cdot}025.10^{23}$. *Electronic charge* $1{\cdot}6.10^{-19}$ *coulomb.*

We shall first find the number of free electrons per unit volume and is given by

$$\frac{\text{Avogardro's number} \times \text{density}}{\text{Atomic weight}} = \frac{6{\cdot}025 \times 10^{23} \times 8{\cdot}96}{63{\cdot}54}$$

$$= {\cdot}496 \times 10^{22}.$$

Therefore mobility of electrons is given by

$$\mu_n = \frac{\sigma_n}{ne},$$

where σ_n is the conductivity. If ρ_n be the resistivity, then

$$\rho_n = \frac{1}{\sigma_n},$$

so that $$\mu_n = \frac{1}{\rho_n n e}$$

$$= \frac{1}{1{\cdot}7 \times 10^{-6} \times 8{\cdot}496 \times 10^{22} \times 1{\cdot}6 \times 10^{-19}}$$

$$= 43{\cdot}28 \text{ cm}^2/\text{volt-sec.}$$

2. *Germanium is doped with aluminium so that the concentration of acceptor atom is* 2×10^{21} *atoms/m*^{3}. *Determine at room temperature* (*i*) *hole concentration,* (*ii*) *free electron concentration, and* (*iii*) *conductivity, given that mobility of hole* $\mu_p = 0{\cdot}17$ *m/sec./volt/metre.*

(*i*) At room temperature (300°*K*) all the acceptor atoms are assumed to be ionised giving an equal number of holes

$$N_a = 2 \times 10^{21} \text{ holes/m}^3.$$

(*ii*) Intrinsic concentration of free electrons at room temperature is given by

$$n_i = AT^{3/2}\, e^{-(E_g/2K_BT)}$$

$$= 2\left(\frac{2\pi m K_B}{h^2}\right)^{3/2} T^{3/2}\, e^{-(E_g/2K_BT)}$$

$$= 2\left(\frac{2\pi \times 9{\cdot}1 \times 10^{-31} \times 1{\cdot}381 \times 10^{-23}}{(6{\cdot}624)^2 \times 10^{-68}}\right)^{3/2} (300)^{3/2}$$

$$\times \exp.\left\{-\left(\frac{0{\cdot}72 \times 1{\cdot}59 \times 10^{-19}}{2 \times 1{\cdot}381 \times 10^{-23} \times 300}\right)\right\}$$

where energy gap $E_g = 0{\cdot}72$ *eV*.

$$= 4{\cdot}82 \times 10^{21} \times (300)^{3/2}\, e^{-13.0}$$

$$= \frac{4{\cdot}82 \times 10^{21} \times 300\sqrt{300}}{10{,}59{,}000} = \mathbf{2{\cdot}36 \times 10^{19}\ \text{electrons (or holes)/m}^3}$$

Since the product *np* is constant, (independent of doping) and is equal to n_i^2, therefore, free electron concentration in the doped semiconductor is

$$n = \frac{n_i^2}{p} = \frac{(2{\cdot}36 \times 10^{19})^2}{2 \times 10^{21}} = \mathbf{2{\cdot}78 \times 10^{17}\ \text{electrons/m}^3}$$

This is negligible as compared to hole concentration.

(*iii*) Conductivity $\sigma = q\,(n\mu_n + p\mu_p)$

$$\approx qn\mu_n$$

$$\approx 1{\cdot}59 \times 10^{-19} \times 2 \times 10^{21} \times 0{\cdot}17$$

$$\approx \mathbf{0{\cdot}54 \times 10^2\ \text{mhos/m}}.$$

3. *A p·n junction is formed from germanium of conductivity* 0·8 *ohm*$^{-1}$ *cm*$^{-1}$ *on p side and* 1·6 *ohm*$^{-1}$ *cm*$^{-1}$ *on n side. Calculate potential barrier and depletion width at* 300°*K* ($n = 2{\cdot}1 \times 10^{23}$ *cm*$^{-1}$). $\mu_p = 2000$ *cm*^{2}*/volt sec and* $\mu_n = 4000$ *cm*^{2}*/volt-sec. The relative dielectric constant for germanium is* 16 *and* K_B *is* $1{\cdot}38 \times 10^{-16}$ *erg per* °*K.*

Taking majority carriers contributing mainly to the conductivity, we can write

$$N_d=\frac{\sigma_n}{e\mu_n}=\frac{1\cdot6}{1\cdot6\times10^{-19}\times4000}$$
$$=2\cdot5\times10^{15}\text{ per c.c.}$$
$$N_a=\frac{\sigma_p}{e\mu_p}=\frac{0\cdot8}{1\cdot6\times10^{-19}\times2000}$$
$$=2\cdot5\times10^{15}\text{ per c.c.}$$

At 300°K,

$$\frac{K_BT}{e}=\frac{1\cdot38\times10^{-16}\times300\times10^7}{1\cdot6\times10^{-19}}=0\cdot0258.$$

Multiplication by 10^7 has been done to convert ergs into joules.

The potential barrier is

$$V_B=\frac{K_BT}{e}\log\left(\frac{N_dN_a}{n_i^2}\right)$$
$$=0\cdot0258\log\left(\frac{6\cdot25\times10^{30}}{4\cdot41\times10^{26}}\right)$$
$$=0\cdot2467\text{ volts.}$$

The depletion width is

$$X=\frac{2\epsilon V_B}{e(N_a+N_d)^{1/2}}\left[\left(\frac{N_d}{N_a}\right)^{1\cdot2}+\left(\frac{N_a}{N_d}\right)^{1/2}\right]$$

Taking $V_B=\frac{0\cdot2467}{200}$ e. s. u.

and $e=4\cdot8.10^{-10}$ e. s. u.

Here $\frac{N_d}{N_a}=\frac{N_a}{N_d}=1$ and $(N_a+N_d)=5.10^{15}$

$\epsilon=16$,

$$X=\left[\frac{2\times16\times0\cdot2467}{300\times0\cdot8\times10^{-10}\times5\times10^{15}}\right]^{1/2}[(1)^{1/2}+(1)^{1/2}]$$
$$=\left[\frac{0\cdot4934}{45\times10^6}\right]^{1/2}\times2$$
$$=\left[\frac{49\cdot34\times10^{-8}}{45}\right]^{1/2}\times2$$
$$=1\cdot9\times10^{-4}\text{ cm.}$$

6·16. JUNCTION TRANSISTOR :

The two forms used are :

(*i*) *PNP type* : This is formed by sandwitching a thin layer of *N*-type semiconductor between two layers of *P*-type semiconductor.

(*ii*) *NPN type* ; This is formed by sandwitching a thin layer

of *P*-type semiconductor between two layers of *N*-type semiconductor.

Both types are shown in figure 19(*a*), with their symbols for representation.

The centre section is called the base, one of the outside sections —the emitter, and the other outside section—the collector.

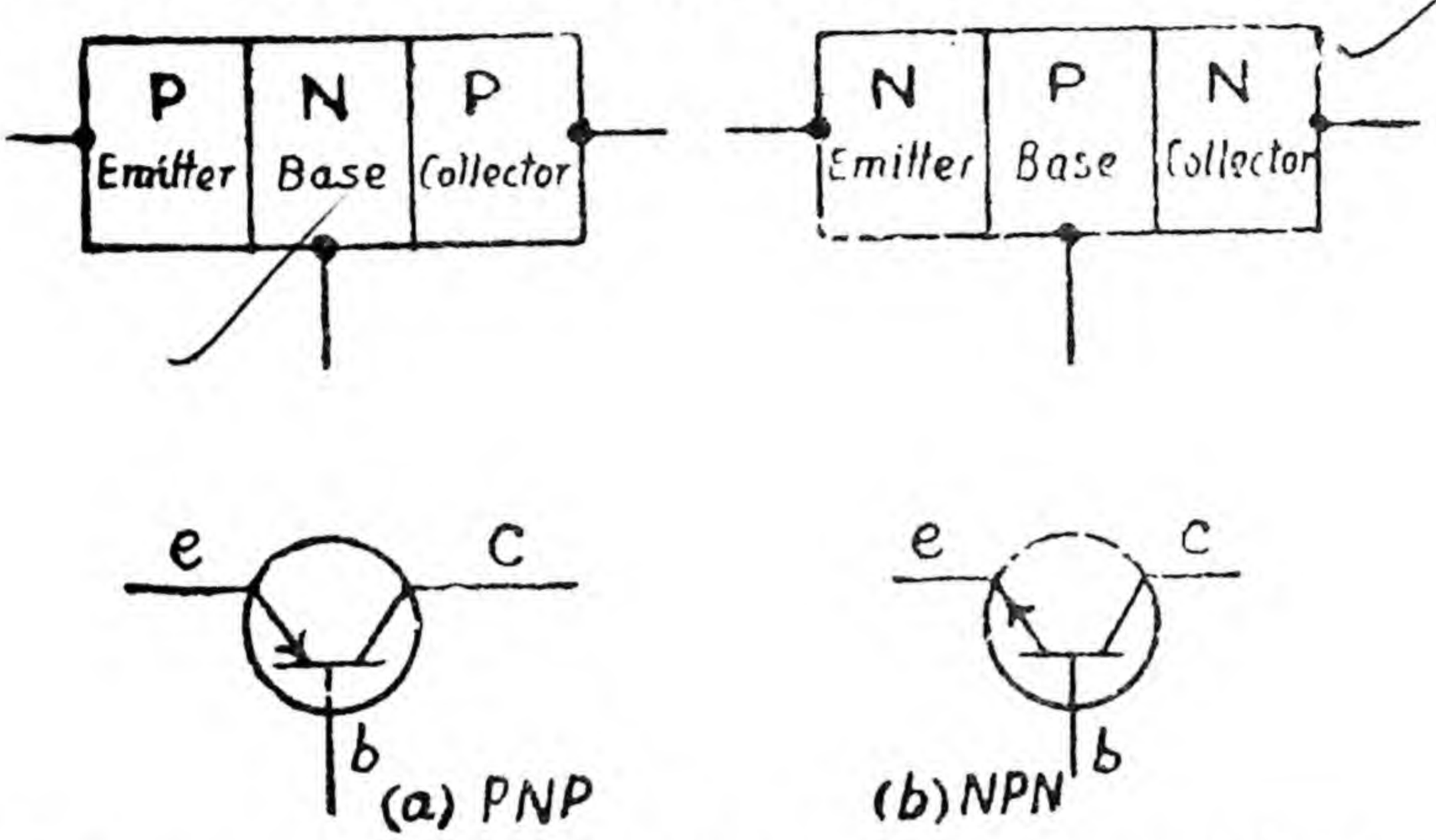
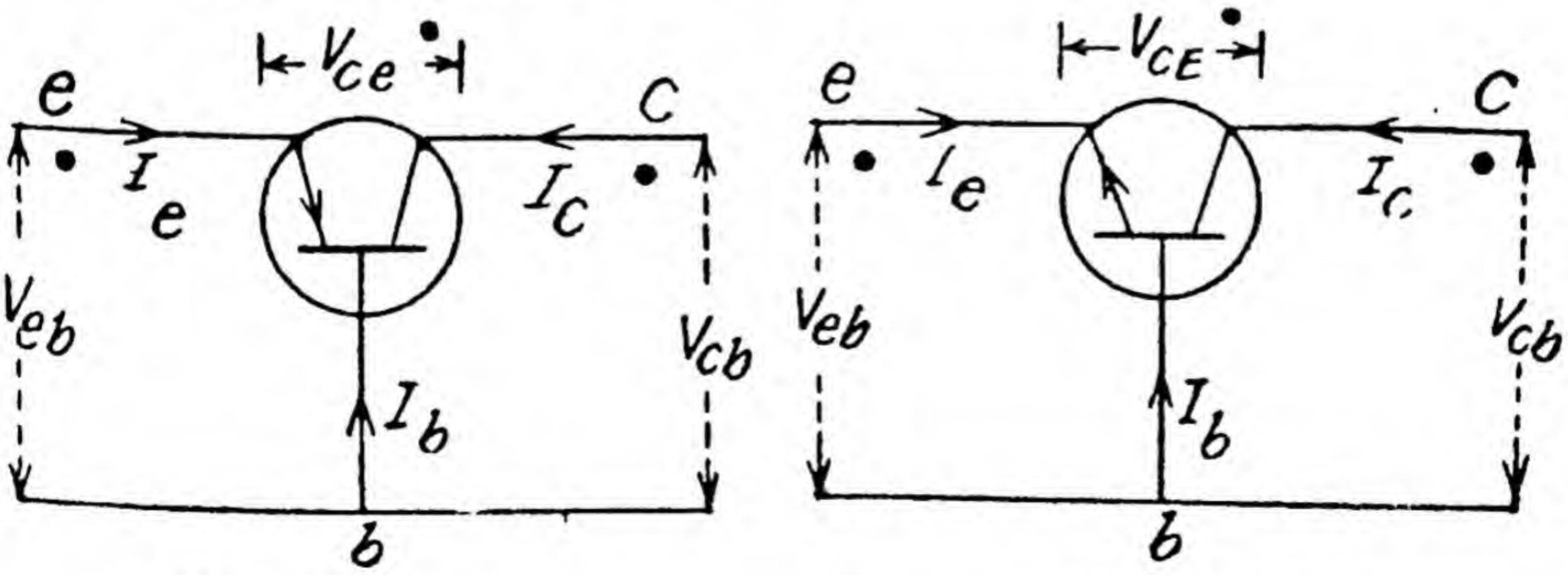


Fig. 19. (*a*) Illustration of *PNP* and *NPN* junction transistors.

The direction of the arrowhead gives the direction of the conventional current with forward bias on the emitter. The conventional flow is opposite in direction to the electron flow. Because in *PNP* transistor, the emitter current is constituted by the flow of holes from emitter into base region, hence the conventional current flows from emitter into base. Thus the arrowhead points towards the base. On the other hand, in *NPN* transistors the emitter current is constituted by the flow of electrons from emitter into base region, hence the conventional current flows from base into emitter. Thus the arrowhead points outward the base.

Fig. 19 (*b*) shows the conventional polarities of voltages and currents in both types of transistors.



PNP transistor NPN transistor

Fig. 19 (*b*) Conventional polarities of currents and voltages

The emitter current I_e, base current I_b and collector current

I_c are taken positive when they flow into the transistor. Considering the common base configuration, the emitter and collector voltages are refered to base *i.e.*, the voltage between emitter and base is represented by V_{eb} while the voltage between collector and base is represented by V_{cb}. The voltage between collector and emitter is expressed by V_{cc}. A dot is used on the arrowhead to indicate the positive polarity.

Operation of PNP Transistor :

A *PNP* junction transistor is made by sandwitching two *PN* germanium or silicon diodes, placed back to back. The centre of *N* type portion is extremely thin in comparison to *P* regions.

The *P* region of the left is connected to the positive terminal and *N* region to the negative terminal, *i.e.*, *PN* is biased in the forward direction while *P* region of right is biased negatively *i.e.*, in the reverse direction as shown in fig. 20 (*a*). The *P* region in the forward biased circuit is called *emitter* and the *P* region on the right, biased negatively is called *collector*. The centre is called base.

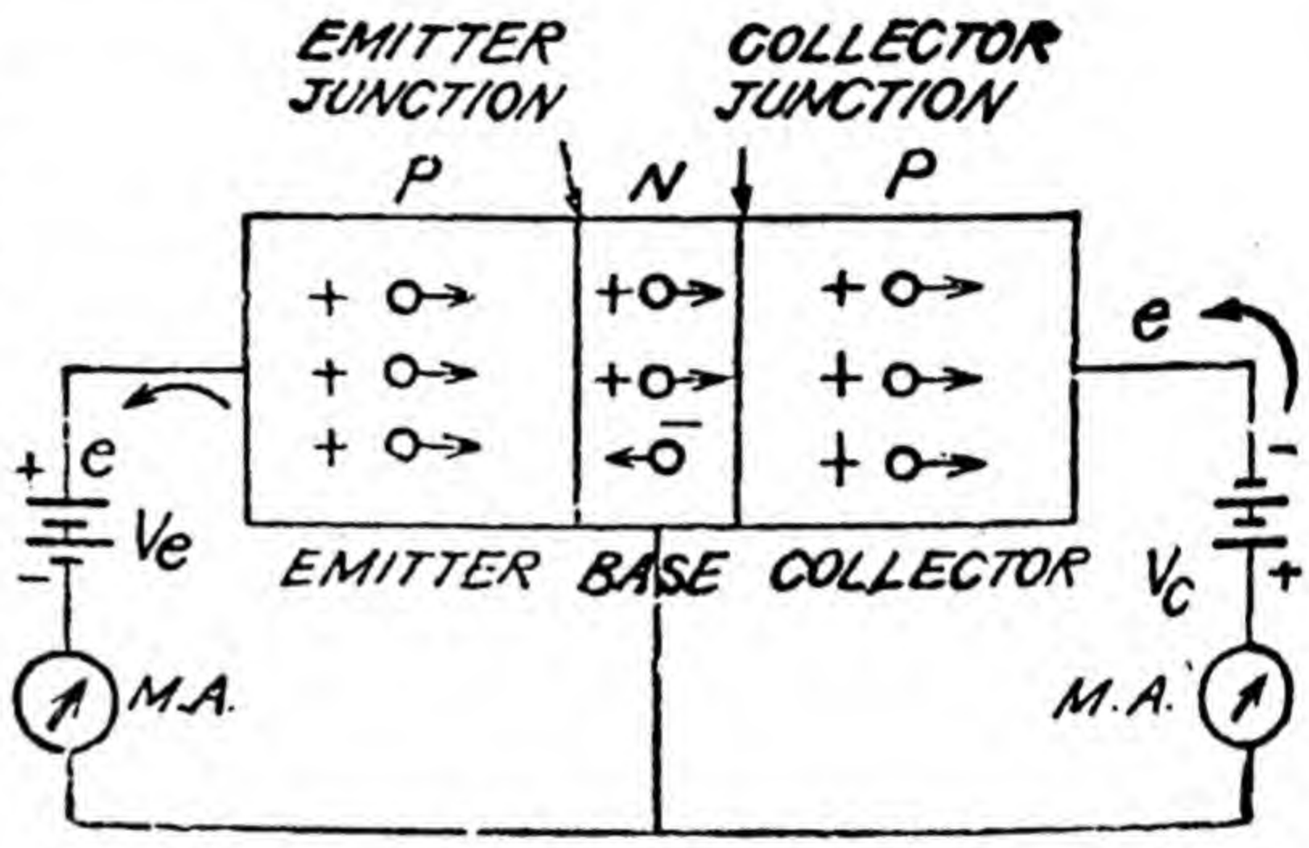


Fig 20 (*a*) Basic connection of *PNP* junction transistor

The majority carriers (holes) of *P* region (known as emitter) move to *N* region as they are repelled by the positive terminal of battery while the electrons of *N* region are attracted by the positive terminal. The holes overcome the barrier and cross the emitter junction into *N* region. As the width of base region is extremely thin, two to five percent of the holes recombine with the free electrons of *N* region which result in a small base current, while the remaining holes (95% to 98%) reach the collector junction. The collector is biased negatively and the negative collector voltage aids in sweeping the hole into collector region.

As the *P* region at the right is biased negatively, a very small current should flow but the following facts are observed :

(*i*) a substantial current flows through it when the emitter junction is biased in a forward direction.

(*ii*) the current flowing across the collector is slightly less than that of the emitter, and

(*iii*) the collector current is a function of emitter current, *i.e.* with the decrease or increase in the emitter current a corresponding change in the collector current is observed.

The facts can be explained as follows :

(*i*) The collector current increase as the holes reaching the collector junction are attracted by negative potential applied to the collector.

(*ii*) As already discussed 2 to 5% of the holes are lost in recombination with electrons in base region which result in a small base current, and hence the collector current is slightly less than the emitter current.

(*ii*) When the emitter current increases more holes are injected into the base region which are attracted by the negative potential of the collector and hence results in increasing the collector current. In this way emitter current controls the collector current. The control of collector current by the emitter is analogous to the control of plate current by small grid voltage in vacuum triode.

Hence we can say that when the emitter is forward biased and collector is negatively biased, a substantial current flows in both the circuits. Since a small emitter voltage of about 0·1 to 0·5 volt permits the flow of an appreciable emitter current, the input power is very small. The collector voltage can be as high as 45 volts, thus permitting relatively large output power.

Beside hole current another component of current flows from base region to emitter region due to the flow of electrons. This current is a function of emitter base potential When the width of base region is very small, the ratio between hole current and electron current is very large and for the practical purposes the electron current can be neglected Thus only the hole current plays the important role in the operation of *PNP* transistor.

NPN Transistor Biassing : As shown in figure 20 (*b*) for an amplifying action of *NPN* transistor, emitter-base junction is biased with a forward voltage and base-collector junction with a reverse voltage. A relatively high current will flow through the emitter base junction and a relatively low current through the base collector junction. The base is made thin so that its width is less than the mean path taken by an electron in the base before combining the hole. The effect of narrow base will be the diffusion of electrons into base-collector junction, where the applied voltage will draw them to collector terminal. The number of emitter electrons combining with base holes will, therefore, by very small due to which base current is always small while, as stated above, emitter-collector current is large. This explains the amplifying action of the transistor.

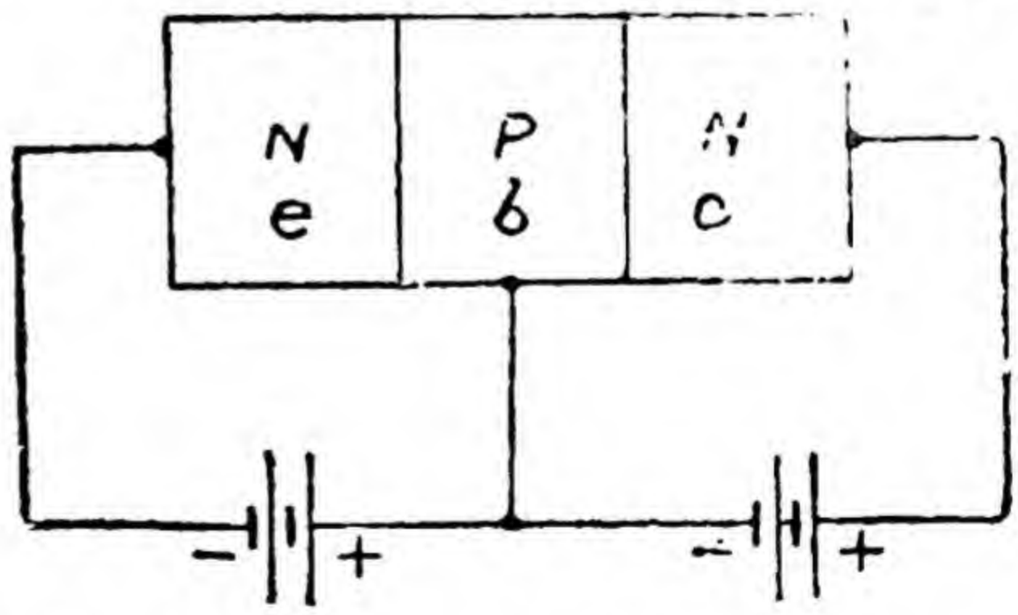


Fig. 20. (*b*). *NPN* transistor with biassing potential.

From the above analysis following points are observed in case of junction transistors :

(i) *Holes* are the major charge carriers in the *PNP* junction.

(ii) *Electrons* are the major charge carriers in the *NPN* junction.

(iii) In either type of junction transistors, the collector current is always less than the emitter current. This is due to the recombination of holes and electrons in base area. The ratio of collector to emitter current is known as current gain.

6 17. CHARACTERISTICS OF TRANSISTOR :

The method of connèction of a transistors into a circuit largely affects its input and output impedances. Therefore, the characteristics of the transistor will vary according to the methods of connection. There are three methods of connection, (*i*) common base, (*ii*) common emitter, (*iii*) common collector.

As compared to vacuum tube, base serves the purpose of contro grid, collector that of anode and emitter that of cathode. The word 'common' is used to denote the electrode that is common to both the input and output circuits. We have seen that common electrode is usually grounded; therefore for 'common' we can also write the word ' grounded". All the three types of connections are shown below in fig. 21.

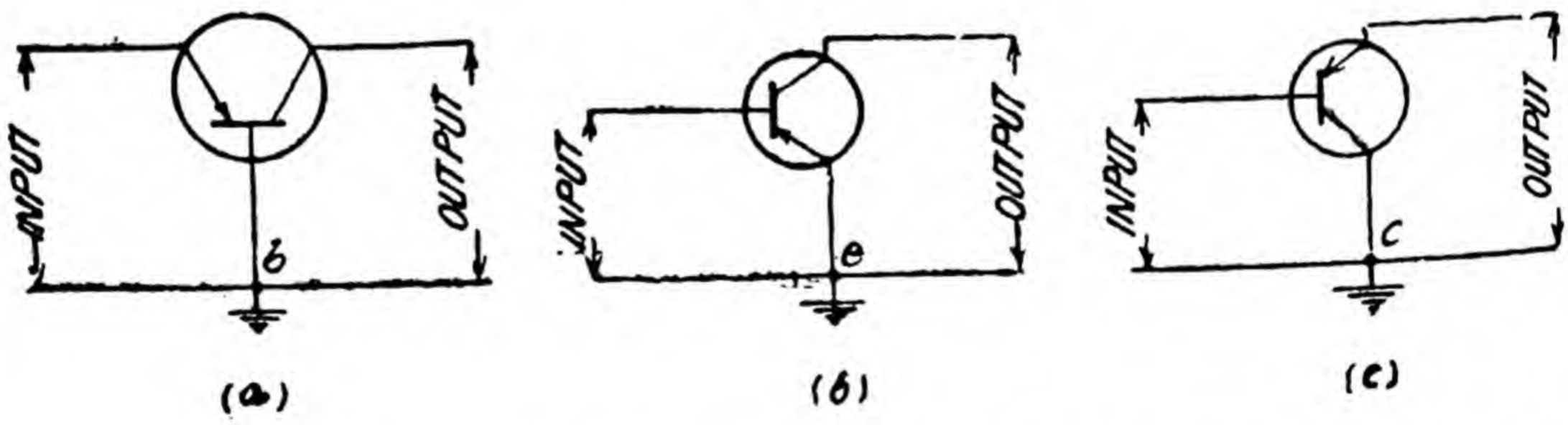


Fig. 21. Illustrating three methods of connection.

Although there are several ways of plotting the static characteristics of a transistor, but the most commonly used characteristics are those which are drawn in collector current *versus* collector voltage, one for constant values of emitter current and the other for constant values of base current. The first family of curves is generally used for grounded-base connected transistors, while the second is generally used for grounded-emitter connected transistors.

Common Base Configuration Static Characteristics :

Figure 22 shows the circuit arrangement for determining the static characteristics of a *PNP* transistor in common base configuration.

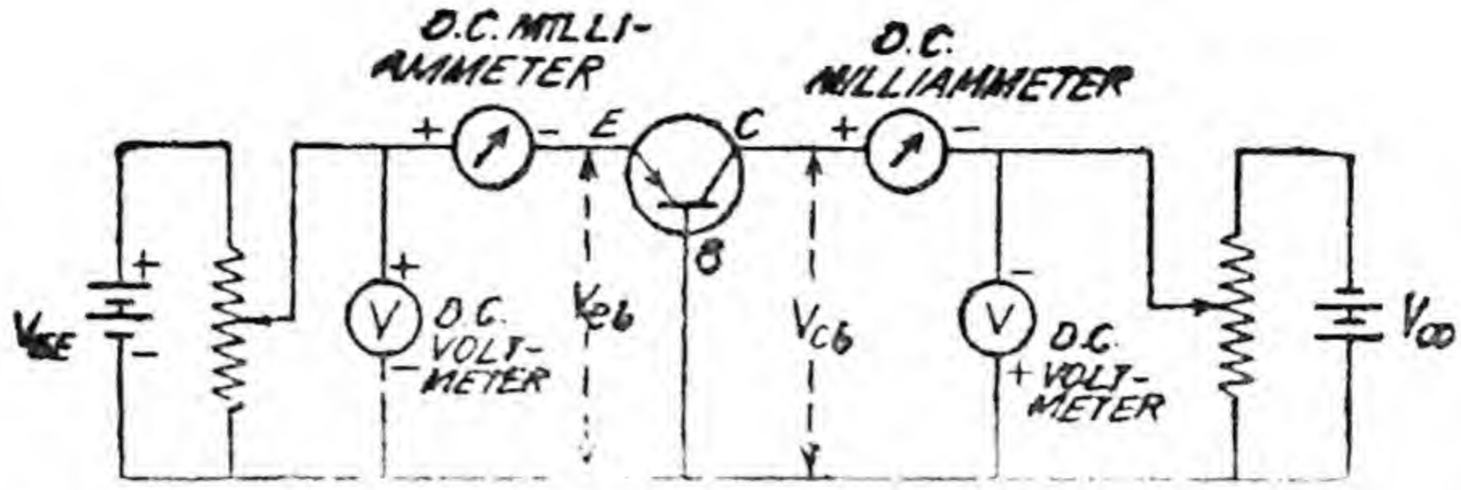


Fig. 22. Circuit arrangement for static characteristics of *PNP* transistor under common base configuration

The curves of collector current I_c *versus* collector voltage V_{cb} or V_c for *PNP* transistor are shown in figure 23.

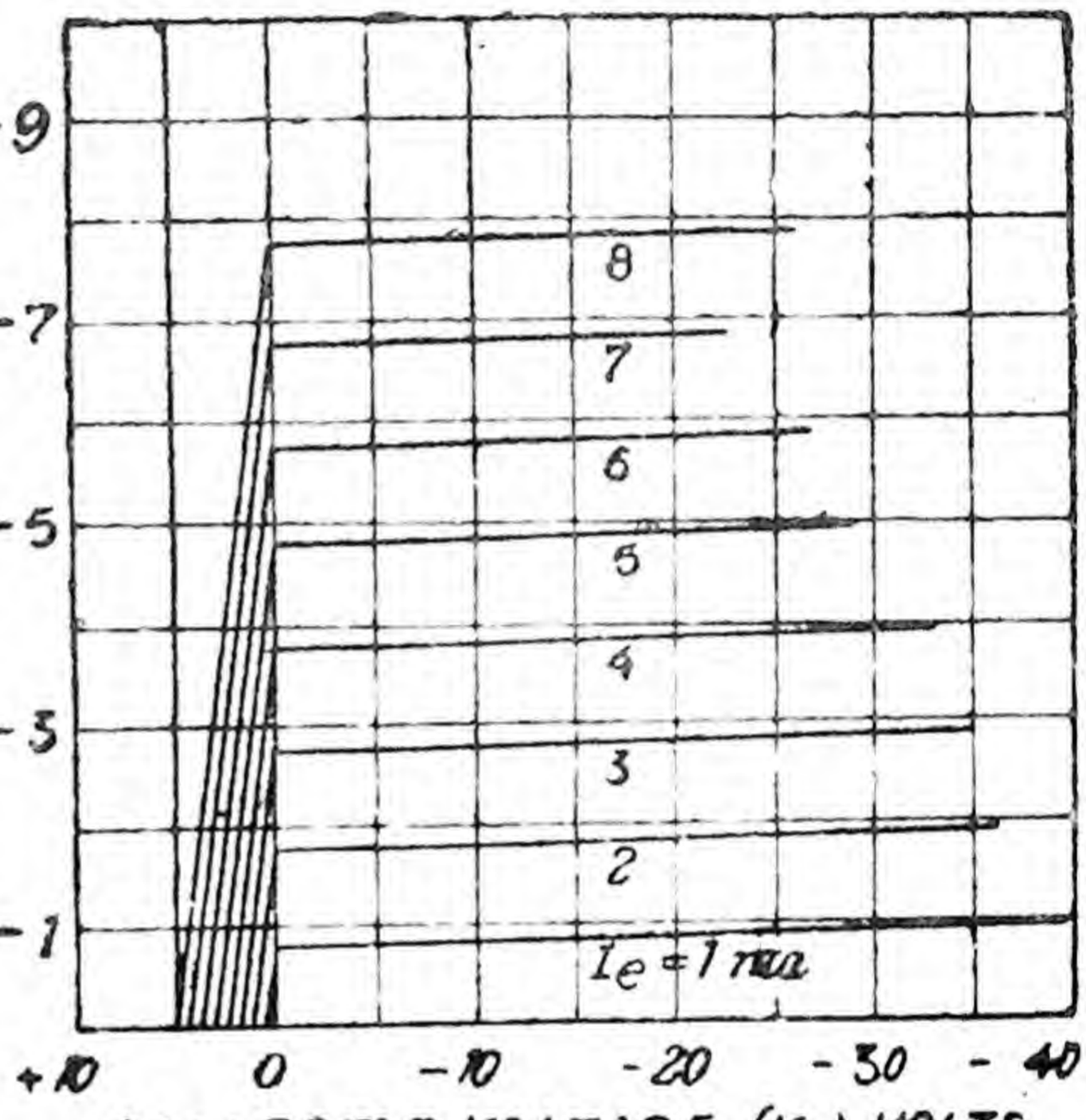


Fig. 23. Static common base collector characteristies of a triode junction transistor.

We see from the curves of figure 23 that collector current in a grounded-base connected transistor is relatively independent of collector voltage and primarily depend upon emitter current This behaviour is similar to that of pentode vacuum tube curves. If the collector voltage is reversed, then we see that for its very low values, the collector current is dropped sharply to zero; but if collector voltage is increased further, then collector current rapidly reverses itself and starts increasing in forward direction. If the resulting current becomes excessive, the transistor may be damaged.

Common Emitter Configuration Static Characteristics :

Figure 24 shows the circuit arrangement for determining the static characteristics of a *PNP* transistor in common emitter configuration.

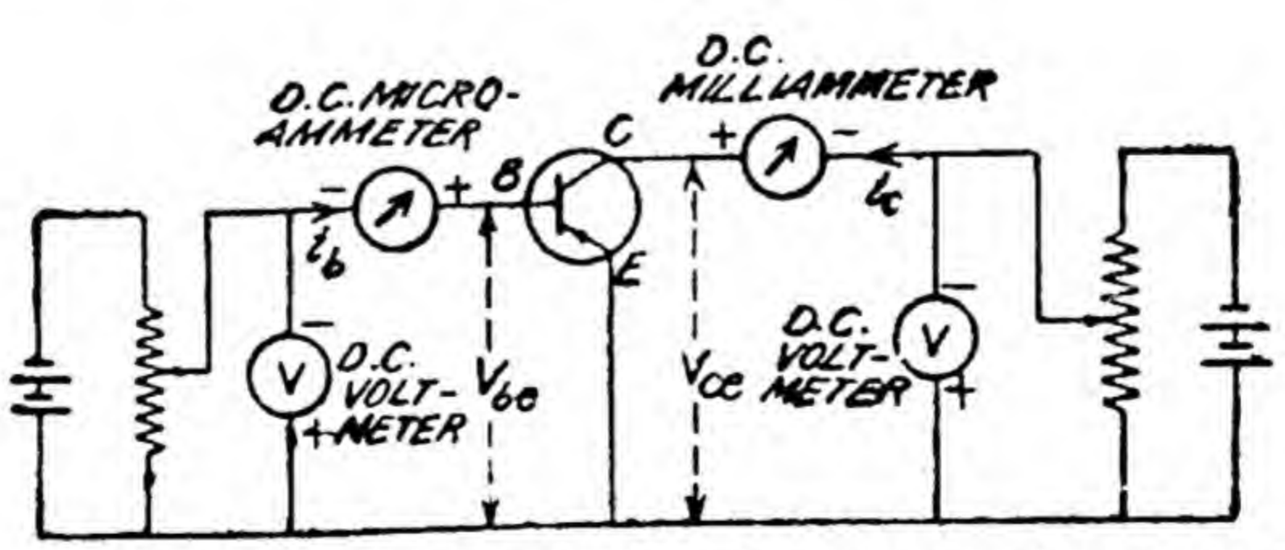


Fig. 24. Circuit for determining the static characteristics of a *PNP* transistor in common emitter configuration.

The output characteristics of a *PNP* transistor under common emitter configuration are shown in fig. 25, in which collector

current is plotted against collector to emitter voltage V_{ce} or V_o simply using the base current as parameter.

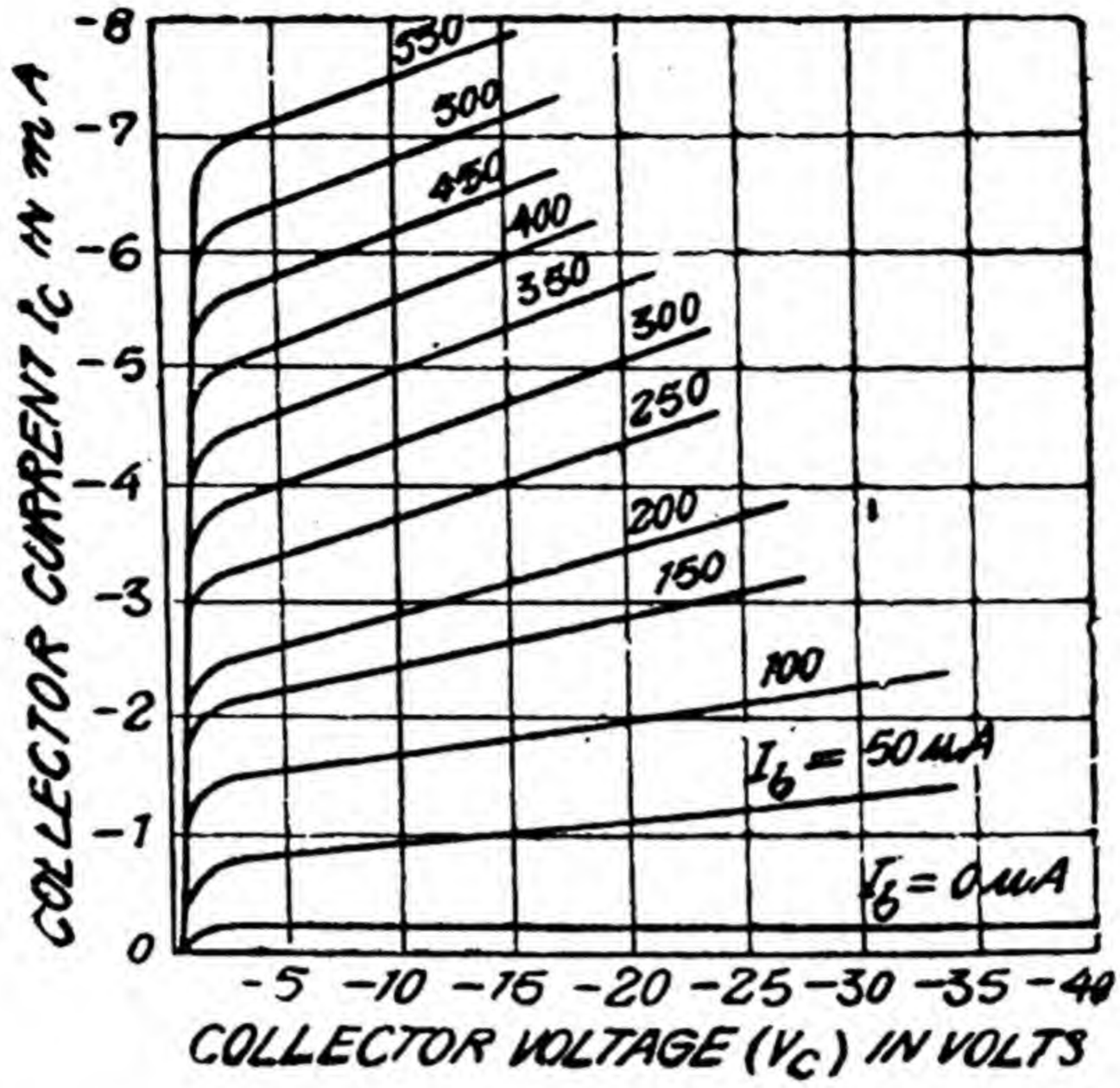


Fig. 25. Output characteristic of a *PNP* transistor under common emitter configuration.

It is observed from the figure 25 that for small values of base current, the effect of collector voltage over collector current is small while for large base current values this effect increases. Secondly, these characteristics have the same general pattern as the output characteristics in the common base configuration. However, the output current *i.e.* collector current is larger than input current *i.e.* base current. Thus the current gain in this configuration is greater than unity.

6·18. IMPORTANCE OF CHARACTERISTICS :

The characteristics are useful in calculating the current gain of a transistor circuit. The current gain of a common base transistor circuit is defined as ratio of the change in collector current to the change in emitter current for constant value of collector voltage. It is denoted by a symbol α called alfa direct current gain when specified for static or large emitter and collector currents.

$$\text{So } \alpha = \frac{\text{change in the collector current}}{\text{change in emitter current}} \text{ for constant collector voltage}$$

$$= \left(\frac{dI_c}{dI_e}\right)_{V_o}.$$

The value of α can be estimated from curves of fig. 23. It is seen that α is always less than 1 for the junction transistors.

For common emitter circuits the current gain is defined as the ratio of the change in collector current to the change in base current for a constant value of collector voltage. It is called the beta direct current gain and is denoted by β for static or large values of collector currents and base currents.

$$\therefore \quad \beta = \left(\frac{dI_c}{dI_b}\right)_{V_c} \text{ constant.}$$

The value of β is always greater than unity and can be estimated from figure 25.

By definition,

$$dI_c = \alpha\ dI_e.$$

We know that the emitter current is equal to the sum of collector and base currents

$$\therefore \quad dI_e = dI_c + dI_b$$
$$= \alpha\ dI_e + dI_b$$

or $$dI_b = dI_e - \alpha\ dI_e = (1-\alpha)\ dI_e.$$

Dividing by dI_c on both sides,

$$\left(\frac{dI_b}{dI_c}\right)_{V_c} = (1-\alpha)\left(\frac{dI_e}{dI_c}\right)_{V_c}.$$

or $$\frac{1}{\beta} = \frac{1-\alpha}{\alpha}.$$

$$\therefore \quad \beta = \frac{\alpha}{1-\alpha}.$$

6·19. LOAD LINE :

A common emitter amplifying stage is shown in figure 26 (*a*). R_L is the load resistor into which the transistor is working. Now writing in a conventional manner with terminal indicated by first subscript as positive, we have

$$I_c R_L + V_{EC} = V_{CC}$$

or $$I_c = -\frac{V_{EC}}{R_L} + \frac{V_{CC}}{R_L}, \qquad \ldots(1)$$

where V_{CC} denotes the supply voltage and V_{EC} the voltage across emitter and collector.

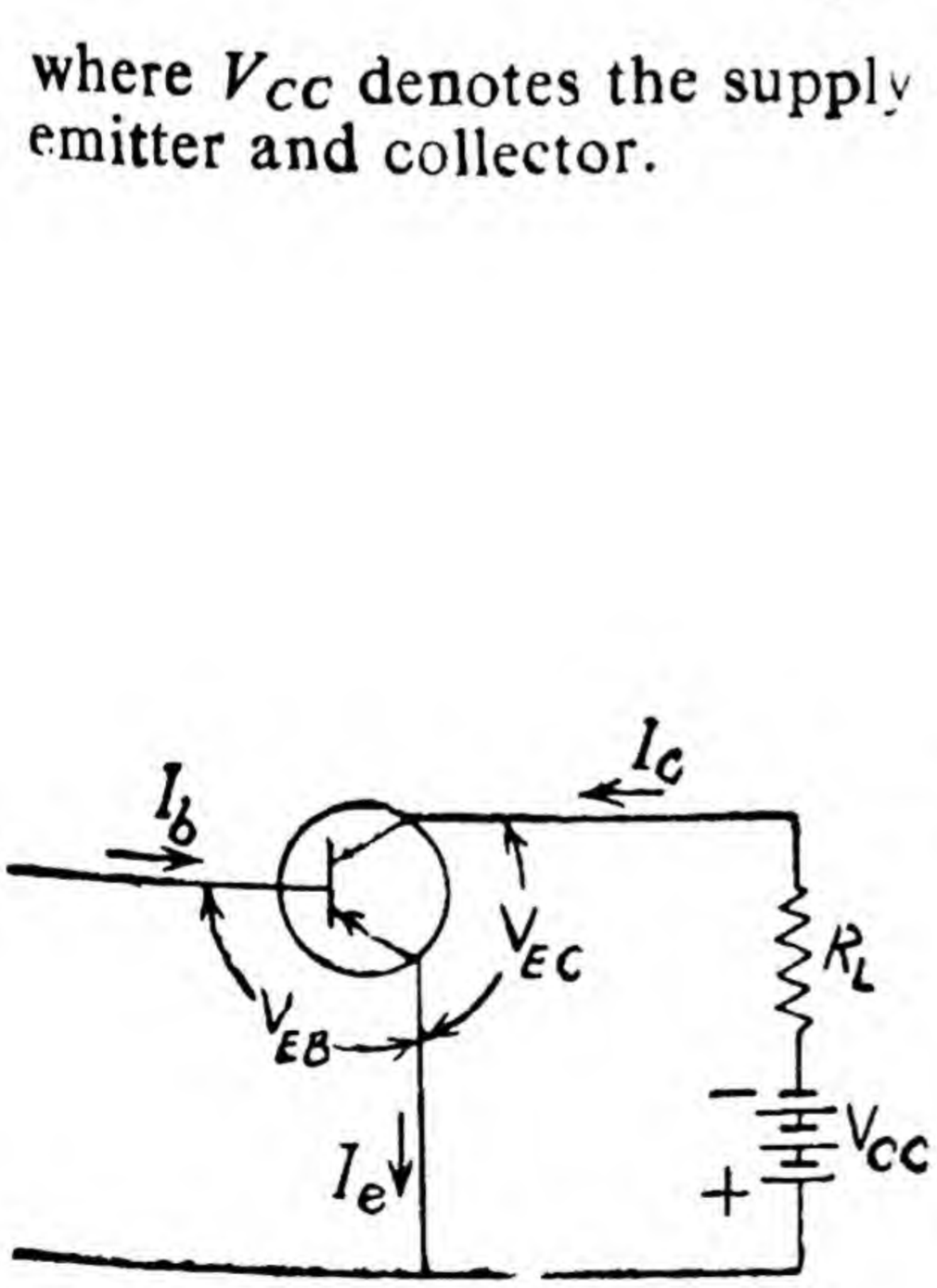


Fig. 26. (a) Simple common emitter circuit.

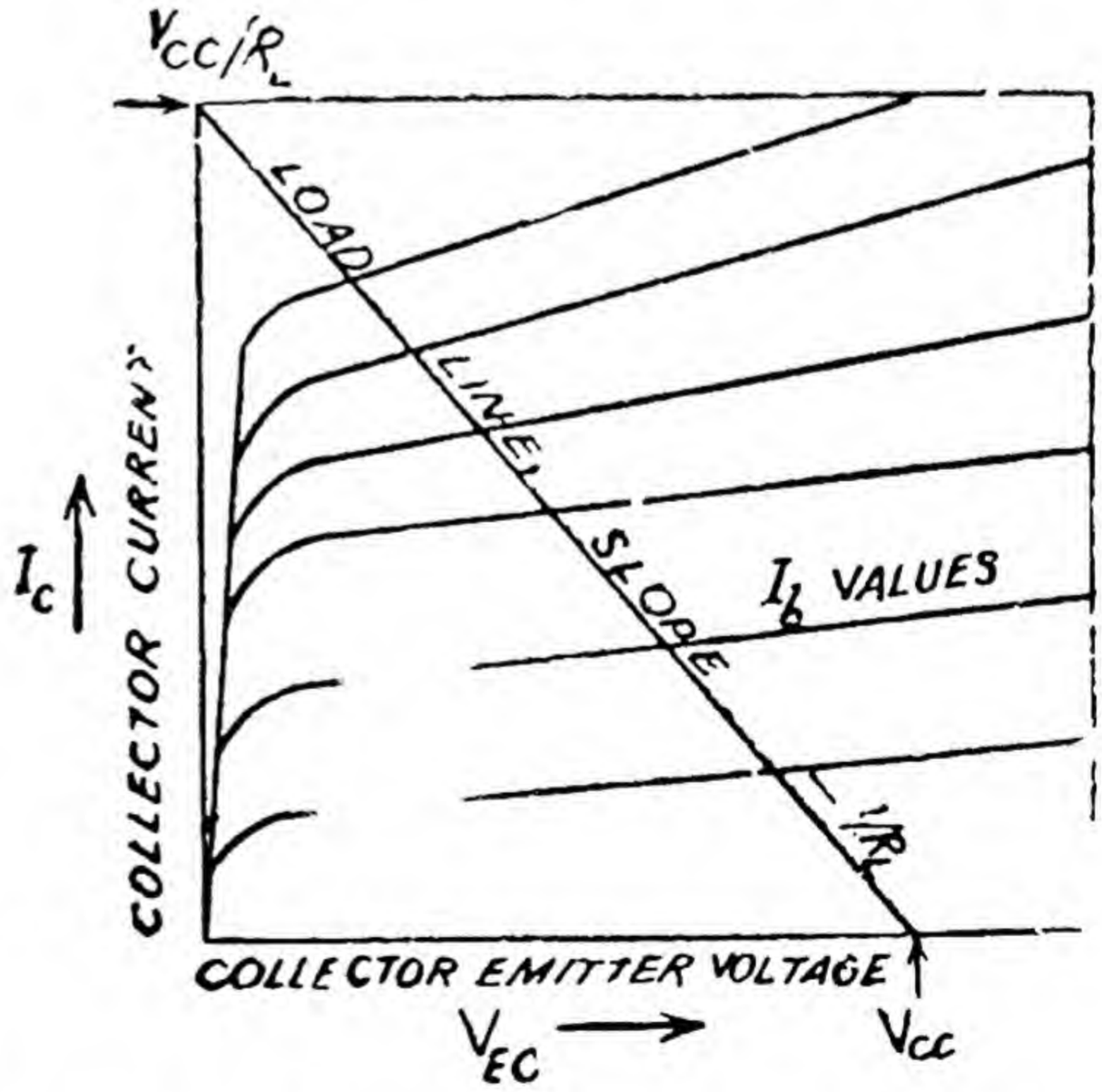


Fig. 26. (b) Load line plotted on output characteristics.

Common emitter configuration characteristics in which I_c is plotted against V_{EC} for different base currents are shown in figure 26 (*b*). The straight line represented by equation (1) is traced in this figure, intercepting V_{EC} axis at V_{CC} and I_c axis at V_{CC}/R_L. The slope of this straight line is $-1/R_L$. Such a straight line is known as *load line* and its intersection with output characteristics gives a set of quiescent operating points for the device.

Selection of the operating point is done according to the use to which the device is put. For small signal amplifiers, in which power is conserved, operating point should be so chosen as to give lowest quiescent value of collector current, while for an amplifier required to deliver sufficient amount of power, operating point is chosen so as to give quiescent current about one half of the maximum permissible collector current of the transistor.

Example : *Figure* 26 (*c*) *shows the collector characteristic of a common base amplifier. The collector battery voltage* V_{CC} *is* 16*V and load resistor is* 400Ω.

(*i*) *Draw the d.c. load line and locate operating point for* $v_{EE}=1{\cdot}5V$ *and* $R_E=60\Omega$.

(*ii*) *If the output is taken across a load* $R_0=1K$ *as shown in fig.* (26*d*) *draw the a.c. load line.*

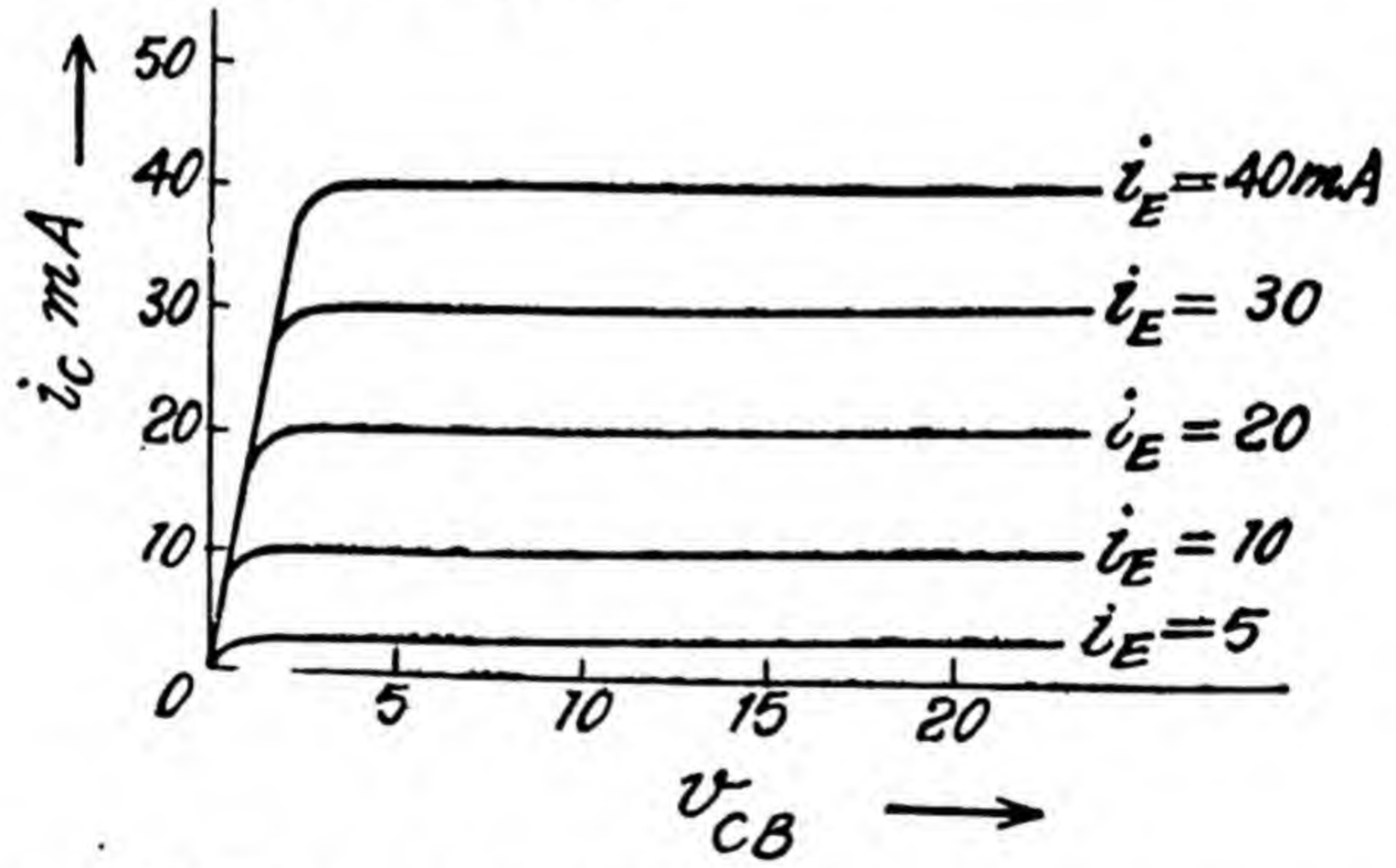


Fig. 26. (c) Collector characteristic of a common base amplifier.

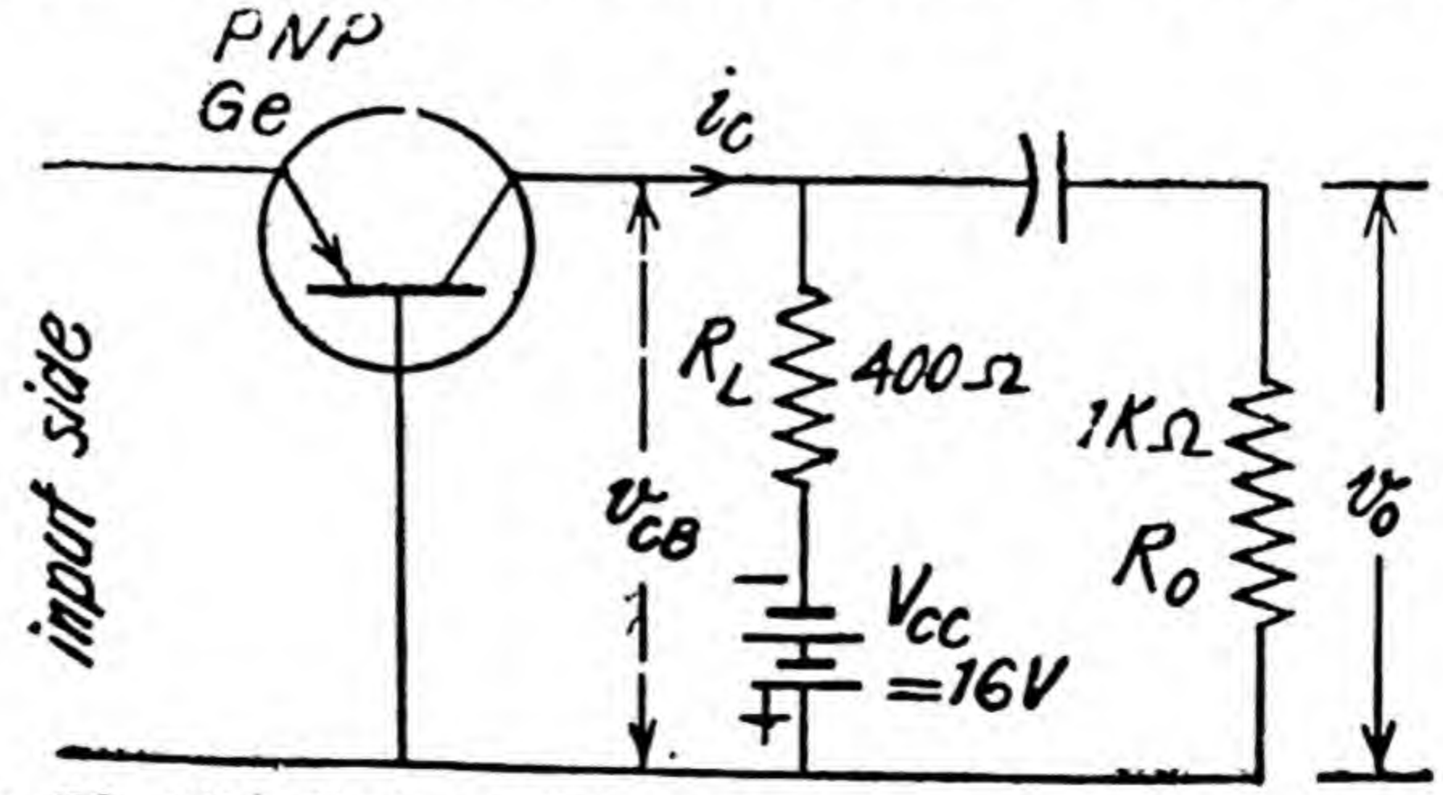


Fig. 26. (d) Output side of a common base amplifier.

(i) d.c. Load line : Here $v_{CB}=i_C R_L-V_{CC}$

One end of the d.c. load line is where $i_C=0$

$$i.e., \qquad v_{CB}=-V_{CC}=-16V$$

Now we plot this point as $i_C=0$ and $|\,v_{CB}\,|=16V$.

The second point is located where $v_{CB}=0$

$$i.e., \; i_C=V_{CC}/R_L=16/400=40\ mA.$$

A line passing through these points will be d.c. load line. This is shown in fig. (26e).

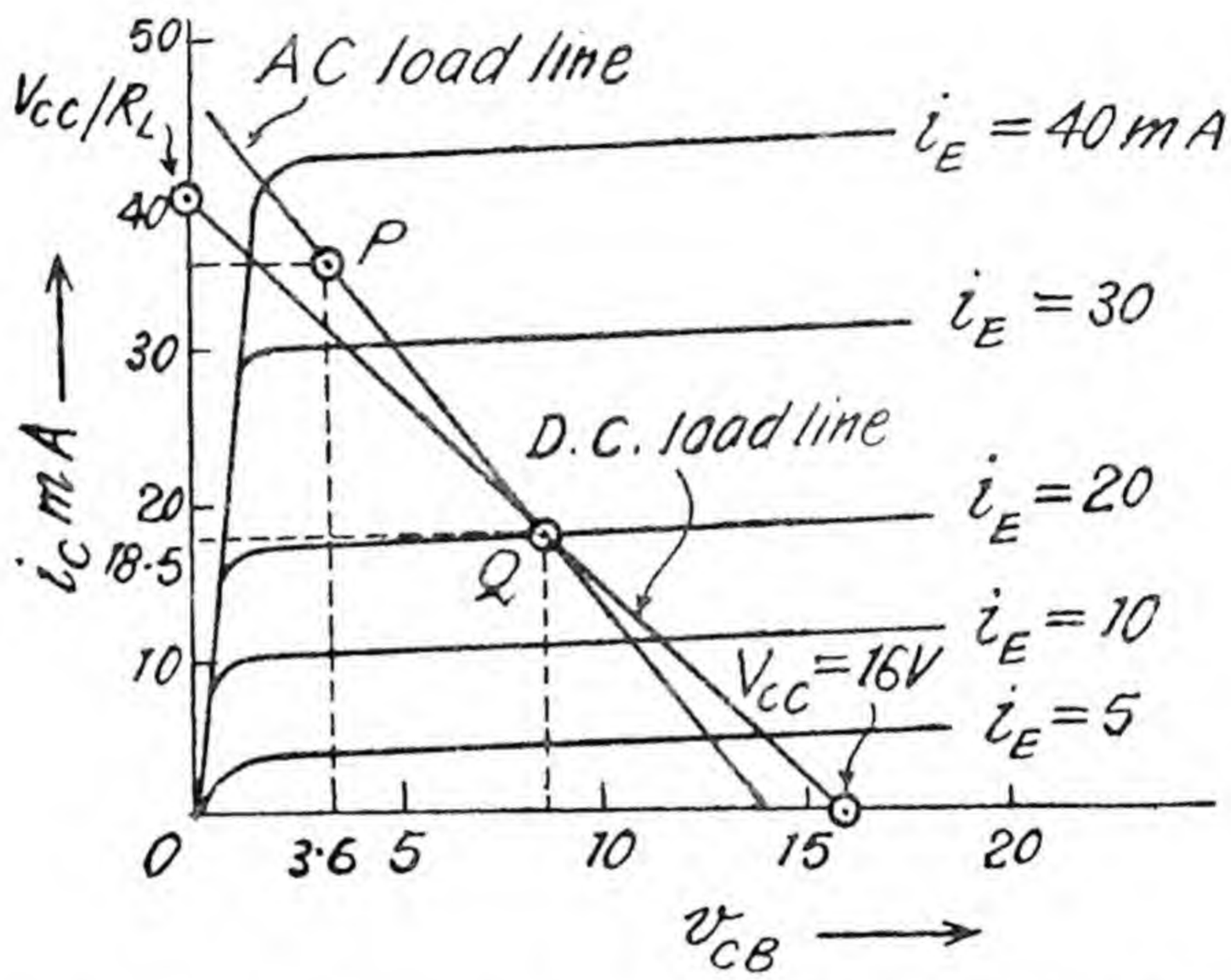


Fig. 26 (e) Showing the d.c and a.c. load lines.

For locating the operating point, we write the equation fo input circuit.

$$V_{EE}=I_E R_E+V_{EB}$$

In germanium, V_{EB} is approximately 0·3 V

$$\therefore \quad 1{\cdot}5=I_E\times 60+0{\cdot}3$$

Solving we get $\qquad I_E=20\ mA.$

So the operating point Q is selected corresponding to $i_E=20$ mA characteristic where the d.c. load line intersect it. Correspondin to operating point Q, $v_{CB}=8{\cdot}6\ V$ and $i_C=18{\cdot}5$ mA.

(ii) a.c. load line : In this case a load resistance $R_o=1\ K\Omega$ is connected at the output. Assuming that reactance of C_2 is negligible, the output a.c. load will be a parallel combination of R and R_L *i.e.* equivalent resistance R_e is given by

$$R_e=\frac{R_L R_O}{R_L+R_O}=\frac{400\,\Omega\times 10^3\,\Omega}{400\,\Omega+10^3\,\Omega}=286 \text{ ohm.}$$

This load demands a new load line through the operatin point Q with a different slope. The slope is given by $1/R_e=1/28$ ohm or 0·0035 amps. per volt because $1/R=I/V=$amp./volt.

Now for each 5 volts, the slope is equivalent to $0{\cdot}0035 \times 5$ amp. or 17·5 mA. So to draw a.c. load line, we start at the Q point on the characteristic curve (8·6 V, 18·5 mA) and move to left 5 volts and vertically upwards 17·5 mA. Now the new point will be $8{\cdot}6\ V - 5V = 3{\cdot}6$ volts and $18{\cdot}5 + 17{\cdot}5 = 36\ mA$. This point is shown by P in fig. (26 e). The straight line passing through P and Q will be the a c. load line.

6·20. CLASSIFICATION OF SEMICONDUCTOR DEVICES :

The semiconductor devices may be classified as follows :

(1) Semiconductor diodes :

(a) Junction diodes : (i) Grown junction diodes, and
(ii) Fused junction diodes.

(b) Point contact diodes :

(2) Transistors :

(c) Junction transistors : (i) Grown junction transistors, and
(ii) Fused junction transistors.

6·21. MANUFACTURE OF GROWN JUNCTION DIODES AND TRANSISTORS :

For diode and transistor fabrication, single crystals with controlled amounts of impurity are used. Crystals are grown by touching a seed crystal to the surface of liquid semiconductor material and then slowly withdrawing it as new lattice layers form by freezing out of the melt. A crystal growing furnace is shown in figure 27. The limitation of this process is that impurity concentration increases as more of the melt is used up, because impurities tend to remain in solu ion.

In order to record the temperature of the furnace, a thermocouple is used and thus the temperature can be very closely controlled. A large crystal is grown from a rod of polycrystalline germanium feeding into melt. The purified polycrystalline setmiconductor is placed in a chamber and heated in an atmosphere of hydrogen and nitrogen to prevent oxidation. There is a vertical shaft which can be rotated as well as pulled up and the seed crystal is attached to it such that it touches the molten semiconductor. As the seed crystal is gradually pulled up, the crystal starts growing and the rotation produces agitation of the melt which results in symmetrical crystal.

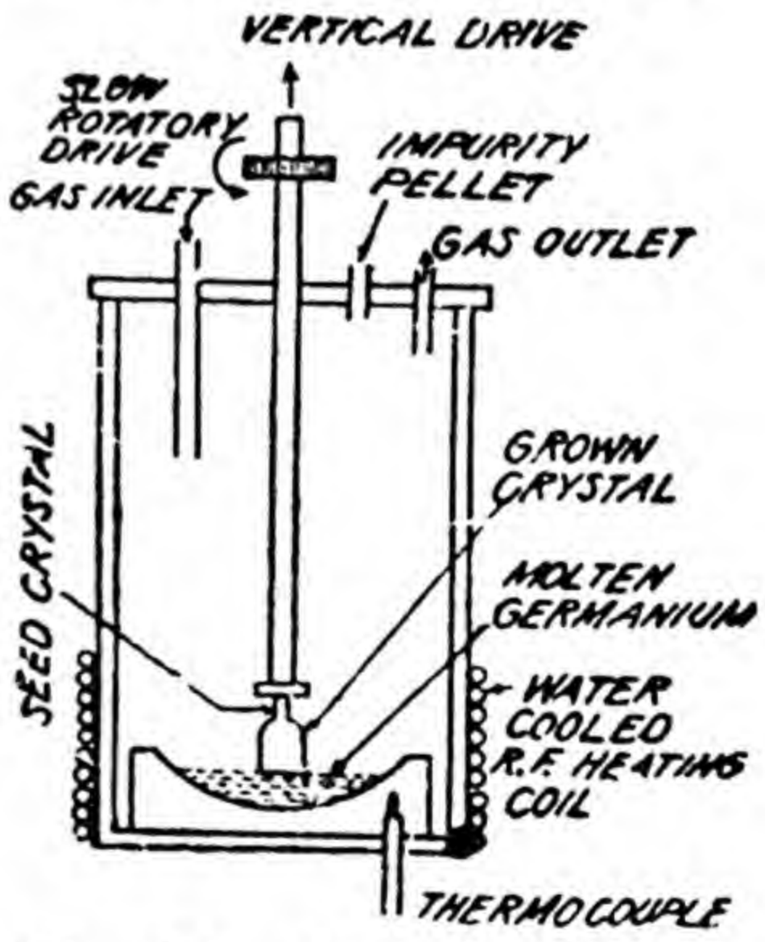


Fig. 27. Furnace used for growing single large crystal.

In order to form junction diode or transistor, two methods are used.

(i) impurity variation method, and
(ii) speed variation method (or growth rate variation method).

(*i*) *Impurity variation method :*

In this method, the impurity content of the metal is altered in its type as well as quantity. For example, in making NPN germanium grown junction transistor, first of all a small amount of *N* type impurity is added to molten germanium. The shaft is lowered suitably such that the seed crystal be dipped into the melt and crystal growth starts. The crystal which now grows is of high resistivity *N* type due to the low impurity donor concentration and forms the collector part of *NPN* transistor. Now keeping the pulling continuous, enough *P* type impurity is added to the melt to overbalance the *N* type impurity. This *P* type region forms the base and is kept very thin as the seed crystal is pulled for a very short distance. Next, the impurity is again made *N* type by adding a large amount of *N* type impurity to the melt. The part of the crystal that now grows is of low resistivity *N* type and forms the emitter part.

(*ii*) *Speed variation method* (*or growth variation method*) :

The grown junction may be formed by suddenly varying the rate of pulling the seed crystal from the melt. For example, arsenic (*N* type impurity) crystallises such that its fraction in the grown semiconductor crystal increases with the increase in the rate of pulling. On the other hand, the fraction on indium or gallium (*P* type impurity) in the grown crystal increases with the dec ease of pulling rate. When the germanium melt contains proper amounts of impurities (*N* and *P* types) and seed crystal is pulled slowly, a *P* type crystal is grown and if seed crystal is pulled fast then crystal grown is *N* type. In this way *PN* junction can be formed simply by altering the rate of pulling the crystal.

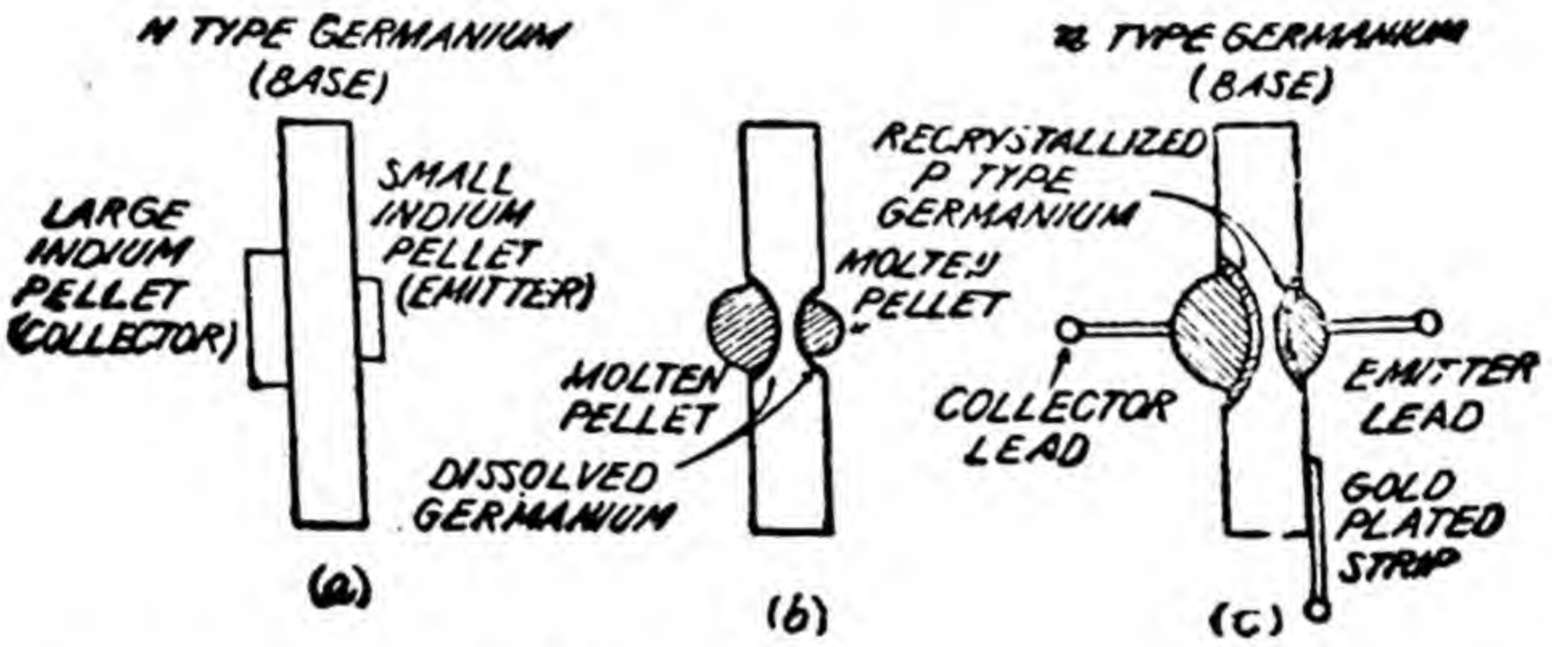
In order to prepare *NPN* transistors, impurities of both types (*i.e.* arsenic or antimony and indium) are added in proper quantities. First of all the rate of pulling is kept higher, then slowed down for a short interval and again made high. In this way *PNP* transistor is formed. The crystal so formed is cut into rods along its length. Each rod contains two junctions and constitutes one grown junction. Now non-rectifying electrodes are deposited at each end and wires are attracted to end electrodes to form the emitter and collecto. leads.

6·22. MANUFACTURE OF FUSED JUNCTION OR ALLOY JUNCTION DIODES AND TRANSISTORS :

Fused junction transistors are generally of *PNP* type. First of all the semiconductor is refined to a high degree and then a large crystal is grown from this polycrystalline crystal as discussed in the previous article. The impurity added in case of *PNP* transistor is of *N* type. A crystal of *N* type is formed by controlling the temperature and rate of seed crystal. The crystal so formed is

cut into thin wafers or plates by thin diamond cutting wheels of a size $0{\cdot}1'' \times 0{\cdot}1'' \times 0{\cdot}01''$ thick. Each such plate is then made into a semi-conductor or a transistor. When it is made into *PNP* transistor, it acts as a base. After cutting, the plate is grounded, polished and etched. The cleaning agent, known as "etch" removes all contaminants and also surface irregularities caused by the cutting process.

To prepare a *PNP* fuse junction transistor, on one side of the slice a small pellets of indium and on the other side a large pellet (about 3 times of previous) are placed. The larger pellet are used as collector and the smaller as emitter as shown in fig. 28 (*a*). The assembly is heated to about 500°C in an atmosphere of hydrogen. The indium pellets (melting point 115°C) now melt while germanium does not melt as its melting point is about 500°C.



(*a*) Germanium plate with pellet before heating. (*b*) Molten indium pellets. (*c*) Formation of *P* type germanium.

Fig. 28. Formation *PNP* fused Junction transistor.

The molten indium dissolve some of the germanium from the plate and forms a bead of saturation solution as shown in fig. 28(*b*). On cooling, the two separate a little. As cooling proceeds molten germanium with indium content recrystallizes to form a crystal of *P* type germanium at the solid liquid interface On further cooling the indium pellet solidifies with a little germanium. Leads are soldered as shown in fig. 28 (*c*).

Fused junction diodes are manufactured exactly in a similar way except indium pellet is placed only on one side of the plate.

6·23. POINT CONTACT DIODE AND TRANSISTORS :

Point contact diode :

In case of a junction diode the barrier is very thin, and represents a capacitance between the two electrodes. The capacitance tends to shunt out the rectifying action at high frequencies. The point contact form provides a low capacitance diode suitable for operation at microwave, while the junction diodes are used for low frequency work.

Early diodes were of point contact form as shown in figure 29. The semiconductor wafer (fragment) is only a few millimeter

square and a fraction of a millimeter thick. The contact wire or catwhisker is usually phosphor bronze or tungston.

The electrical forming consists in placing a fine whisker or tungsten to the crystal surface and passing a pulse of large current through the system, generally by a discharge of a capacitor. The result of forming is to lower the forward resistance and raise the reverse resistance of the contact. During the passage of forming pulse some whisker material melts. The molten metal welds the whisker to the crystal surface which gives the contact good mechanical and electrical stability. The details of electrical forming are not understood. However, it is believed that a *PN* junction is formed at the contact because of melting of crystal surface and diffusion of whisker material into the surface at the point. Thus point contact diode and junction diode basically operate on the same principle.

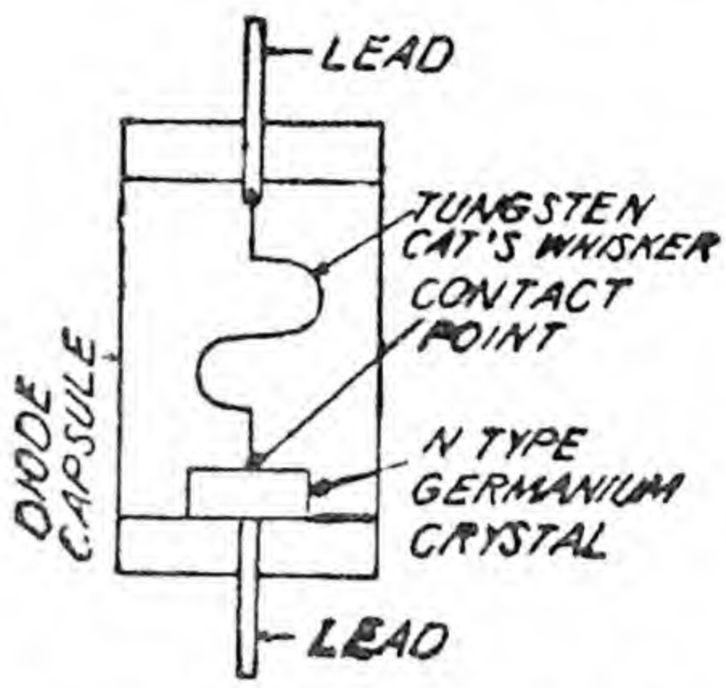


Fig. 29. Point contact semiconductor diode.

The detailed characteristics of a point contact diode depend on whisker composition, contact area, whisker pressure, crystal structure and crystal surface treatment.

Typical voltage-current curves for a point contact diode are shown in figure 30. It is observed that temperature has an important effect on the characteristics and there is sudden increase in current when the reverse voltage reaches the reverse breakdown voltage. The larger slope of the reverse current voltage characteristics of a point contact diode, at high reverse voltages, results because of local heating at the contact point.

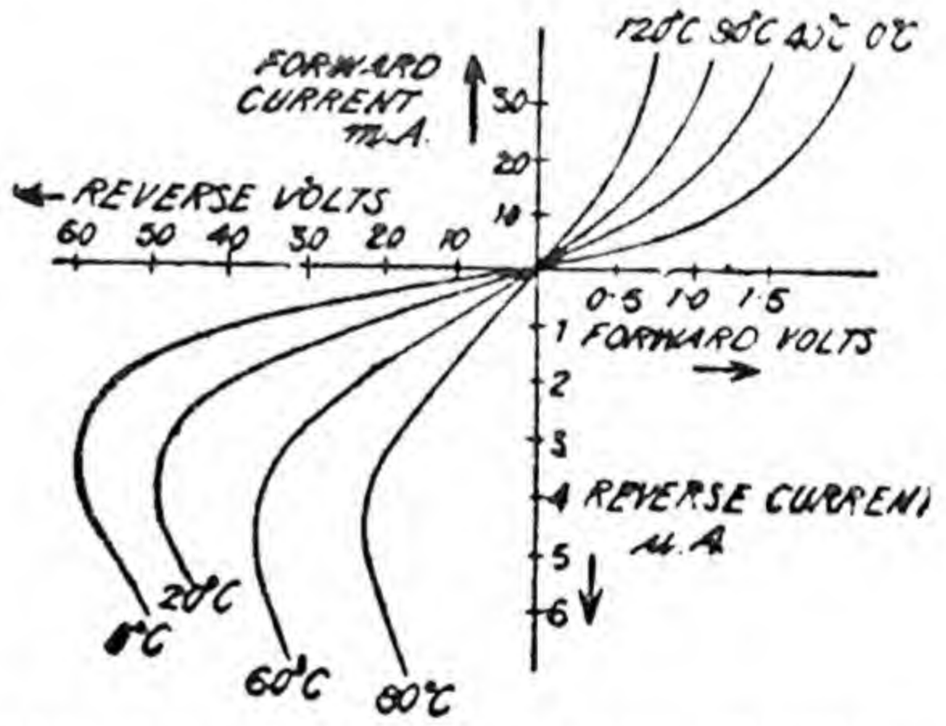


Fig 30. Current voltage characteristic of point contact diode.

Point contact transistors :—

A *PNP* point contact transistor consists of a crystal of *N* type germanium or silicon as base. The point contacts, one of tungston and the other of phosphor bronze (collector) are made on semiconductor base a few thousands on an inch apart as shown in figure 31. The contacts are formed in a similar way as in point contact diodes. The current may be passed either between each whisker and the base or it may be passed between two whiskers. In this case, the current gain exceeds unity which is not found in junction transistor, and is commonly between two and four.

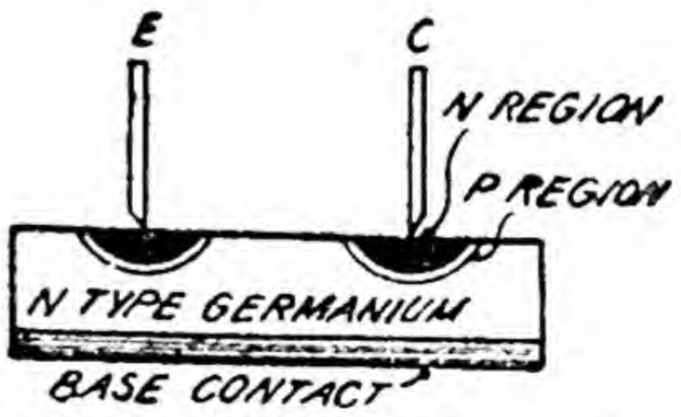


Fig. 31. Junction formed a *PNP* transistor.

9.29 THE TUNNEL DIODE :

A two terminal device that exhibits a region of incremental negative resistance is the *tunnel diode*, also referred to as the Esaki diode, after its inventer. They are usually made of germanium or gallium arsenide because of high electron mobility and reasonable gap energy. It is a junction diode. Both *P* and *N* regions in tunnel diode are heavily doped and consequently, the depletion region is extremely narrow. It is possible for carriers to tunnel through the potential barrier if it is narrow enough (typically 10^{-6} cm.) and if available energy level exists on the other side.

The current voltage characteristic of a typical tunnel diode is shown in figure 32. The right hand rising portion is the normal forward biased diode region of the device. The current variation in the vicinity of the origin is due to quantum mechanical tunneling of electrons through narrow space charge region of the junction. As the applied voltage is increased from zero, tunneling (current) first increases and then decreases towards zero. This decrease in current with increasing voltage results in a negative resistance region. When the forward voltage is further increased, the tunneling effect ceases and current increases in a manner similar to a conventional diode.

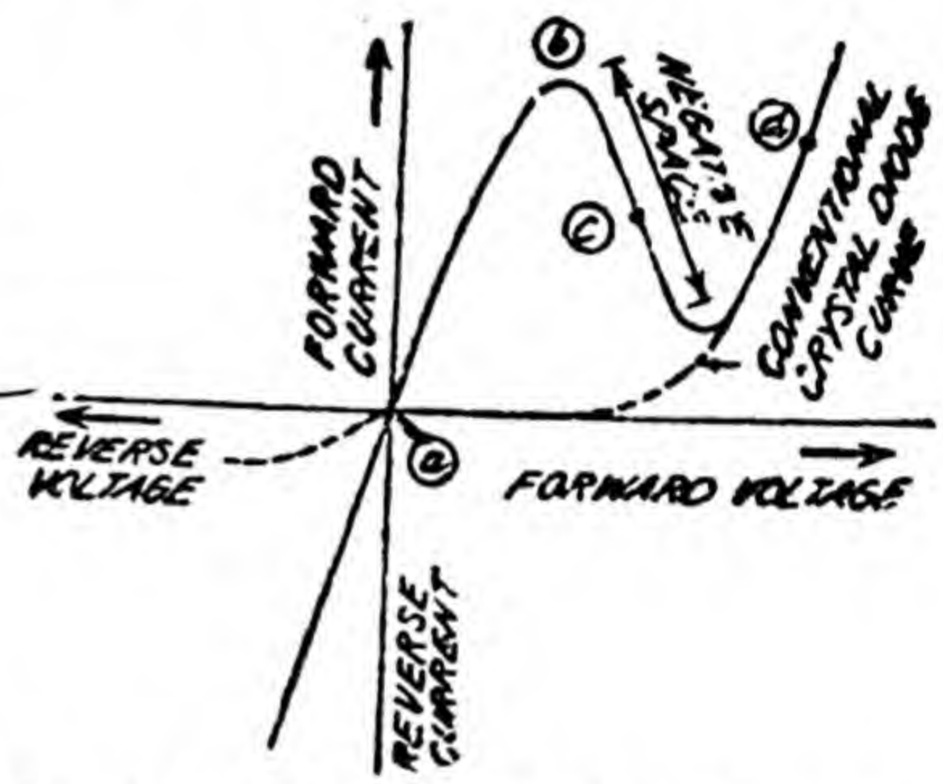


Fig. 32. Tunnel diode characteristic.

The physical operation of a tunnel diode will be considered with reference to the regions (*a*), (*b*), (*c*) and (*d*) of volt-ampere characteristic. The corresponding energy level diagrams are shown in fig. 33.

When the bias is zero, corresponding to point (*a*), it is observed that the tunneling is zero, because there is just the same probability of electrons going from states in the conduction band of *N* sides to states in the valence band of *P* side as in the opposite direction.

The current reaches maximum at point (*b*). From fig. 33 (*b*) it is observed that all the electrons in the conduction band see empty state just across the barrier and

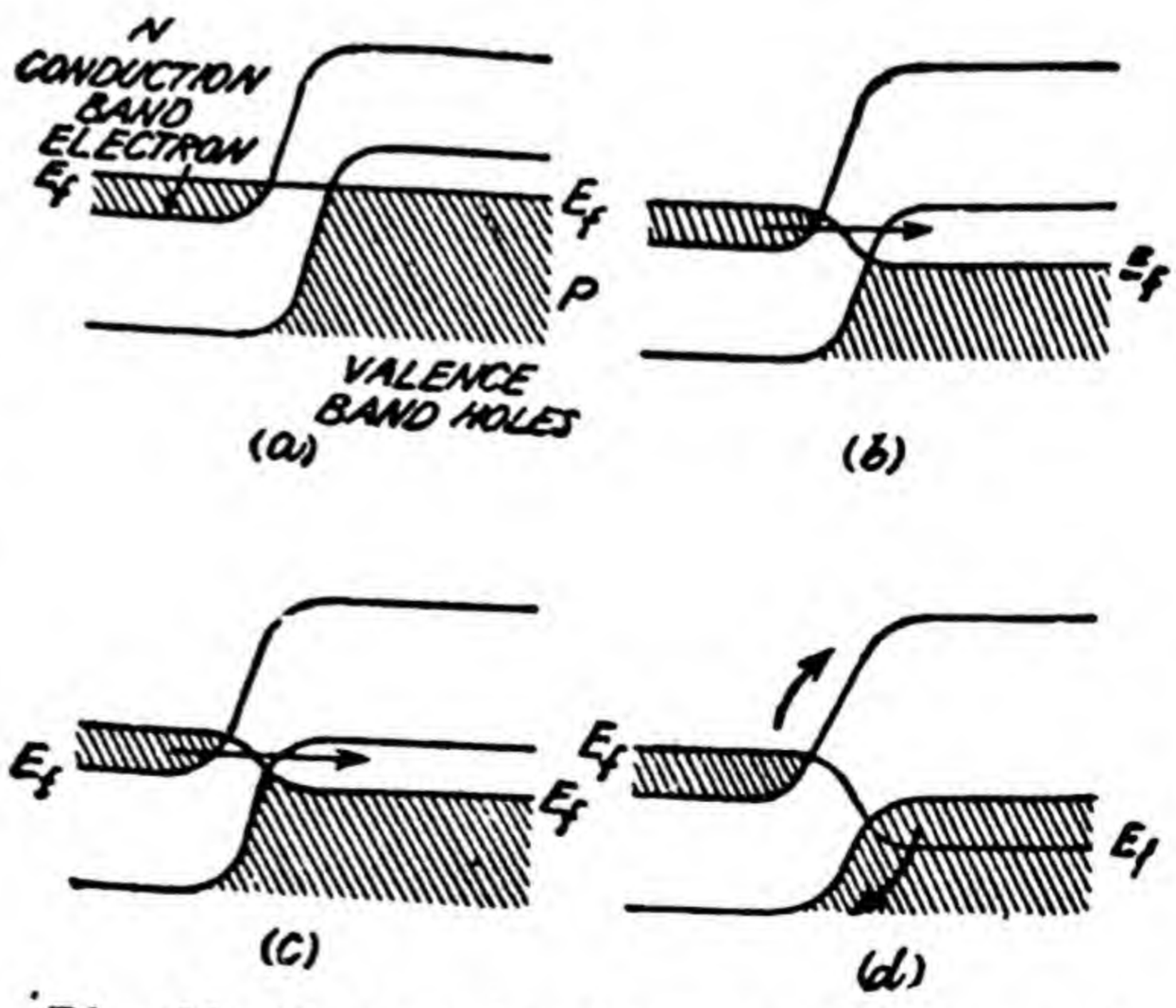


Fig 33. Energy level diagram of tunnel diode.

hence a large tunneling takes place. Again, the valence electrons in *P* types see forbidden levels to the left and can tunnel.

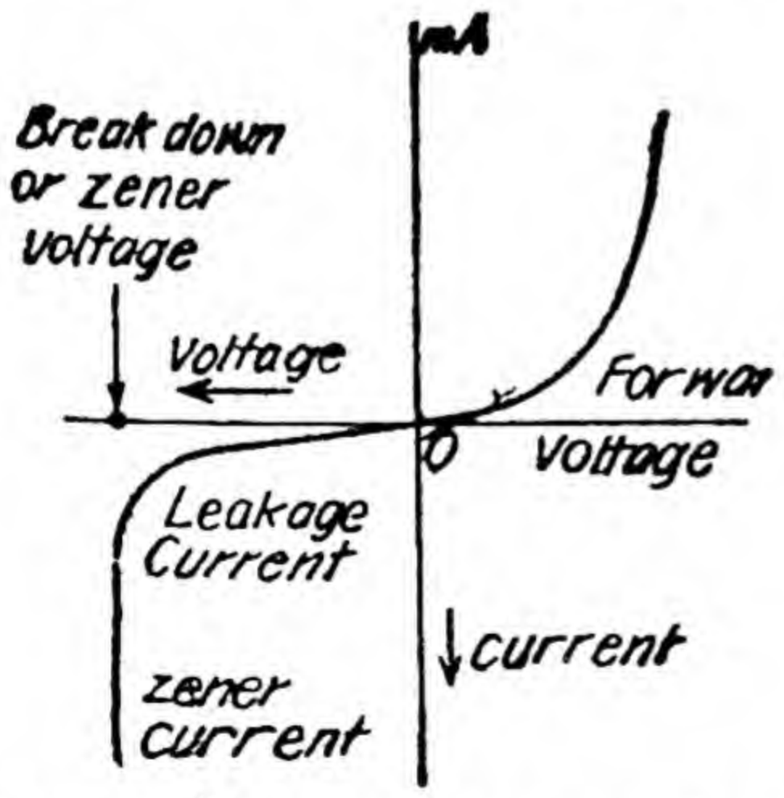
Corresponding to a point (*c*) when the forward bias is again raised, the conduction band electrons are raised to a level so high that only a part of them see available energy levels across the barrier. Ihus tunneling current is smaller than above.

Now with refcrence to point (*a*), the tunneling has been completely suppressed and ordinary injection currents are making up the current as is seen from fig. 33 (*d*).

Tunnel diodes are useful in high frequency circuits. There is a very little capacitance associated with the junction and hence the temperature effects are small.

6·25. ZENER DIODE OR BREAKDOWN DIODE :

Zener diode is a silicon crystal diode having an unusual reverse current characteristic which is particularly suitable for voltage regulating purposes. The characteristic is shown in figure (34*a*). It is obvious from the figure that as the reverse voltage applied to *PN* junction is increased, a value is reached at which the current increases greatly from its normal cut off value. This voltage is called a *zener voltage* or *breakdown voltage* and the reverse current *as zener current*. There are two mechanisms of break-down. According to the first mechanism, new carriers are liberated when thermally excited carriers acquire sufficient energy from applied potential and break the bonds. The newly generated carriers, in turn, generate carriers by further breaking the bonds. In this way current increases tremendously without increasing applied voltage. According to the second mechanism, the breakdown is initiated by stron g electric fields even if initially carriers with sufficient energy are not available. In case of zener diode, both breakdown exists simultaneously. The breakdown can be controlled by the an ount of impurity in the semiconductor, the nature of electrodes, and the forming treatment employed.



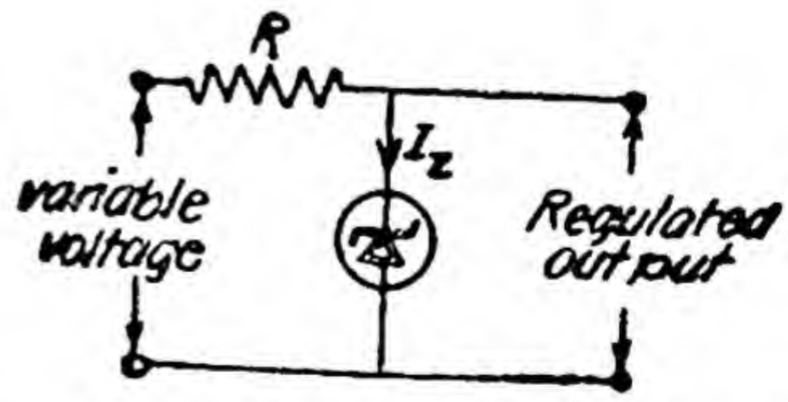


Fig. 34 (a) Characteristics of a zener diode. Fig. 34 (b) Regulating circuit.

The figure (34 *b*) shows a basic zener diode D.C. regulating circuit. The diode is selected in such a way that its breakdown voltage is equal to the desired regulating output. When the input

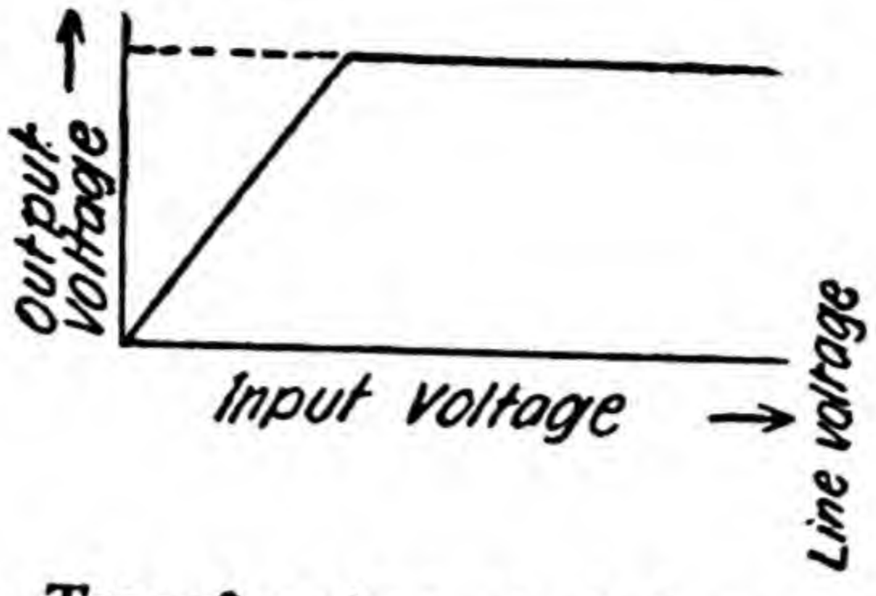


Fig. 34 (c) Transfer characteristic of zener diode.

voltage is less than the breakdown voltage, very small current flows through the diode and the whole input appears across the output terminals. But when the input voltage reaches the breakdown voltage, breakdown takes place and current through the diode is adjusted itself such that voltage across the zener diode remains constant. The transfer characteristic is shown in figure (34*c*).

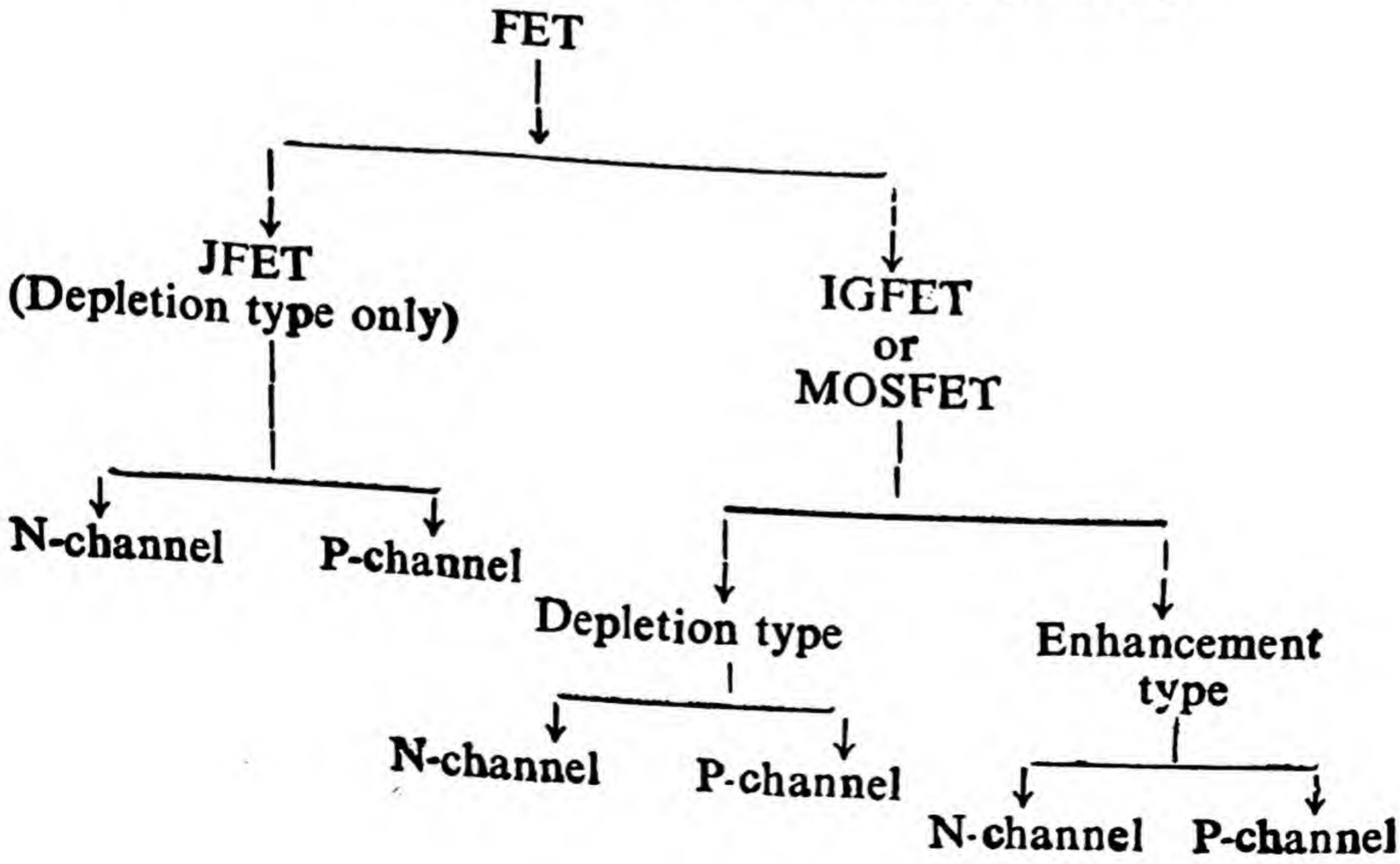
6·26. FIELD EFFECT TRANSISTOR (FET) :

The field effect transistor is a device in which the number of current carriers available in conducting region is controlled by the application of an electric field at the surface of semiconductor or *a field effect transistor operates on the principle that the thickness and hence the resistance of a conducting channel of semiconductor material may be regulated by the magnitude of a potential applied to its input terminals.* They are of two types :

(1) Junction field effect transistor (JFET or simply FET),

(2) Insulated gate field effect transistor (IGFET) or metal oxide semiconductor field effect transistor (MOSFET).

Thus we have the following classification of FET :



Differences in field effect transistor and junction type transistor :

FET's	Junction transistor
1. FET's have very high input resistance. It ranges from 10^8 to 10^{12} ohms in JFET and from 10^4 M. ohms to 10^7M ohms in MOSFET. Thus FET is a voltage control device.	Junction transistors have low input resistance. It ranges from 10^3 to 10^6 ohms. Thus it is a current controlled device.
2. FET is unipolar (can operate only in forward direction) *i.e.*, operation depends upon the flow of majority carriers only.	The operation depends upon the interaction of both the majority and minority carriers and hence they are named as bipolar.
3. Transconductance is low (10,000μ mhos) and hence voltage gain is not large.	Transconductance is very high so the voltage gain is high
4. Lessnoisy as compared to junction transistor.	More noisy.
5. Immune to radiation.	Sensitive to radiation.
6. Thermal stability is good.	Thermal stability is poor.
7. More costly than junction transistor.	Less costly.
8 The direction of arrow in gate terminal indicates the direction of gate current flow when gate junction is forward biased	The arrow in the emitter terminal points the direction of forward junction current.
9. The output voltage is linearly related to input voltage when used as small signal amplifier. With large signal non-linear distortion appears.	The major application in the field of electronics is as amplifier.
The FET has a relatively small gain bondwidth product.	

6 26-1. JUNCTION FIELD EFFECT TRANSISTOR (JFET)

The field effect transistor is a single impurity semiconductor,

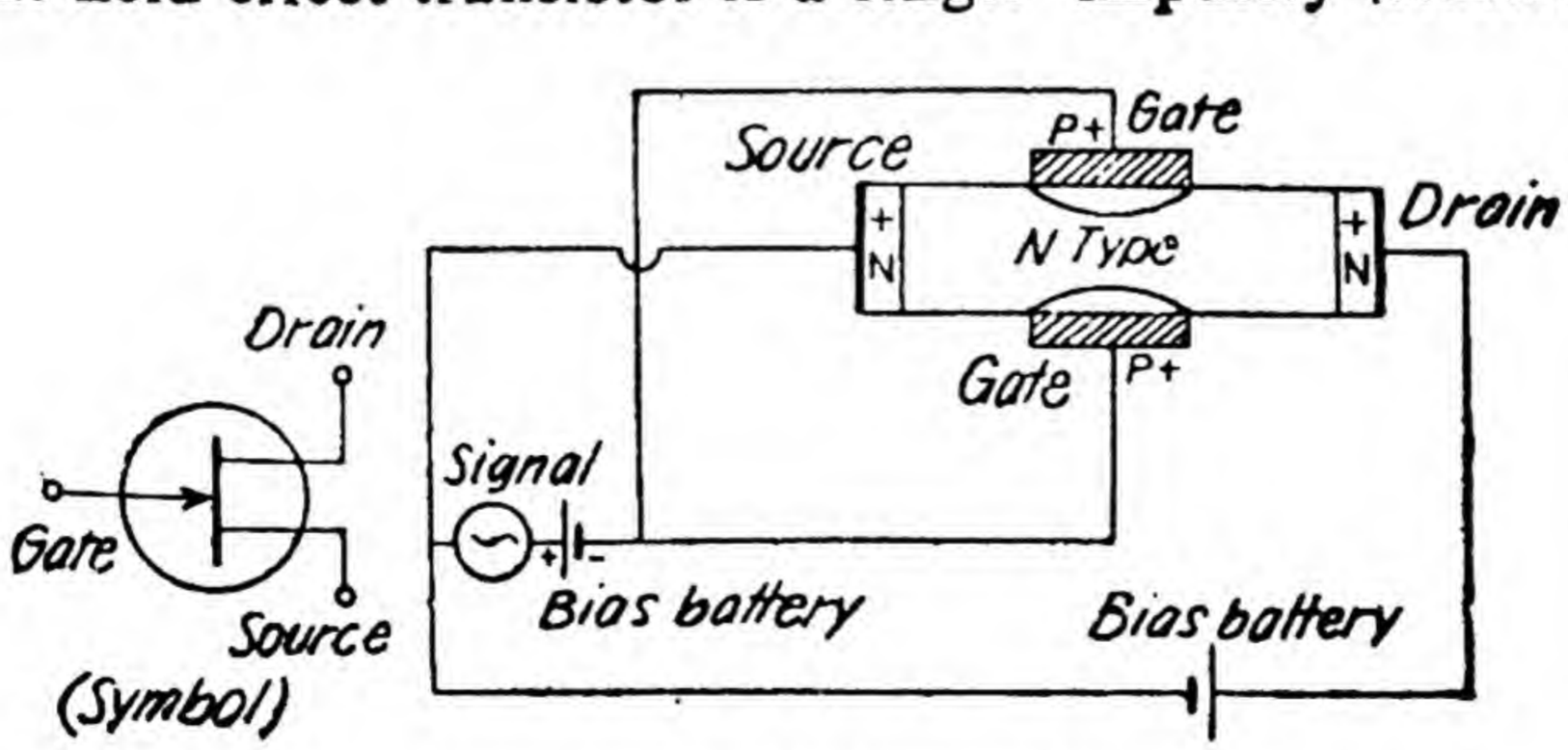


Fig. 35. An *N*-channel JFET.

to which impurity material of opposite type is formed on both the sides. The construction is shown in figure (35). The basic material is *N*-type to which heavily doped *P* material (denoted by *P*+) is added. The major *N* type area forms the channel. The ends of the semiconductor have heavily doped area (*N*+). The current input side of the semiconductor is known as source (comparable to vacuum tube cathode) and opposite end as drain (comparable to plate of a vacuum tube). The opposite impurity control regions are known as gate (like a grid in vacuum tube). Between the gate and the source a voltage is applied in a direction which is reverse bais for the *PN* junction. A voltage is also applied between source and drain so that current flows along the length of the transistor.

Circuit symbol and notation. The circuit symbol and polarity notations for an *N* channel FET are shown in fig. (36 *a*). It should be noted that the direction of the arrow in the gate terminal indicates the direction in which gate current would flow if the gate junction were forward biased. When a *P*-type semiconductor bar is used, the gate is formed by making heavily doped *N* regions on both side of the bar. Such FET is called a *P*-channel FET. The circuit symbol and polarity notations for a *P*-channel FET are shown in fig. (36 *b*). It can be observed from fig. (36 *b*) that the polarity of voltages are reversed as compared to *N*-channel FET and thus I_D and N_{DS} are negative and V_{GS} is positive.

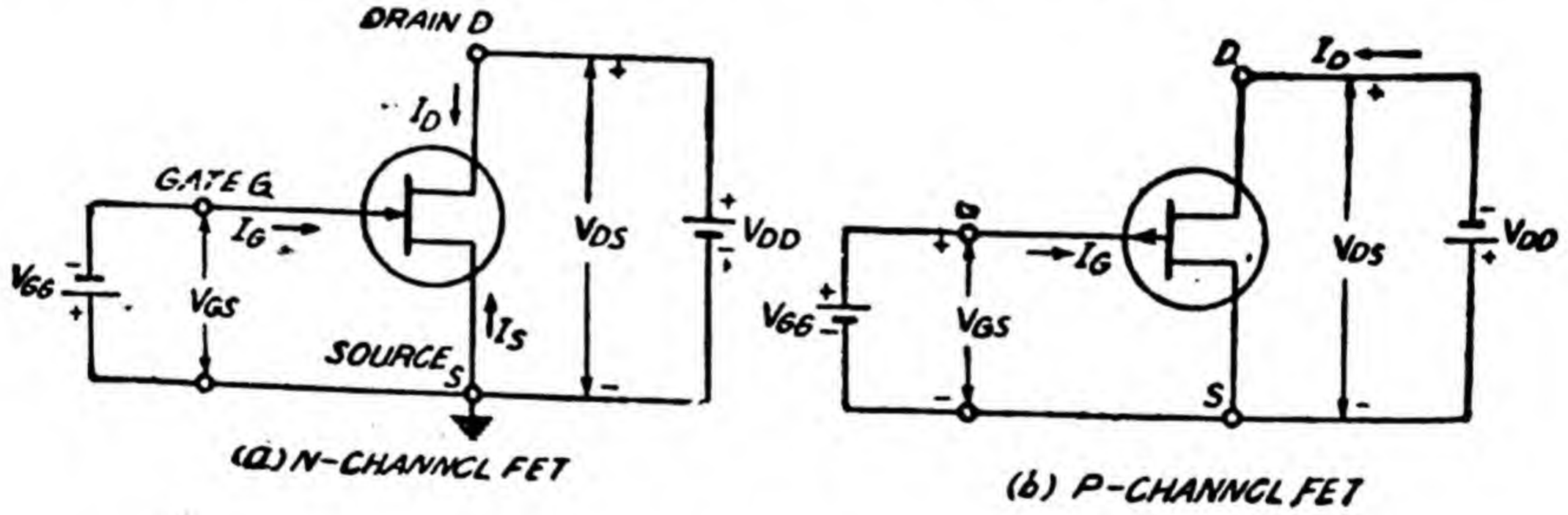
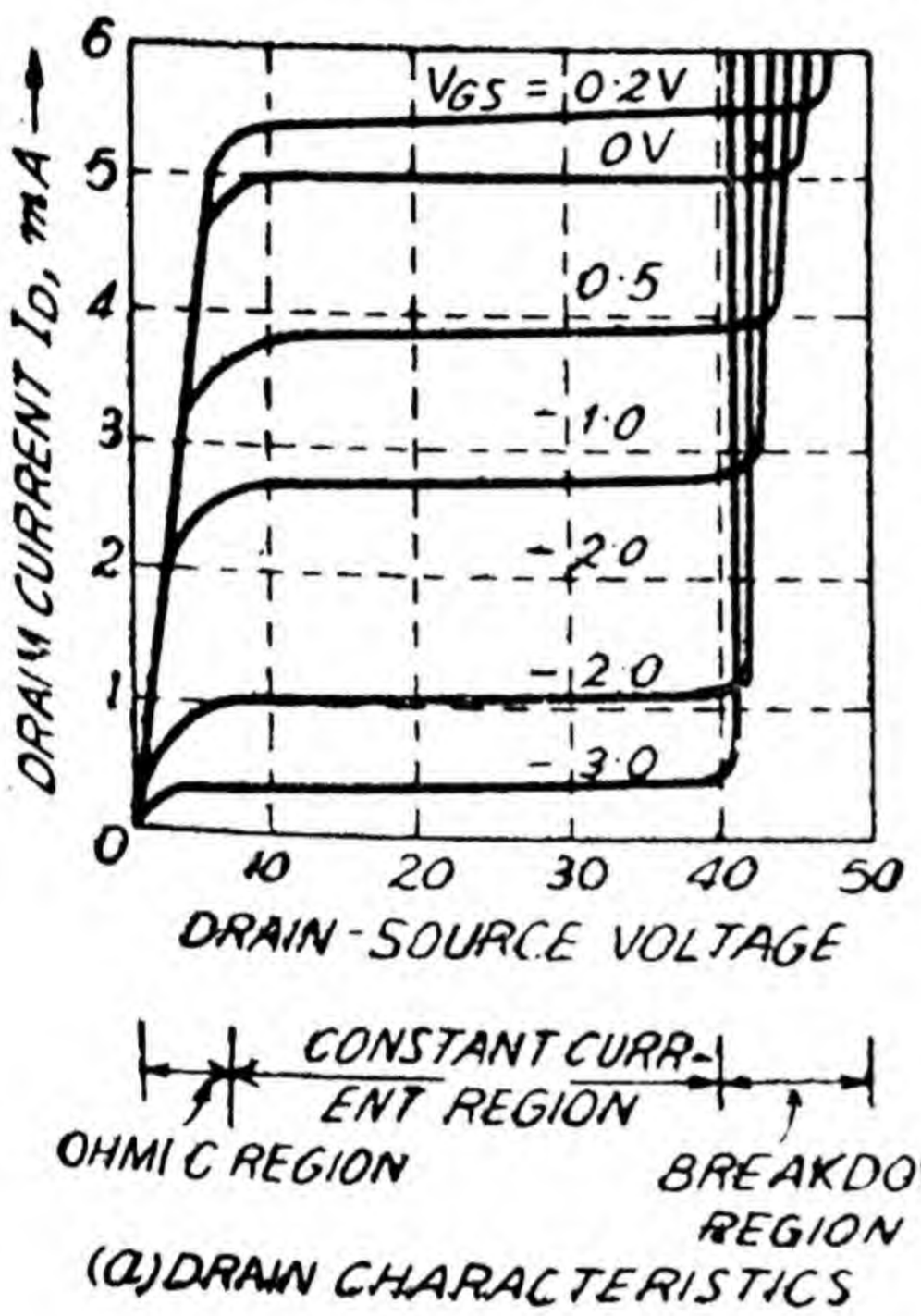


Fig. 36.

Operation : The operation of this transistor can be explained by considering the main current and a controlled voltage. In this transistor the main current is composed of electrons as they are the majority carrier in *N*-type. The *N* type layer provides a path to the main current. The main current is controlled by the signal voltage by applying a reverse bias to both the *PN* junctions. Now we shall explain that how a reverse bias affects the control of the current. Due to the reverse bias voltage across *PN* junction a layer of immobile carriers (called depletion layer) is formed with uncovered positive ions on *N*-side and negative ions on *P*-side. The thickness of depletion layer in *N*-type channel increases as the reverse bias voltage across junction increases *Thus the effective width*

of the channel reduces with the increase of reverse bias. Actually, the depletion region acts like a capacitor. Thus for a fixed drain to source voltage, the drain current is a function of reverse bias voltage. A small signal which provides a reverse bias controls a much large output signal.

Static Characteristics—The static drain characteristics for a typical *N* channel *FET* in its common source configuration is shown in fig. (37*a*). It gives variation of I_D versus V_{DS} with V_{GS} as a



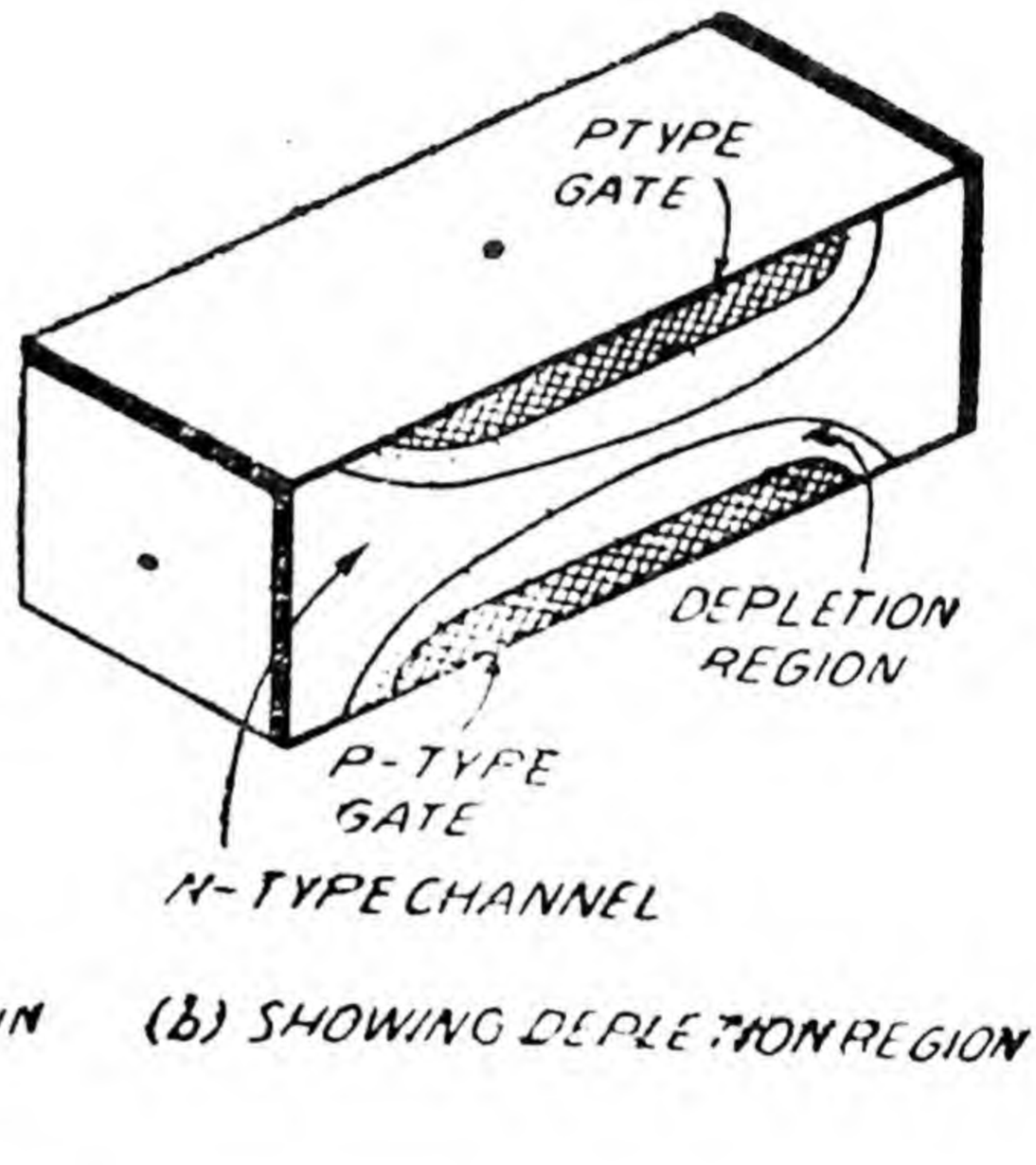


Fig. 37.

parameter. To see qualitatively, why the characteristics have the form shown, let us consider the case for which $V_{GS}=0$. When V_{DS} is zero, the channel between gate junctions is entirely open, but due to absence of any attracting potential at the drain the majority carriers are not drained by the drain terminal and consequently the drain current I_D is zero. In response to a small applied voltage V_{DS}, the *N*-type bar acts as a simple semiconductor resistor and the current I_D increases linearly with V_{DS}. With increasing current, the ohmic voltage drop along the *N* type channel region reverse biases the gate junction and the conducting portion of the channel begins to constrict. This is due to the fact that when a junction is reverse biased a depletion layer is formed and the width of the depletion layer is proportional to the square root of the reverse voltage. The depletion layer extends to the channel as *P* region is heavily doped. The current through the channel is not proportional to the drain source voltage as the resistance changes due to the movement of the depletion layer into the channel and a consequence decrease in the area available for the carriers to move. The depletion layer

is not uniform but is more pronounced at distances farther from the source as indicated in fig. (37*b*). Thus when the drain source voltage is increased from zero value, the channel current increases at first but reaches a maximum value when the pepletion layer extends throughout the channel. As the applied voltage is further increased, the channel is pinched off (all the free charges from the channel are removed) and the current I_D approaches a constant value. The voltage V_{DS} at which the channel is pinched off is called the pinch off voltage V_P. Further increase in voltage V_{DS} eventually causes the avalanche breakdown across the reverse biased gate junction and the current I_D shoots to a high value. Other characteristics are similar to the one for $V_{GS}=0$ except that of a pinch-off voltage.

Cut off Condition : It is seen from the drain characteristics [fig. (37*a*)], that at a fixed value of V_{DS} in the constant current region, the drain current I_D reduces as the reverse-bias at the gate (V_{GS}) is increased. This is due to the fact that as the magnitude of V_{GS} in the reverse-bias direction is progressively increased, the conducting channel width decreases correspondingly. Finally, when V_{GS} is equal to V_P (Pinch off voltage), the channel width is reduced to zero and the drain current I_D becomes zero. However, in practice it is observed that some leakage current still flows even when the magnitude of V_{GS} is made larger than that of V_P. The manufacturer usually specifies a maximum value of leakage current at some value of V_{DS} and V_{GS}. The leakage current is of the order of a few nanoamperes.

Another important parameter is the gate cut off current or gate reverse current. It is gate to source current with the drain shorted to the source *i.e.* $V_{DS}=0$ and with $|V_{GS}| > |V_P|$ under cut off condition. Typically, the gate cut off current is of the order of a few nanoampere for a silicon FET.

Shorted-Gate Drain Current. The current in the shorted gate condition is represented by I_{DSS}. The subscripts stand for Drain to Source with shorted gate. Since the drain current is almost constant in normal range of operation, I_{DSS} represents the approximate drain current for any drain voltage between V_P and V_{DSS} (max.) I_{DSS} represents the maximum drain current we can get for normal operation of a FET.

Transconductance Curve. A curve which relates the output current to input voltage is known as transconductance curve. For FET, it is a graph of I_D versus V_{GS}. By reading the values of I_D and V_{GS} from fig. (37*a*) a transconductance curve can be drawn. The shape of the graph is a parabolic curve *i.e.*. a part of parabola. Shockley demonstrated that the equation of transconductance curve is given by

$$I_D=I_{DSS}\left[1-\frac{V_{GS}}{V_{GS\,(off)}}\right]^2 \qquad \ldots(1)$$

For example, when $I_{DSS}=9mA$, and $V_{GS(off)}=-3V$ then for $V_{GS}=-1{\cdot}5V$, the drain current I_D is given by

$$I_D=0{\cdot}009\left(1+\frac{1{\cdot}5}{-3}\right)^2=\mathbf{2{\cdot}22mA}$$

Parameters relationships and transconductance. The parameters are related as follows :

V_P=Pinch off voltage

I_D=Drain current at some value of $V_{GS}\neq 0$

I_{DSS}=Drain current at $V_{GS}=0$

V_{GS}=Gate to source voltage

V_{DS}=Drain to source voltage

g_{fs}=Transconductance at $I_D\neq I_{DSS}$ and $V_{GS}\neq 0$

g'_{fs}=Transconductance at $I_D=I_{DSS}$ and $V_{GS}=0$

Now $g_{fs}=\dfrac{\triangle I_D}{\triangle V_{GS}}=\dfrac{2I_{DSS}}{V_P}\left(1-\dfrac{V_{GS}}{V_P}\right)$ using eq. (1) ...(2)

and $g'_{fs}=2I_{DSS}/V_P \quad (\because V_{GS}=0)$ ...(3)

FET Parameters. In analogy with the vacuum triode or bipolar transistor, the FET has the following important parameters :

Amplification factor μ. This is defined as

$$\mu=-\left(\frac{\partial V_{DS}}{\partial V_{GS}}\right)_{I_D \text{ constant}} \quad ...(1)$$

Dynamic drain resistance r_D. This is defined as

$$r_D=\left(\frac{\partial V_{DS}}{\partial I_D}\right)_{V_{GS} \text{ constant}} \quad ...(2)$$

The reciprocal of r_d is drain conductance g_D. This is also called as common source output conductance.

Mutual Conductance g_m. The mutual conductance or the transconductance is defined as

$$g_m=\left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS} \text{ constant}} \quad ...(3)$$

This is also called as common source forward transadmittance. The above parameters are related as

$$\mu=r_D g_m \quad ...(4)$$

Comparison of N and P Channel JFETs

N–Channel JFET	P-Channel JFET
1. The current carriers are electrons.	Current carriers are holes.
2. The mobility of electrons is almost twice that of holes in P-channel.	The mobility of holes is poor.
3. Low input noise.	Large input noise.
4. Large transconductance.	Low transconductance.

Thus in electronic circuits, *N*-channel device is superior to P-Channel device.

6·26-2. METAL OXIDE SEMICONDUCTOR FET :

There are two types of metal oxide semiconductor FET. The first type is known as enhancement type MOSFET, while the second type as depletion MOSFET. Here we shall describe both types of MOSFET.

(*i*) **Enhancement type MOSFET :** An *N* channel MOSFET is shown in figure (38 *a*). There are two highly doped (*N*+) regions (one source and the other drain) on a lightly doped *P* type (substrate). A layer of insulating silicon di-oxide is placed over it. Now over the insulating SiO_2 layer a gate is formed. Three different terminals for source, gate and drain are taken out as shown in figure. The insulating layer between gate and channels provides a very high input impedance. When a positive voltage is applied at the gate, it induces negative charges in insulating layer and correspondingly negative charges in semiconductor. The negative charge region in semiconductor increases as the positive voltage at the gate increases This negative charge in semiconductor helps in conducting between source and drain. In this way the drain current is enhanced by the positive gate voltage.

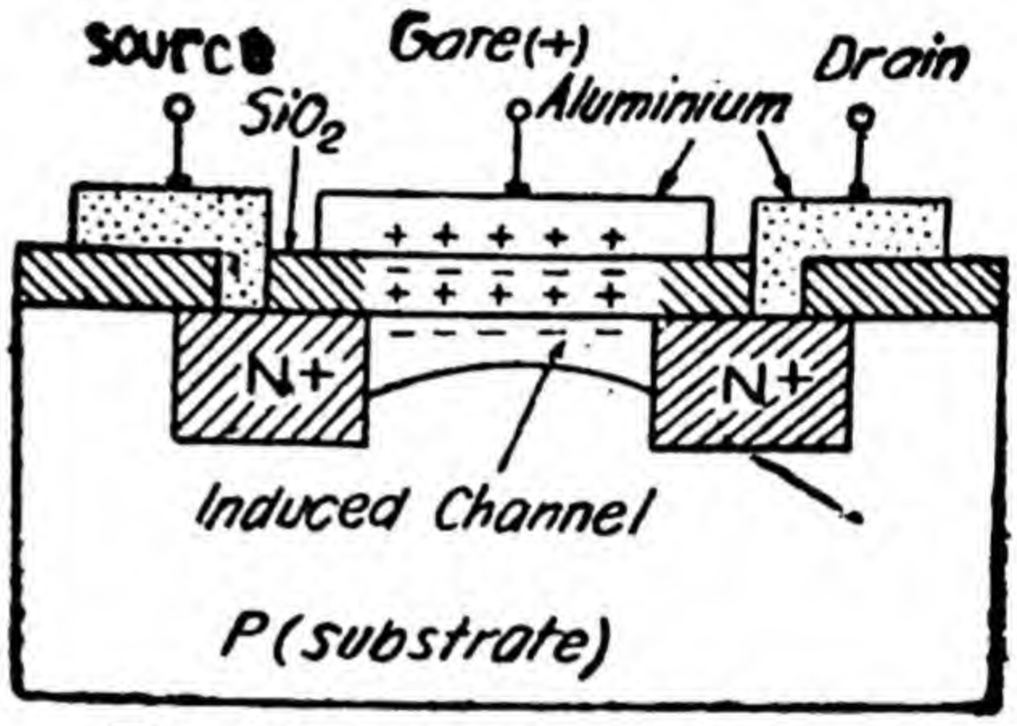


Fig. 38*a*. An *N* channel MOSFET

Fig. 38*b* and 38*c* show the drain characteristics and transfer characteristics of *N*-channel enhancement type MOSFET.

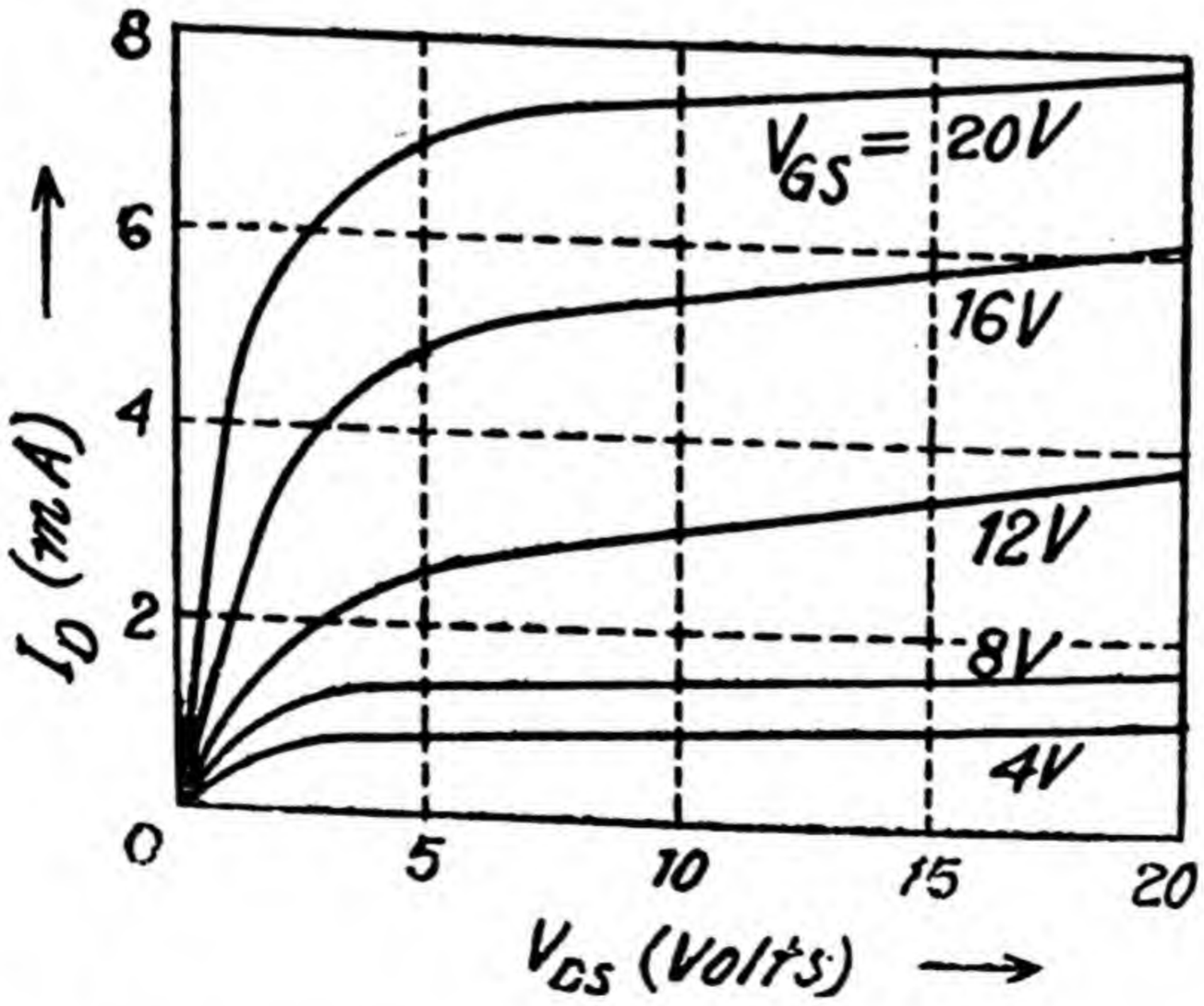


Fig. 38*b*. The drain static characteristics of *N*-channel enhancement type MOSFET

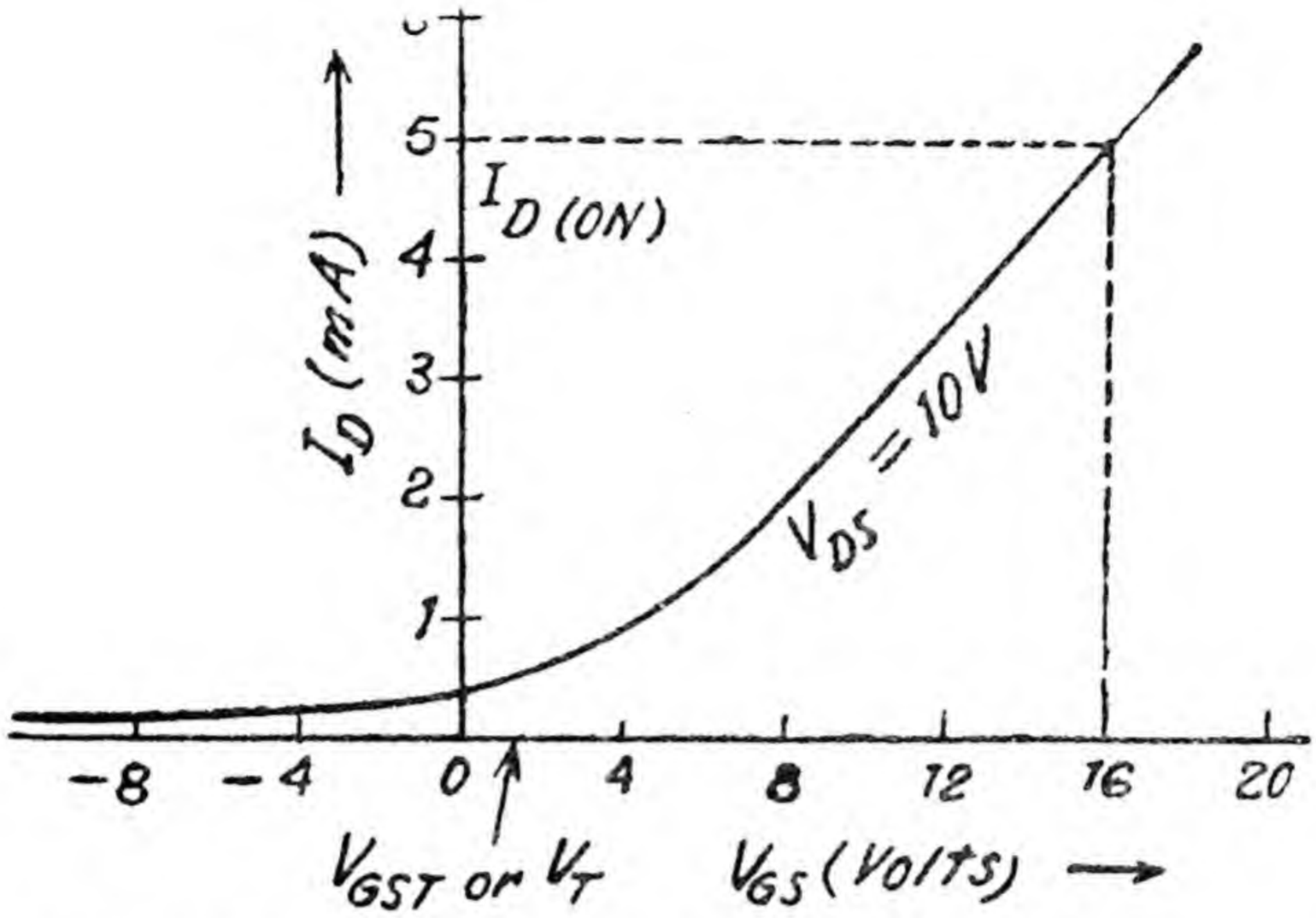


Fig. 38*c*. Transfer characteristics of *N* channel enhancement type MOSFET.

It may be noted from fig. 38*b*, that the drain current for V_{GS} greater than zero is very small and as V_{GS} is made more positive the current increases slowly at first and then at a much more rapid rate. The gate source voltage at which the channel is formed to let through the flow of drain current of predefined value (say $10\mu A$) is called the gate-source threshold value. The gate-source threshold value is denoted by V_{GST} or simply V_T as shown in fig 38*c*. The current $I_{D(ON)}$ represents the maximum permissible current on the drain characteristics of fig. 38*c*. These values (V_{GST} and $I_{D(ON)}$) are sometimes specified by the manufacturer for particular MOSFET.

(*ii*) **Depletion MOSFET :** An *N*-channel depletion MOSFET with its symbol is shown in figure (39*a*). The difference between depletion MOSFET and enhancement MOSFET is that in depletion MOSFET **an *N*-type layer** is also introduced between source and

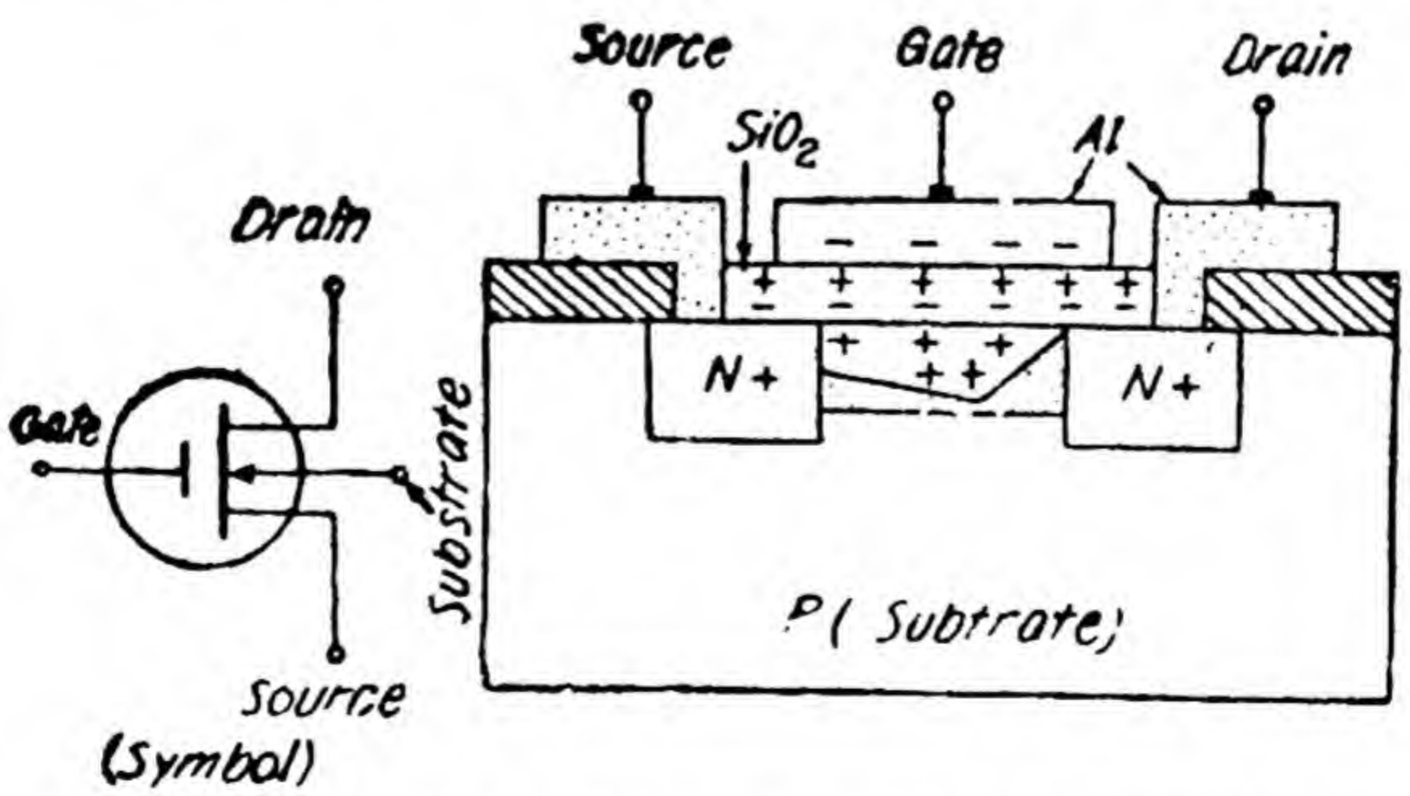


Fig. 39*a*. Depletion type MOSFET.

drain. When a negative voltage is applied at the gate, positive charges are induced in the channel. We know that the electrons

are the majority carriers in *N*-type material but the positive charges reduce the conductivity of this channel by forming a depletion layer. In this way the drain current can be reduced by applying a negative voltage at the gate. However, depletion MOSFET can be used in enhancement mode by applying positive grid voltage.

Fig. (39*b*) and (39*c*) show the drain characteristic and transfer charcteristic for an-*N*-channel MOS FET which can be operated in both depletion mode and enhancement mode corresponding to negative and positive gate voltages respectively. So more appropriately, MOSFET may be designated as dual mode MOSFET.

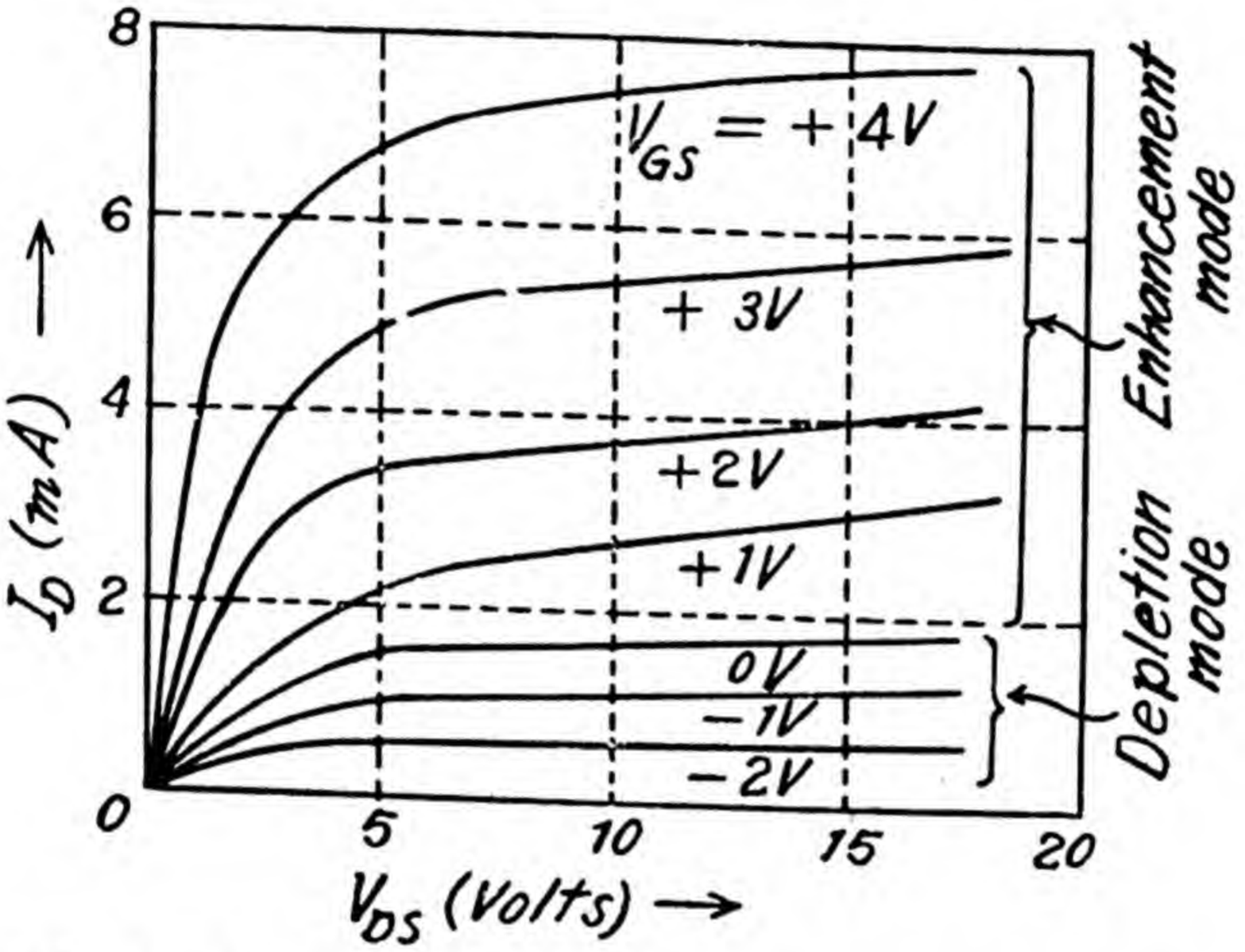


Fig. 39*b*. Drain characteristics of *N*-channel dual mode MOSFET

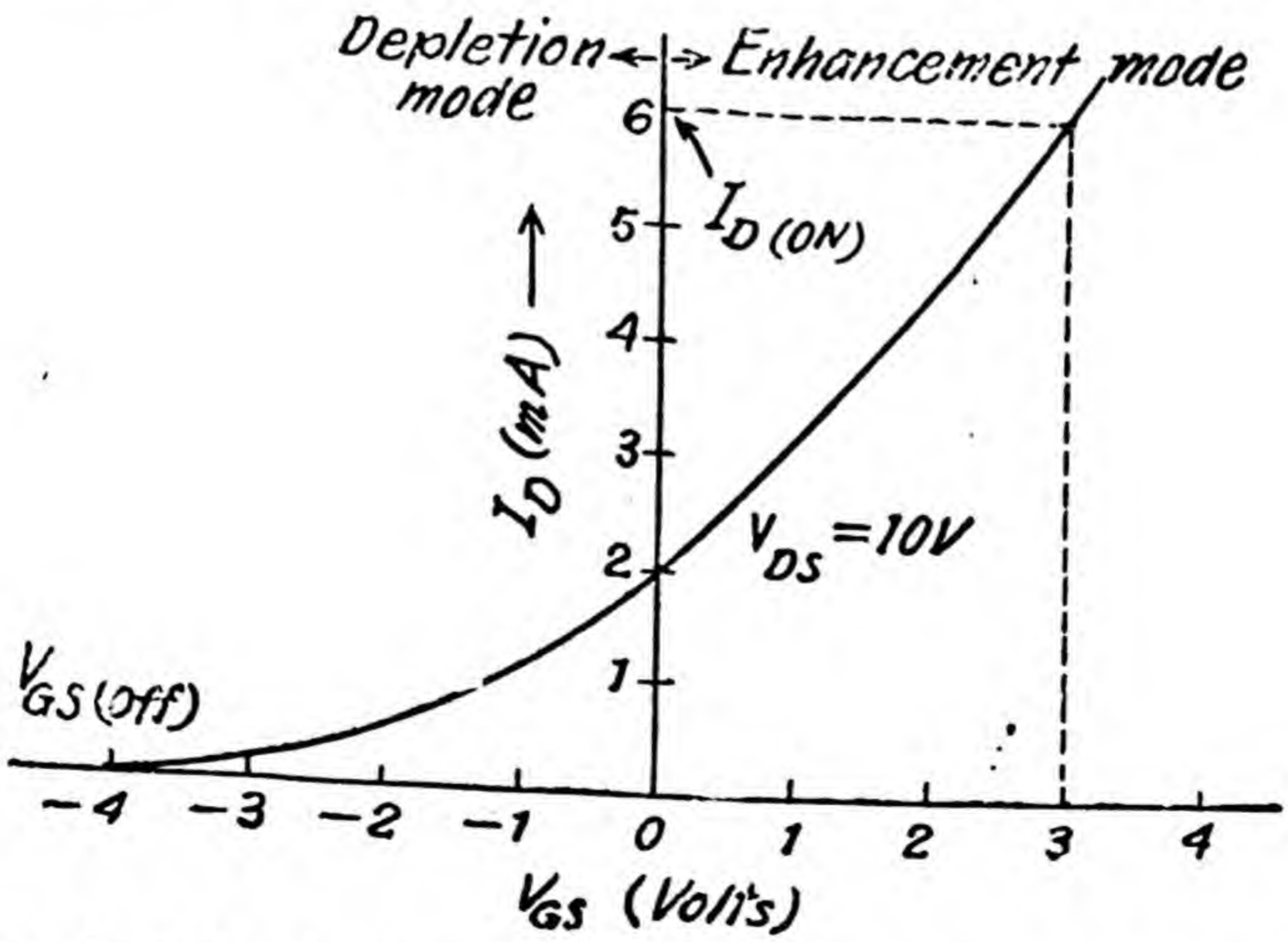


Fig. 39*c*. Transfer characteristics of *N*-channel dual mode MOSFET

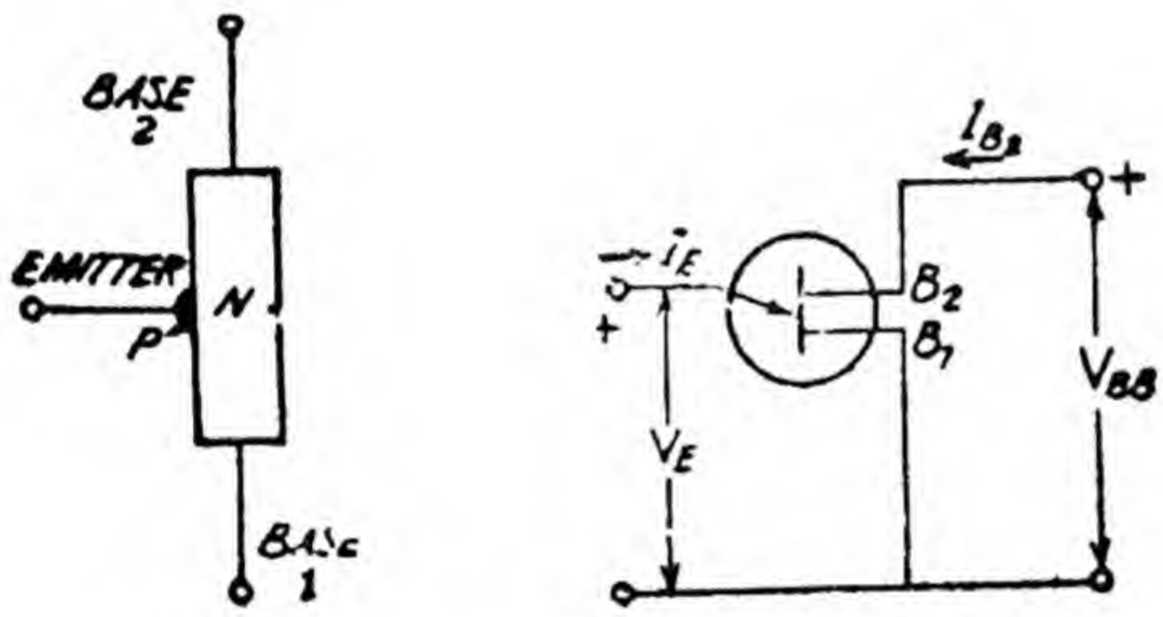
It is obvious from fig. (39*c*), that for V_{GS}=0, the drain current is not zero but has appreciable value. The figure shows the gate

source cut off voltage $V_{GS(off)}$ at which the drain current I_D reduces to some specified negligible value at a recommended drain voltage V_{DS}. This voltage $V_{GS(off)}$ corresponds to a pinch off voltage V_P of a JFET.

The operation and the characteristics of a *P*-channel MOSFET are similar to *N*-channel MOSFET except that the signs of currents and voltages are reversed.

6 27. THE UNIJUNCTION TRANSISTOR (double base diodes):

The device that leads itself to calculation on a negative impedance basis is the unijunction transistor. The device resembles the FET in the sense that it has only *P-N* junction and differs from FET because the junction is forward biased. The transistor and biasing are shown in fig. 40 (*a*) and 40 (*b*). It is a double base diode with a lightly doped bar of *N*-type silicon and has a junction located at the centre. The resistance between base 1 and base 2 is of the order of 5 to 10 Ω when the junction is reversed biased.



(*a*) (*b*)

Fig. 40. Unijunction transistor construction and circuit diagram symbol.

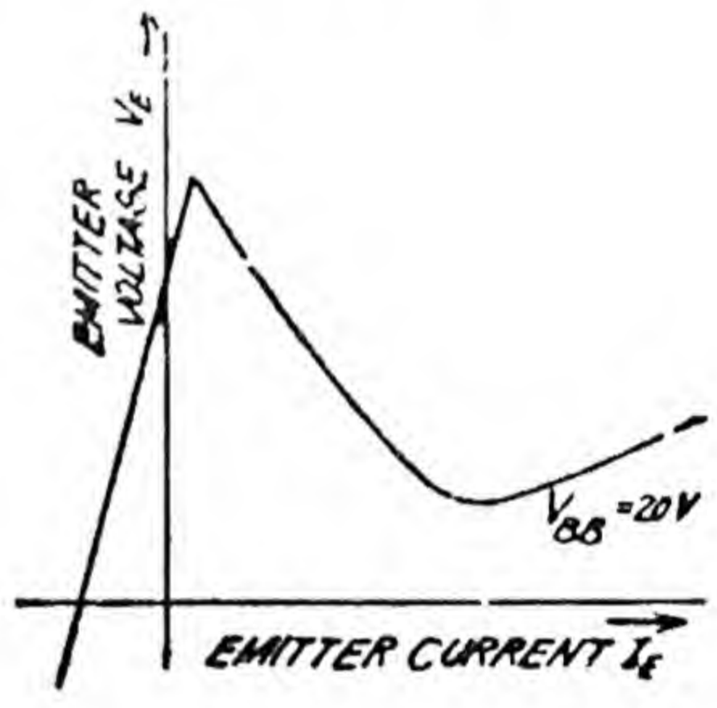


Fig. 41. Characteristic of unijunction transistor.

Figure 41 shows the static emitter characteristic of the transistor. For values of V_E below (0·5–0·75) V_{BB}, the emitter junction is reversed biased, and only a low level of emitter current is possible. When the emitter diode is forward biased, more holes are injected into the bar which are repelled by base 2, but their presence increases the conductivity between emitter and base 1. This increased conductivity results in a repuced voltage drop V_E required to support a given current level I_E. Thus when the voltage from emitter to base 1 is falling, a negative resistance results. The end of this negative resistance region occurs when saturation is reached. For currents above saturation, the voltage again increases but only slowly.

6 28. GUNN EFFECT AND DIODES

With the help of junction diodes, high microwave power can not be obtained. While studying the properties of thin specimens of gallium arsenide, J. B. Gunn, in 1963 discovered that under high electric stress, there is periodic fluctuations in the current passed by the material. This effect is known as Gunn effect. This

is also known as *bulk transferred electron effect*. Besides gallium arsenide, there are few other semiconductors like indium phosphide, cadmium telluride and indium arsenide which exhibit this property.

Fig. (42) shows the construction of a typical Gunn diode. When a small d.c. voltage is applied across the thin slice of gallium arsenide, negative resistance manifests itself under certain conditions. If the slice is connected to a suitably tuned circuit, oscillations occur. As the slice is very thin, the voltage gradient across the slice of GaAs is very high. Due to this reason, the electrons velocity is also very high which causes the oscillations to occur at microwave frequencies.

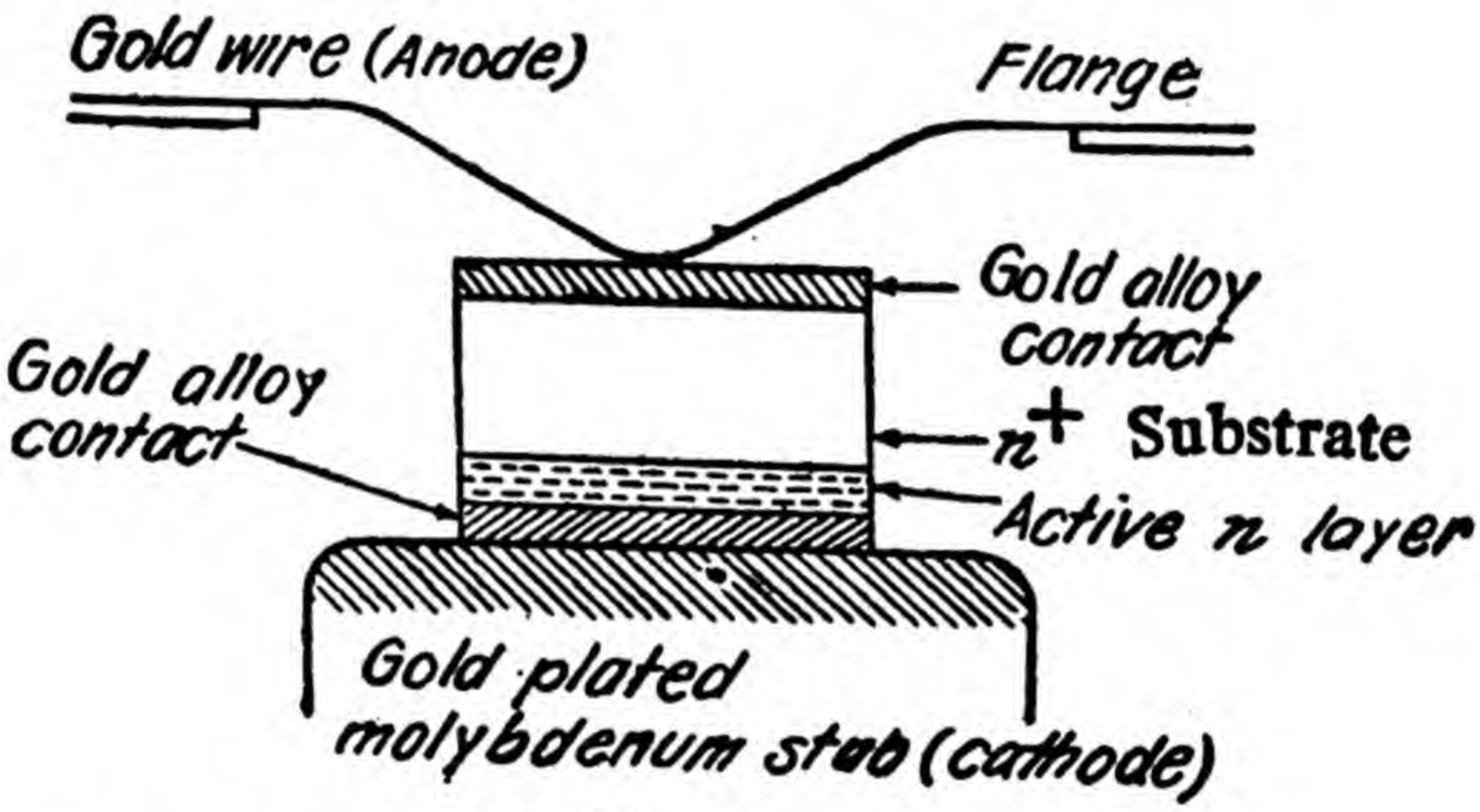


Fig. 42.

Gunn effect is a bulk property of semiconductors and does not depend on either junction or contact properties. This effect is independent of total voltage or current. This is not affected by magnetic fields or different types of contacts. As this effect is found in *N* type materials, so it is associated with electrons and not with holes. Gunn observed that the voltage required is proportional to the sample length, in volt per centimeter. This is the major factor determining the presence or absence of oscillations. For oscillations to take place, the threshold value of 3·3 KV/cm should be exceeded. The frequency of oscillations produced corresponds to the time taken by the electrons to traverse the width of *N* type material, as a result of voltage applied.

Negative resistance. Fig. 43 shows the important energy levels in gallium arsenide. This is semiconductor which, in an *N* doped sample, has an empty energy band higher in energy than the highest filled or partly filled band. Moreover the size of the forbidden gap between these two is relatively small. This is a peculiar situation which is not found in common semiconductors like germanium and silicon.

On the application of voltage across a slice of GaAs which has an excess of electrons (*N* type), the electrons flow (current) towards the positive end of the slice. The current is a function of the voltage applied *i.e.* higher is the voltage applied, higher would

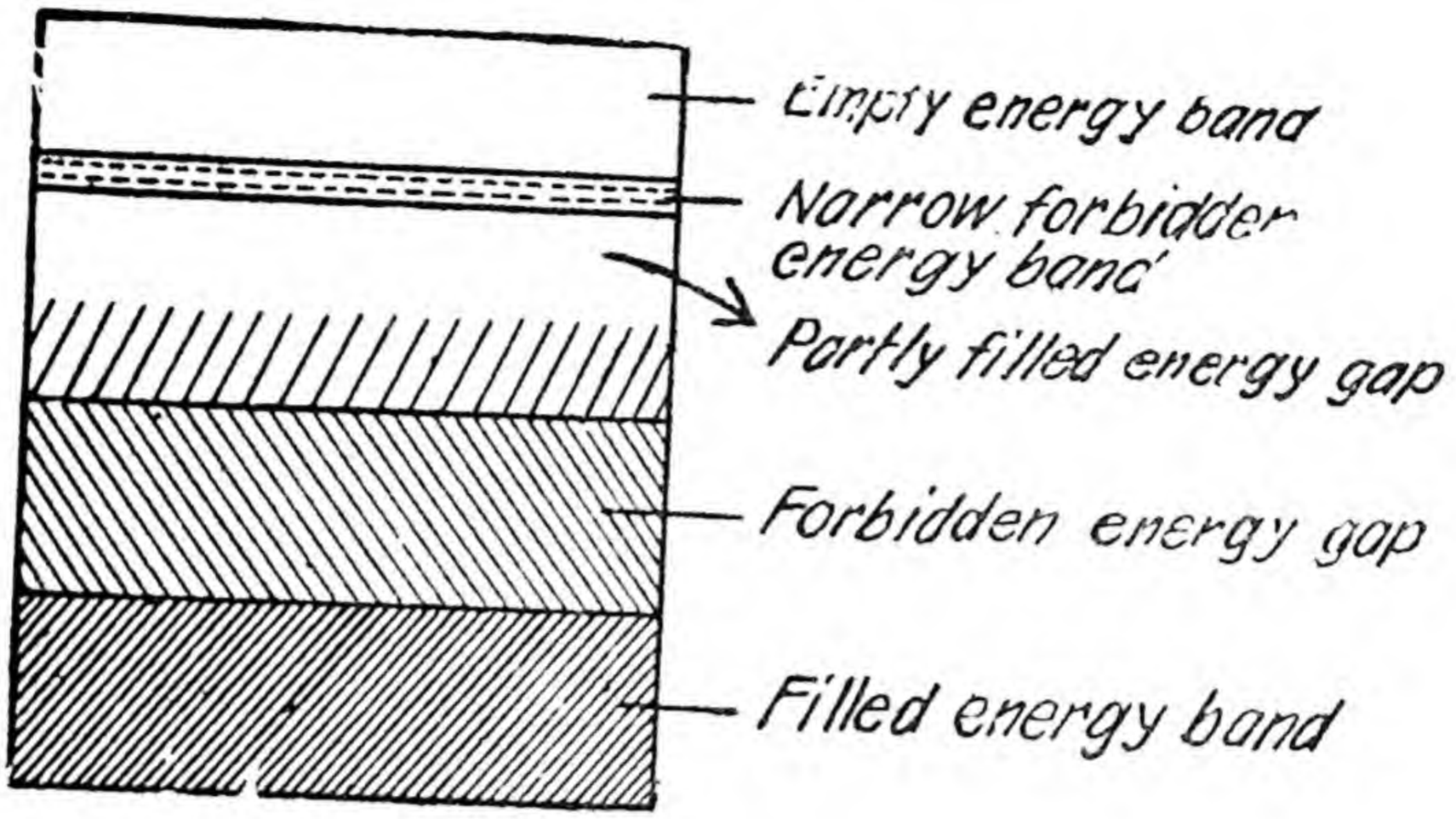


Fig. 43

be the current. So far, the device behaves as a normal positive resistance. In ordinary diodes, the component of velocity towards the positive end, imparted to the electrons due to applied voltage, is quite small compared to the random thermal velocity which these electron possess. However, in case of GaAs Gunn diode, the high voltage gradient impart so much energy to electrons that instead of travelling faster (constituting a large current) they are actually sloewd down. The reason is that due to very high velocity, the electrons get transferred into empty energy band at the top of fig. 43, where they becomes less mobile *i.e.* slowed down. This gives rise to the name transferred electron effect. *The electrons are thus transferred from the conduction band to a higher energy band in which they are much les mobile and so the current is reduced as the voltage is increased.* It is thus seen that as the voltage rises past the threshold level, the current falls and the negative resistance is exhibited. The main difference between gallium arsenide and other semiconductor materials is that it requires a low voltage to transfer electrons to empty energy band compared to other semiconductor materials.

Volt ampere characteristic : The basic meehanism involved in the operation of bulk *N* type GaAs device is to transfer the electrons from lower conduction band to upper empty band. The effective mass of the electron in the lower conduction band is smaller than in the upper empty band. This shows that there are different mobilities in the two bands. In *N* type GaAs. the most of the electrons at low electric fields and at low lattice temperature occupy the lower conduction band with ohmic current

$$J=\sigma E \qquad \ldots(1)$$

where $\sigma \cong e n_1 \mu_1 \cong e n_0 \mu_1$. Here n_1 is the carrier concentration and is assumed to be approximately equal to the total carrier concentration n, and μ_1 is the mobility in the lower conduction band.

As the applied electric field is increased, the electrons gain energy and move upward in empty band. In this band their mobility is decreased as the effective mass increases, thus decreasing the current density J. There is a certain threshold field above which this transfer starts to take place. Now the current density is given by

$$J=\sigma E=e\ (n_1+n_2)\bar{\mu}E$$
$$=en_0\bar{\mu}E \qquad \ldots(2)$$

where $$\bar{\mu}=\frac{n_1\mu_1+n_2\mu_2}{n_0},$$

the average mobility of electrons.

When the field is raised high enough, almost all the electrons transfer to empty energy band. The current density now becomes

$$J=\sigma E\cong en_2\mu_2E \qquad \ldots(3)$$

where n_2=carrier concentration in upper empty band, and

μ_2=mobility in this band.

A graph between J and E is shown in fig. 44. In fig. 44 J_m is the maximum current density, J_V is the empty band current density, E_m is the maximum electric field required before the onset of negative conduction region E_a is the maximum electric field for which eq. (1) is valid, E_b is electric field beyond which eq. (3) holds and E_V is the electric field corresponding to J_V. Since J is proportional to the terminal current I and E is proportional to the applied voltage V, fig. (44) is the V-I characteristic of Gunn diode The region of the characteristic between E_m and E_V where current density decreases with increasing electric field is one of *negative* differential resistivity.

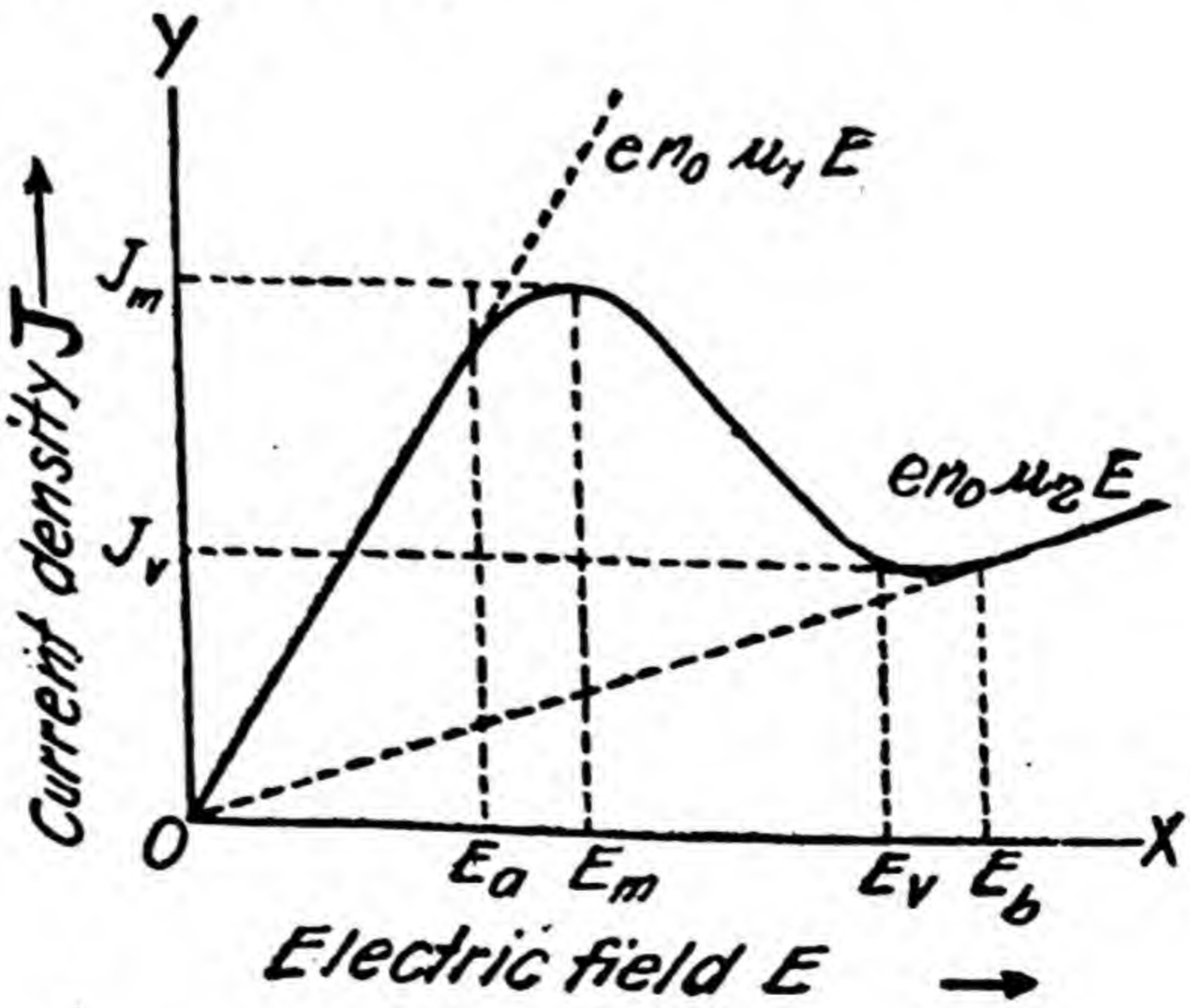


Fig. 44

Gunn Domain : When a d.c. bias of a value equal to or more

than that corresponding to threshold field is applied to a Gunn doide the charge densities and the electric field within the sample become nonuniform Due to this reason domains (bunch of electrons) are created *i.e.* electrons in some region of the sample experience transfer first than the rest of the sample. The reason is as follows :

It is reasonable to except that the density of doping material is not completely uniform throughout the sample of gallium arsenide. So it is possible that there be a region, perhaps somewhere near the negative end, where such an impurity concentration is less than average. In such a region there will be fewer free electrons than in other areas and therefore this region is less conductive than the others. As a result, the field in this area will be greater than anywhere else in the sample *i.e.* the potential at this region will be greater than average potential. When the applied voltage is increased, this region will be the first to have a voltage across it large enough to induce transfer of electrons to the higher energy band. Infact, such a region will then become a *negative resistance domain*. There may be another possibility of the non uniformity in charge density and electric field. This is due to thermal fluctuations in the carrier density. The fluctuations in the carrier density creates a localized excess of electrons. The electric field near the excess of electrons is composed of both the applied field E_m and the field of charges. This is shown in figure (45). It is obvious from the figure that the net electric field on the anode side of the domain is greater than the cathode side. When the field at the anode side exceeds threshold, the domain move with lesser velocity. This creates a slight excess of electrons in that location and a slight deficiency of electrons immediately ahead. The region of excess and deficient electrons form a *dipole layer*. The electric field inside the dipole layer is higher than any where else. The electric field outside the dipole starts dropping as a result of constant applied field. As the dipole drifts along, more electrons in the vicinity will transfer to upper band. This continues until the electric field outside the dipole region is depressed below the threshold value of field and the dipole is collected at the anode. Upon collection, the field in the sample jumps to the original value. Now the next formation of the domain begins as soon as the field value exceeds the threshold value.

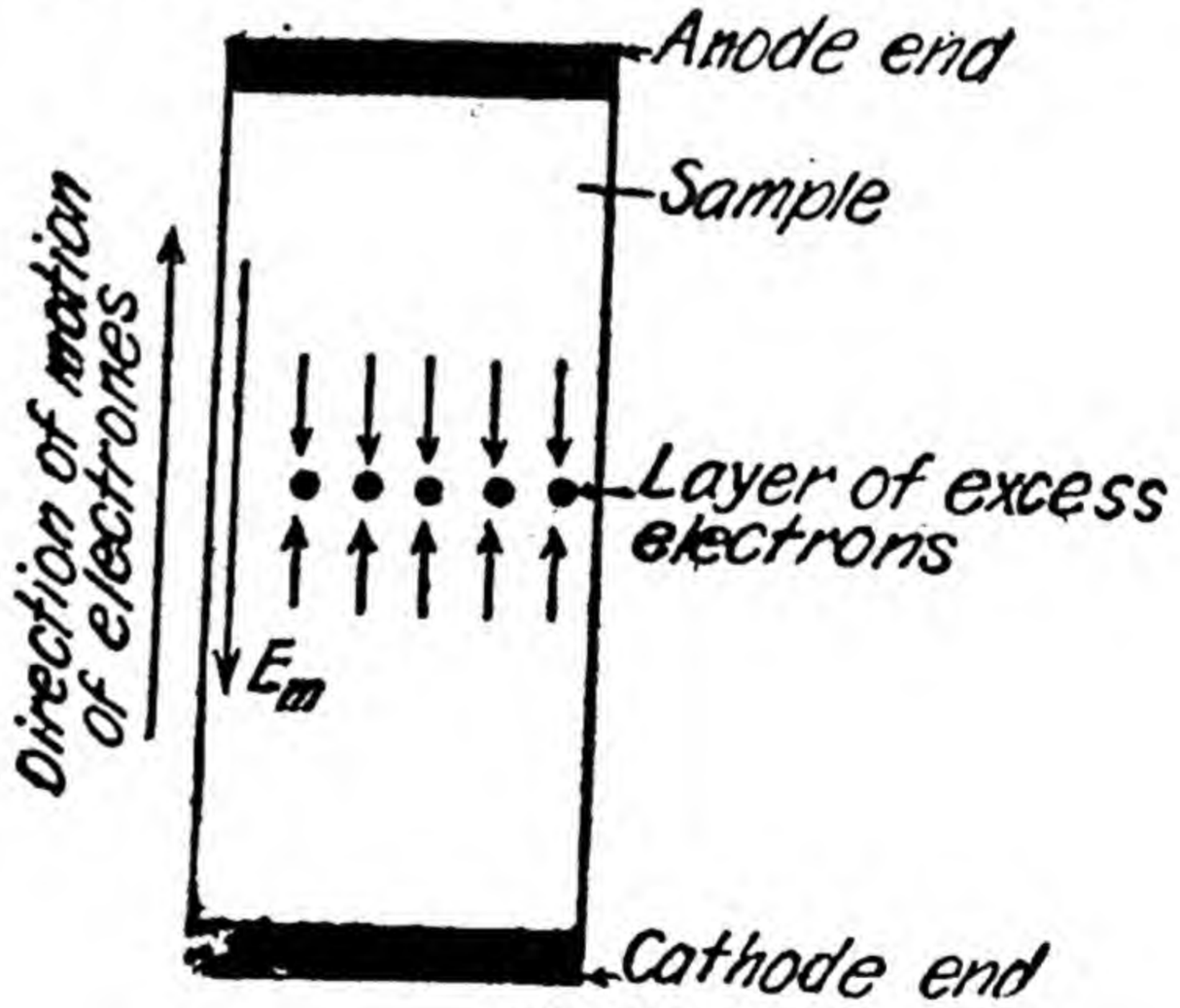


Fig. 45

The domain is quite capable of travelling and may be thought of as a low conductivity, high transfer region, corresponding to a negative pulse of voltage. When it arrives at the positive end of a slice, a pulse is received by the associated tank circuit and shocks into oscillations. Infact, these pulses are responsible for the oscillations in Gunn diode, rather than the negative resistance. It is seen that only one domain, or pulse, is formed per cycle of *RF* oscillations. so that energy is received by the tank circuit in correct phase to permit oscillations to continue.

LSA mode and diodes. The term diode is a misnomer for Gunn devices because neither there is a junction nor the rectification. This is called diode simply because it has two terminals. The Gunn devices have a number of names as transferred electron device (TED), transferred electron oscillator (TEO), bulk effect diode, bulk effect oscillator (BEO) and bulk GaAs oscillator. LSA is an acronym used for Limited Space-charge Accumulation mode. In this mode the domain is not allowed to form at all. The frequency and r.f. voltage are so chosen that the domain is not given sufficient time to form before it is extinguished.

The basic construction of *LSA* diode is shown in fig. (46). The construction is similar to a Gunn diode with difference that the size of active region is much larger in this case. The larger size of the active region permits higher power dissipation which causes higher power output and higher operating frequencies. As the domains are suppressed, negative resistance operation takes place. This is relatively independent of transit time *i.e.* the device length. The voltage across this larger diode can be increased considerably for the same current. Thus the power output is increased significantly.

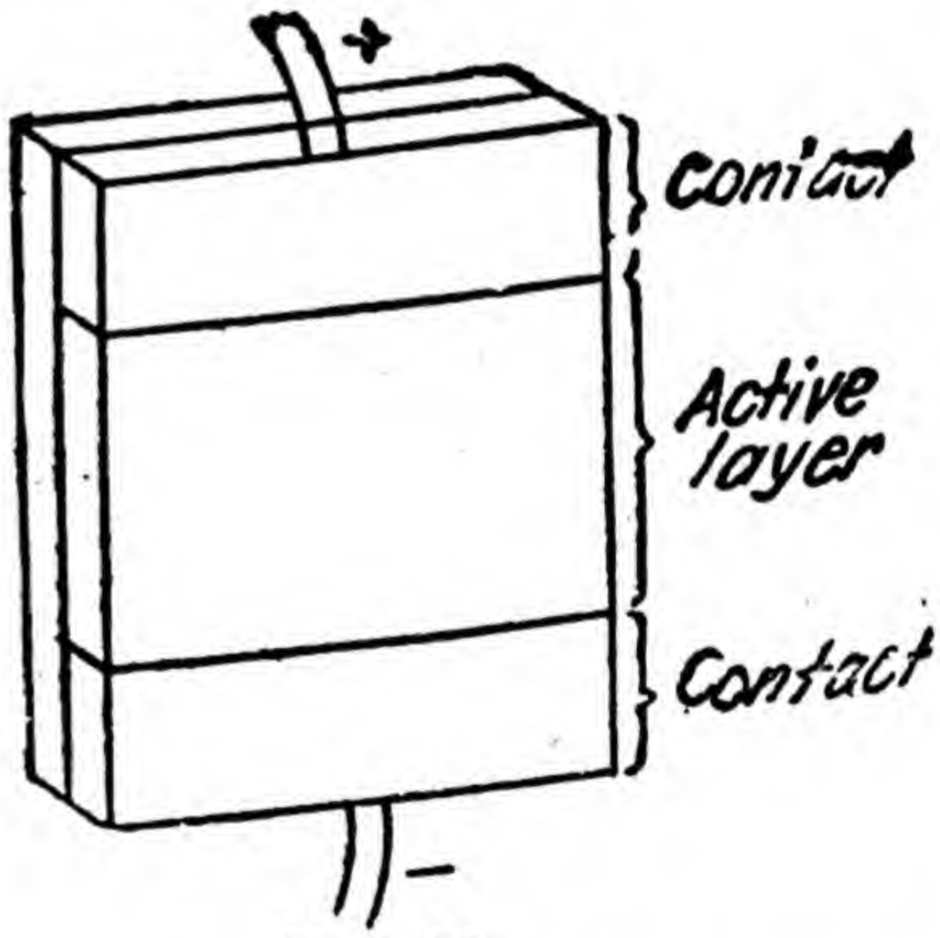


Fig. 46

Following are the steps used to prevent the formation of domains :

(*i*) As shown in fig. (47), the applied voltage is well in excess of the threshold value. This bias together with the external resistance is so arranged that the oscillating voltage exceeds past the negative resistance region. Thus any domain formation will be dissipated.

(*ii*) The resonant frequency of the cavity is made three to twenty times the required value for domain operation. Now the *rf* cycle is too fast to permit the domains formation anywhere.

(*iii*) The domain formation can be also prevented by carefully

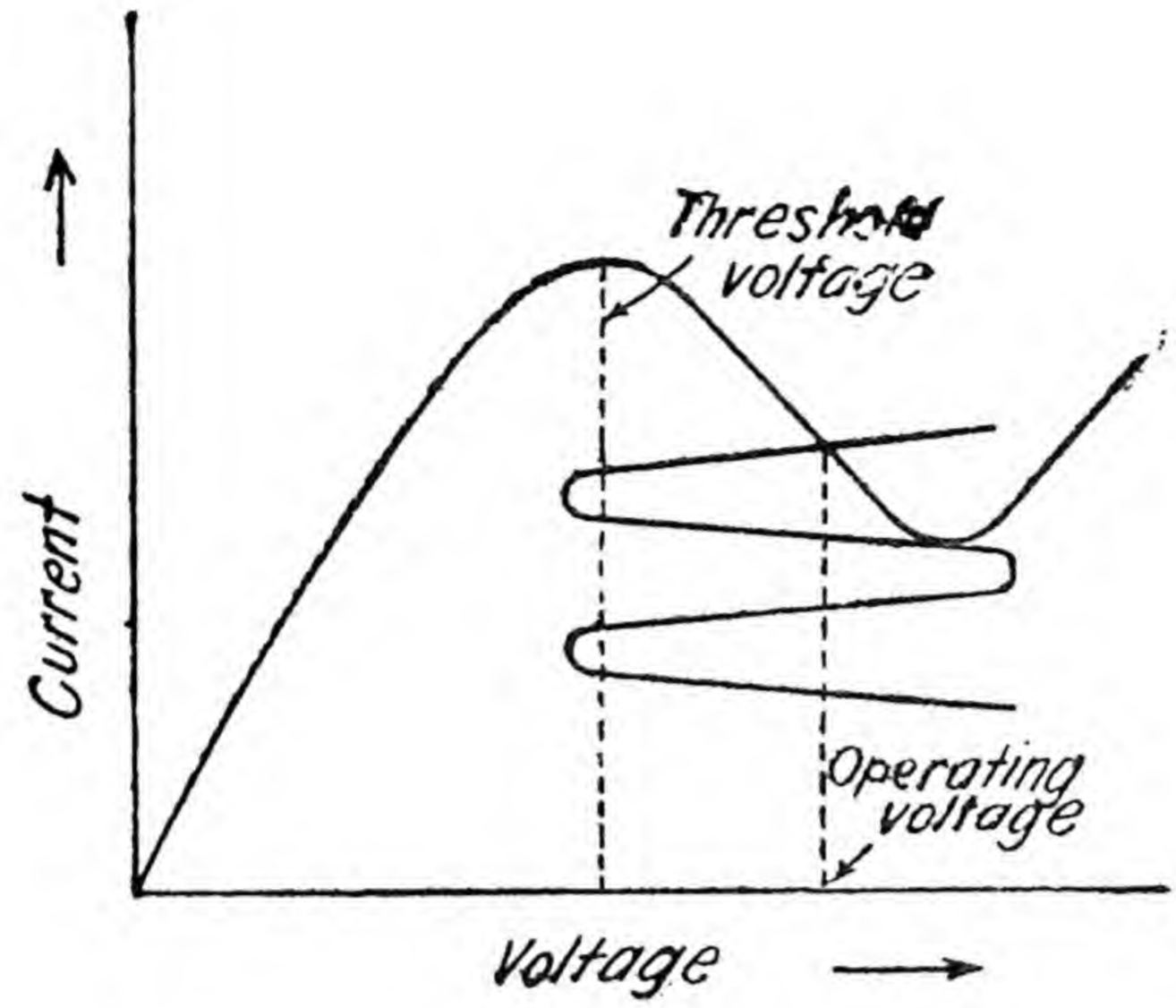


Fig. 47

chosing the ratio of doping level to the operating frequency. *LSA* mode has the following advantages :

(*a*) High efficiency is obtained since most of the devices exhibit negative differential conductance.

(*b*) Since the domains are not formed, high operating voltage is permitted without causing avalanche break down. This suggests high power operation.

(*c*) This mode may be utilised for pulsed oscillations at the highest frequencies.

6·29. IMPATT (IMPACT AVALANCHE AND TRANSIT TIME) DIODES

The IMPATT diode is a two terminal *PN* Junction device operating at reverse bias. The diode uses for its operation *avalanche current multiplication* to achieve a *negative resistance.* The negative resistance may be used at microwave frequencies. The negative resistance of a device means the rise of current with a fall of voltage and vice versa. The negative resistance of a device may also be defined as the property which causes the current through it to be 180° out of phase with the voltage across it. Now we shall explain the term avalanche current multiplication. When a sufficiently high electric field exists in the depletion layer, the velocities of the carriers crossing the depletion layer increases. These carriers (electrons and holes) undergo collisions within the crystal. Out of these collisions some are so violent that electrons are knocked off the crystal atoms. This creates electron-hole pairs. The pairs attain high velocities to cause further pair generation through more collisions. Near the breakdown voltage, the field becomes so large that the chain of collisions can give rise to an almost infinite current with very slight additional increase in voltage. The process

is known as *impact ionization* and production of very large current implies *avalanche multiplication*. A combination of delay involved in generating avalanche current multiplication, together with delay due to transit-time through a drift space, provides the required 180° phase difference between applied voltage and the resulting current in an IMPATT diode. Fig (48) shows the active cross-sectional area of IMPATT diode. Now we shall show that how IMPATT diode shows a 180° phase difference *i.e.* a negative resistance.

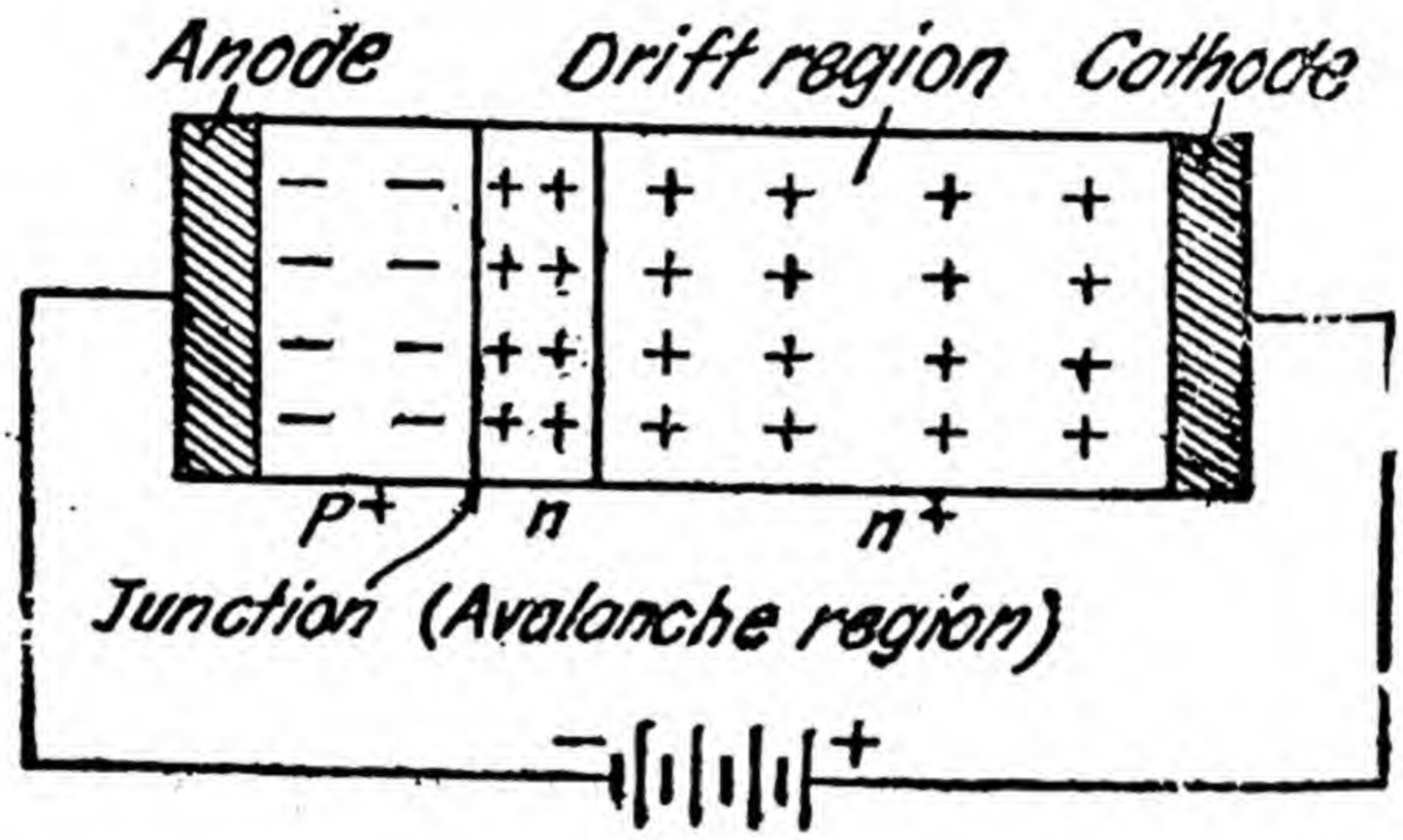


Fig. 48

The back biasing of the diode, results in a flow of minority carriers across the junction. The junction is between p^+ and the n layers. Let us consider the effect of a positive swing of the *r. f.* voltage superimposed on the d.c biasing voltage, assuming that oscillations already exist. Here the process of *impact ionization* continues until avalanche takes place. If the original *dc* biasing voltage is set just equal to the threshold required for this avalanche, then this threshold voltage will be exceeded during the whole of the positive *r. f.* cycle and the avalanche current multiplication will be taking place during this entire time. Here the avalanche is not instantaneous. Fig. (49) shows the voltage and current curve.

It is obvious from the figure that the current pulse reaches its maximum (at the junction) at the instant when the r. *f.* voltage across the diode is zero and going negative. This indicates a 90° phase difference between the voltage and the current. Due to reverse bias, the current pulse drifts towards the cathode. The time taken by the current pulse to reach the cathode depends upon the thickness of n^+ layer and its drift velocity. If the external conditions are suitably arranged, it is possible to ensure a phase difference of 90°. In this way a total phase difference of 180° exists between the applied voltage and the resulting current, so that the device may be said to exhibit negative resistance under these conditions. Such a negative resistance may be used in oscilla-

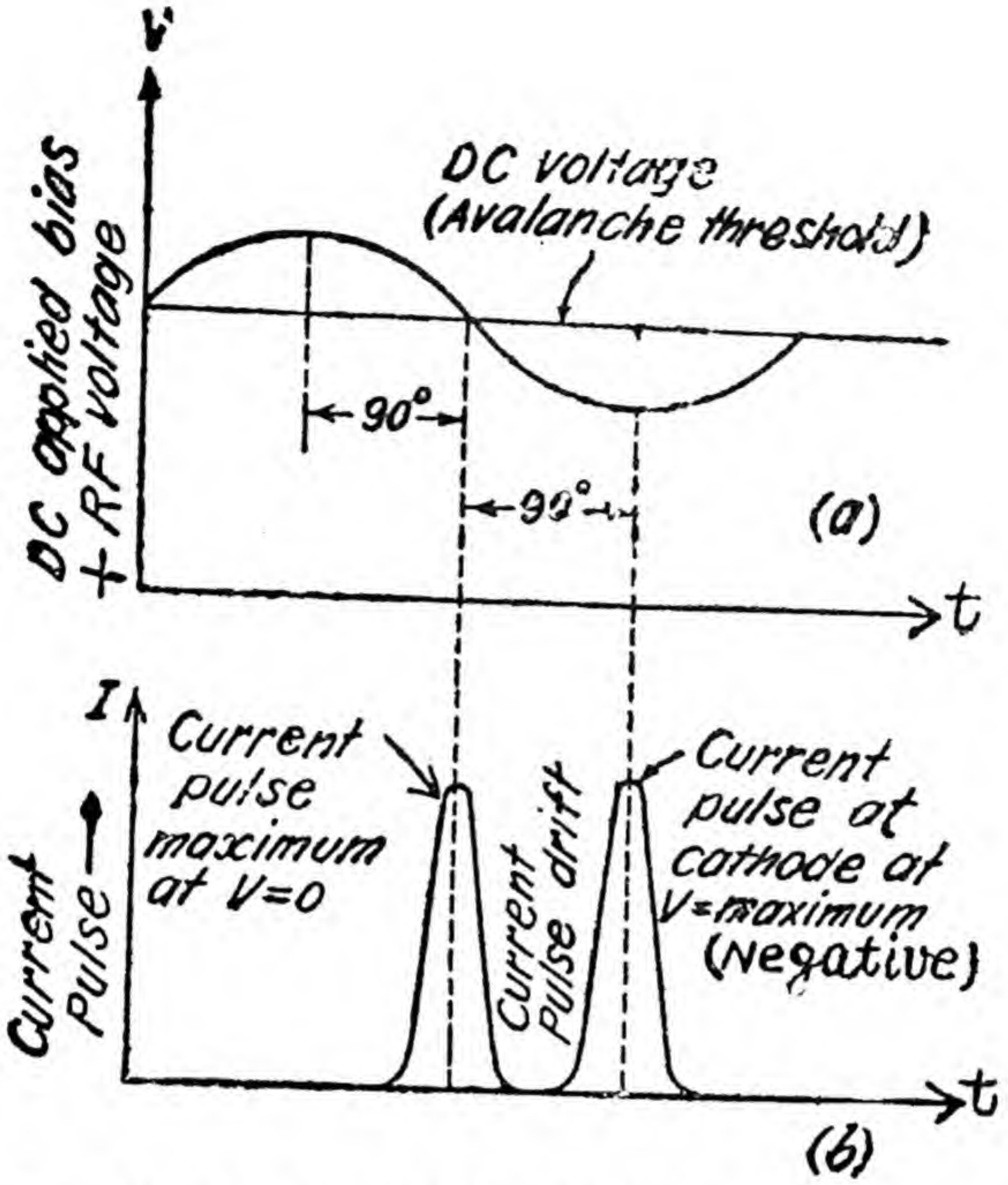


Fig. 49. (a) Applied and *RF* voltage.
(b) Resulting current pulses and its drift across diode.

tors or amplifiers. Because of the short times of drift region, the frequency of oscillation mainly lies in microwave spectrum.

6·30. PIN DIODES

Pin diodes ($p^+ - i - n^+$) find extensive use as microwave switches, with greatly improved switching times as compared to *PN* junctions. This is achieved by inserting a high resistivity intrinsic or very lightly doped material layer of somewhat more thickness between a narrow layer of *p*-type semiconductor and *n*-type semiconductor as shown in fig. (50). Silicon is mainly used in construction of PIN diodes but sometimes gallium arsenide is also used. Metal layers are added to the construction for contact purposes.

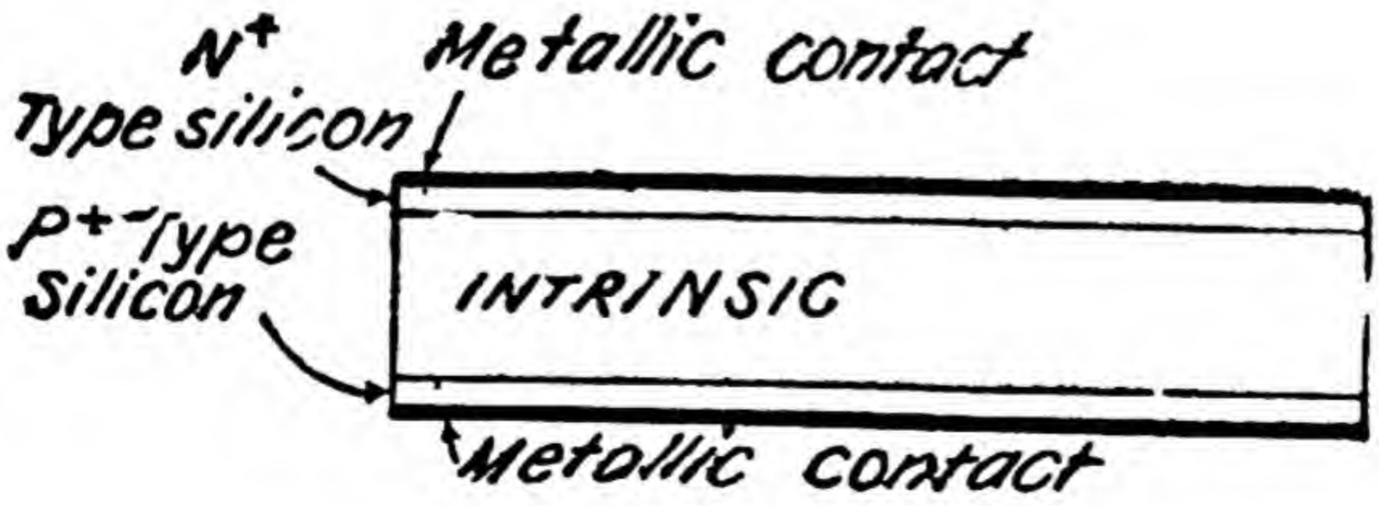
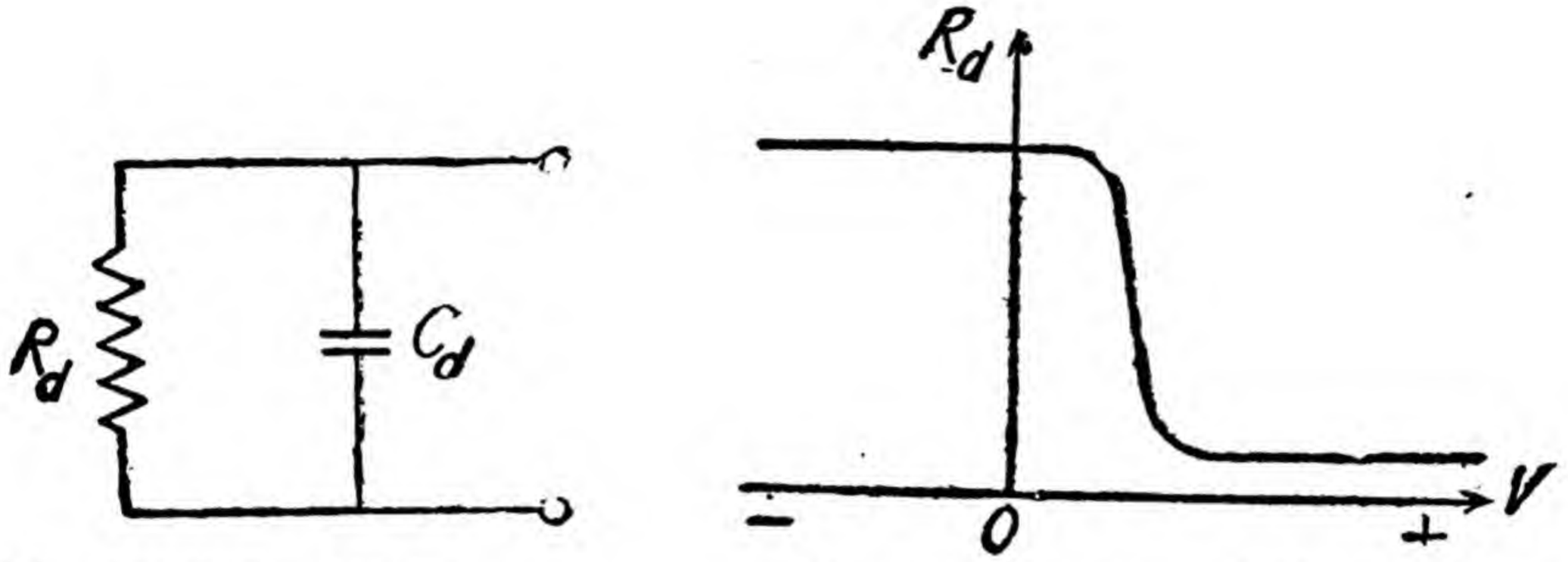


Fig. 50. Pin diode.

PIN diode acts as an ordinary diode upto 100 MH_Z. However

above this frequency it ceases to rectify because of the carrier storage in and transit time across the intrinsic region. At microwave frequencies, the diode behaves as a variable resistance. Fig. 51 shows a simplified equivalent circuit along with its resistance-voltage characteristic.



(*a*) *Equivalent circuit.* (*b*) *Resistance variation with bias voltage.*

Fig. 51.

When a reverse bias is applied on a PIN diode, its microwave resistance lies in the range of 5 to 10 $K\Omega$. On the other hand, under forward bias conditions, resistances as low as 0·5 to 10Ω can also be obtained. Thus, if a PIN diode is mounted across a 50Ω coaxial line, under reverse bias conditions, it will not significantly load the line and therefore the power flow will remain uninterrupted. Moreover, under the forward bias conditions, the resistance of PIN diode becomes very low and as a result of this only a very small part of the power is transmitted *i.e.*, most of the power is reflected. Thus the diode acts as a switch. In a similar way, it may also be used as a modulator and may be placed in series with the line or waveguide.

Following are the typical applications of PIN diode.

(1) *SPST* and *MPST* switches.
(2) Pulse modulators.
(3) Amplitude modulators.
(4) Phase shifters.
(5) Diplexers.
(6) *T-R* switches.
(7) Attenuators.
(8) Constant impedance devices.

6·31. SCHOTTKY BARRIER DIODE.

Schottky barrier diode is an extension of point contact diode. This diode is also known as hot carrier diode, a hot electron diode or an ESBAR diode (short for epitaxial Schottky barrier). The name hot electron diode is given because electrons flowing from the semiconductor to the metal have a higher energy level than electrons in the metal itself, just as the metal would if it were at a higher temperature. The most commonly used semiconductors

are silicon and gallium arsenide. They have their own merits and demerits *i.e.*, silicon is easier to fabricate while GaAs has lower noise and higher operating frequency. *n*-type epitaxial materials are used, and the metal is often a thin layer of titanium surrounded by gold and silver for protection.

Schottky barrier diode shares the advantage of the crystal rectifier in that there are no minority carriers in the reverse direction (*i.e.*, no significant current from the metal to semi-conductor when the diode is reversed biased). The delay due to the hole-electron recombination present in junction diodes is thus absent here. In Schottky barrier diodes, the junction areas are considerably larger than point contact diodes and hence the forward resistance and so the noise is comparatively lower. The applications of the diodes are the same as those of crystal rectifier *i.e.* microwave detection and mixing.

6·32. BACKWARD DIODES

A backward diode is a special form of tunnel diode in which the negative resistance peak and valley region from the tunnel diode characteristic is removed by suitable doping and etching during manufacture. The characteristic of a backward diode is shown in fig. 52. In this case the forward current is very small

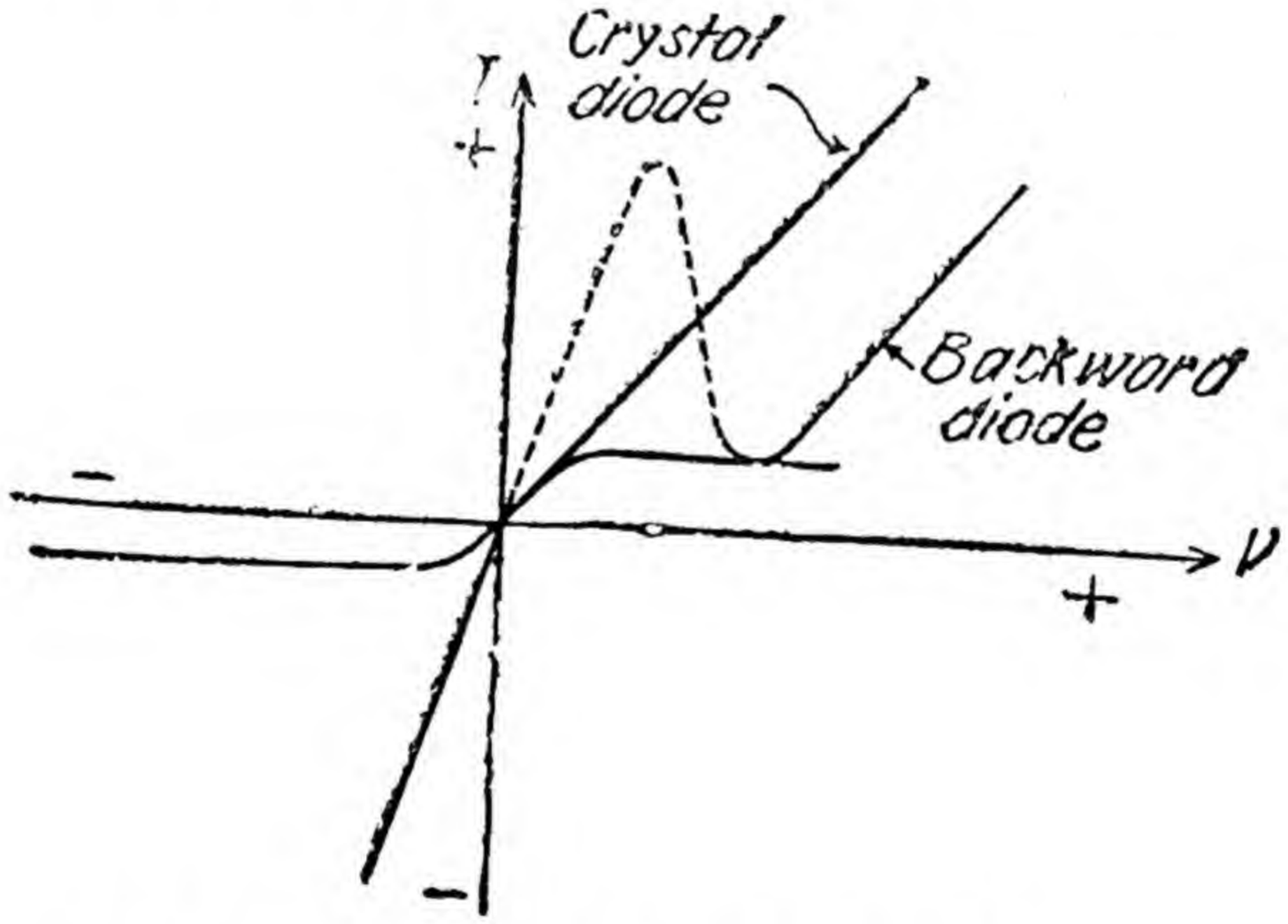


Fig. 52. Backward diode characteristic.

and becomes equivalent to the reverse current of a conventional diode. This diode can be used for a small signal rectifier. It has the advantage of not only narrow junction and therefore a high operating speed and frequency but also of a current ratio (forward to reverse) which is much higher than in conventional rectifier. This diode is sometimes called as *tunnel rectifier* which is quite common in use. This diode is used in video detection, low level mixing etc. as it is a low noise device. Mostly GaAs is used as a semiconductor material for its fabrication.

6·33. SILICON CONTROLLED RECTIFIER (SCR) :

This is a semiconductor equivalent of thyratron tube. It consists of three junctions J_1, J_2 and J_3 (J_1 and J_3 operate in forward direction while middle J_2 operates in reverse direction) and three terminals known as anode A, cathode K and gate G as shown in figure 53. In order to consider the operation of a silicon controlled rectifier, it can be regarded as consisting of one N-P-N transistor and other P-N-P transistor interconnected as shown in figure (54).

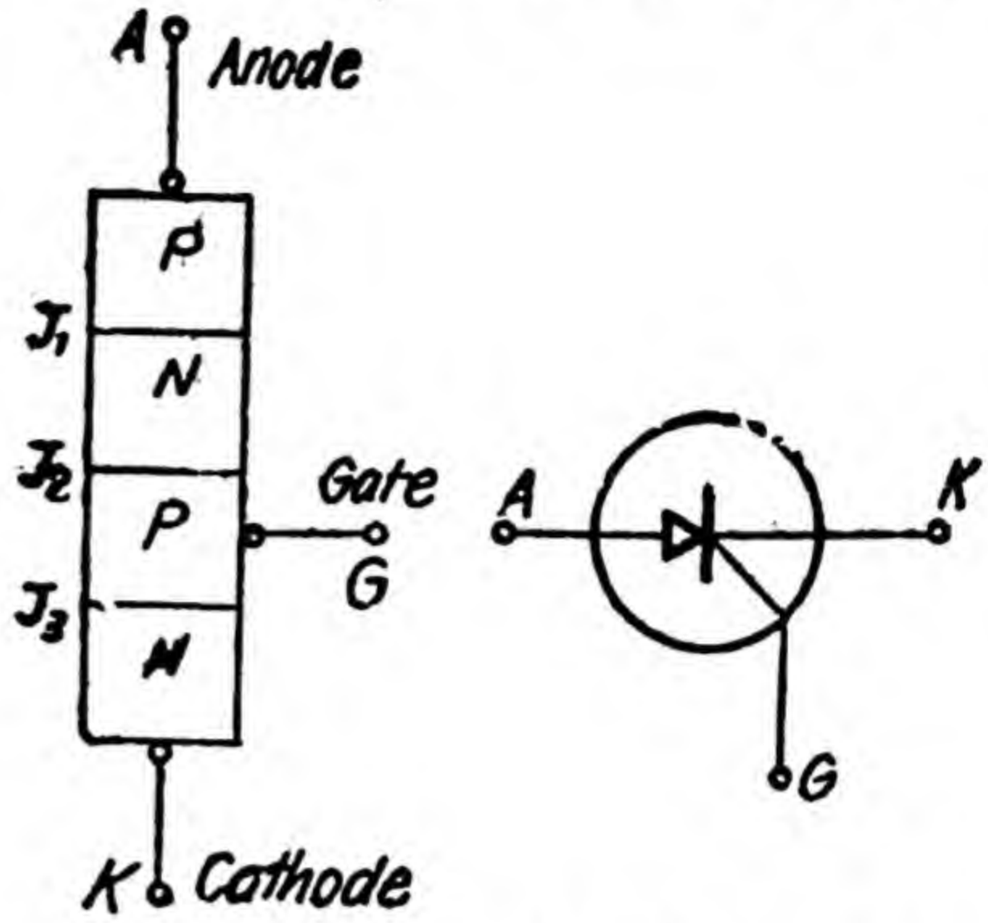


Fig. 53. Structure of SCR symbol.

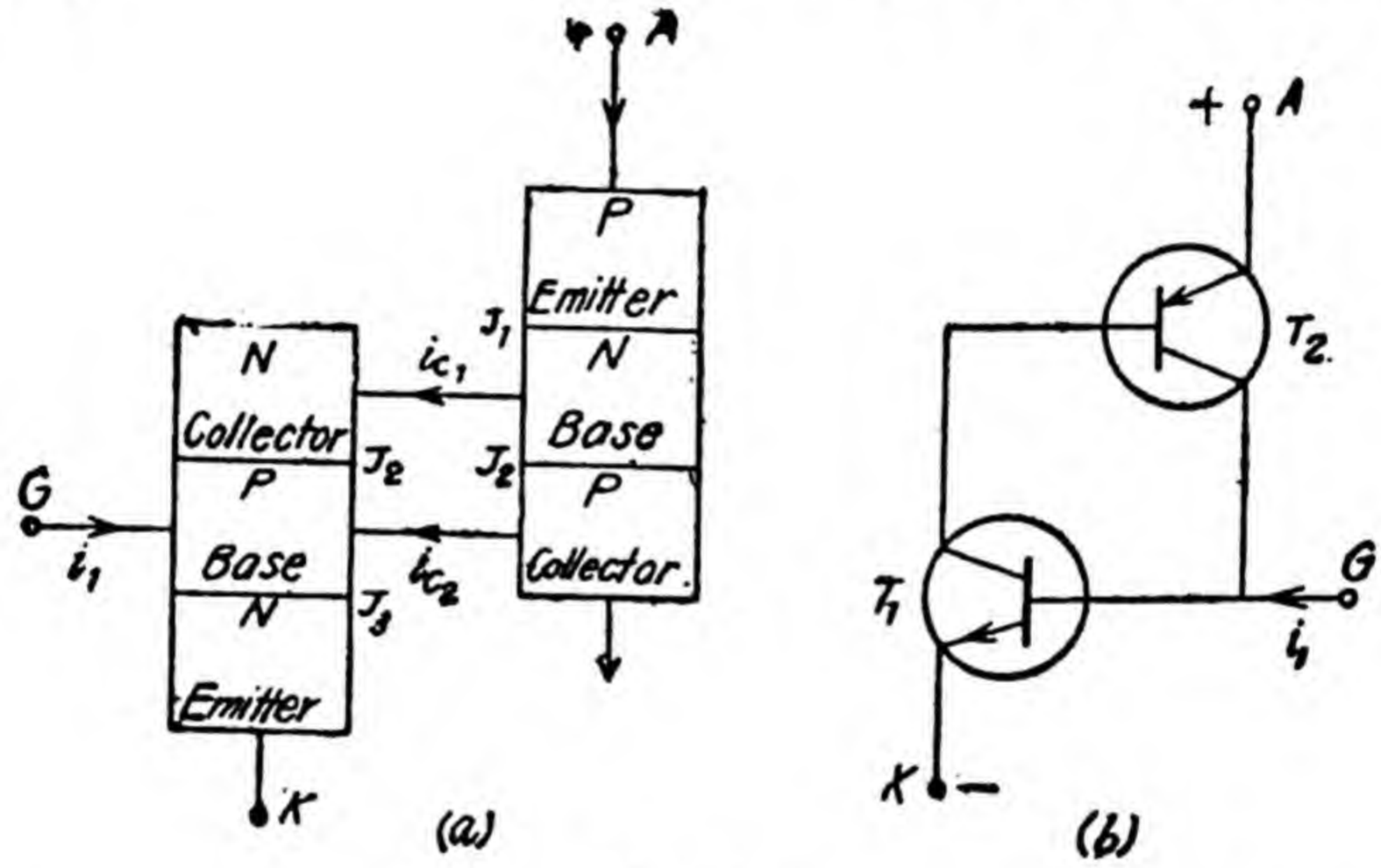


Fig. 54

It is obvious from the figure that the junction J_2 is a collector junction common to both transistors while J_1 and J_3 constitute emitter junctions of these transistors. As long as J_2 is reversed biased, the whole unit behaves as an open circuit between anode and cathode. But when inserting a current pulse at the gate G, the junction J_2 is forward biased, the unit acts as a short circuit so that the anode current is limited only by load resistance. When a positive voltage is applied at A with respect to cathode, each of the emitter junctions is forward biased while J_2 remains reverse biased. Thus the two currents i_{c_1} and i_{c_2} are very small *i.e.*, they are the leakage currents only. When grid is made positive, a large collector current i_{c_1} flows through T_1. The current i_{c_1} is the base current for T_2, hence it causes a large collector current i_{c_2} through the collector of T_2. Ultimately the conduction through both the transistors takes place which leads to saturation state. Now, even if the grid is not supplied any current pulse, the SCR continues to conduct until anode voltage is removed or anode current is reduced below a cetain holding level.

The voltage current characteristics of *SCR* is shown in fig. 55. When a reverse voltage is applied between A and K,

small leakage current flows. Similarly a small leakage current also flows when it is forward biased. When forward voltage is increased a break-down condition is reached which fires *SCR*. The firing of *SCR* can also be made by applying a short positive pulse current at the gate. This is denoted by dotted line *OB*. *SCR* remains conducting as long as the anode current is maintained larger than holding current I_H. If anode current becomes less than I_H, *SCR* turns off.

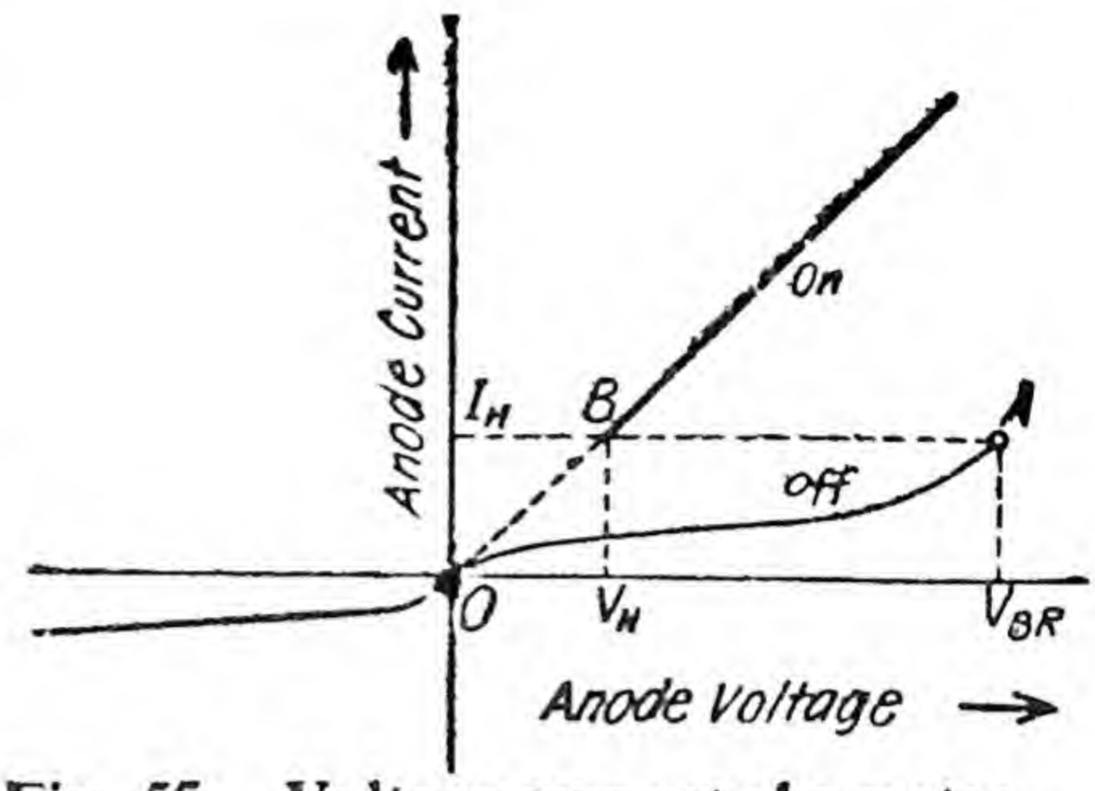


Fig. 55. Voltage current characteristic of *SCR*.

6·34. TRIAC

Fig. 56 (*a*) and (*b*) show the structure and symbol of a triac. This is a four layer device with three electrodes : terminal 1, terminal 2 and a control gate G. It is a bidirectional device as it can block voltages of either polarity and conduct current in either direction. The bidirectional property is due to its junction structure. The *N*-type emitter at terminal 2 is located directly opposite *P*-type emitter at 1 and the *P* type emitter at 2 is located directly opposite *N*-type emitter at 1. The gate terminal makes contact (ohmic) with both *P* and *N* type materials. Thus it allows the use of either a positive or a negative pulse as trigger currents. The *V-I* characteristics are shown in fig. (57).

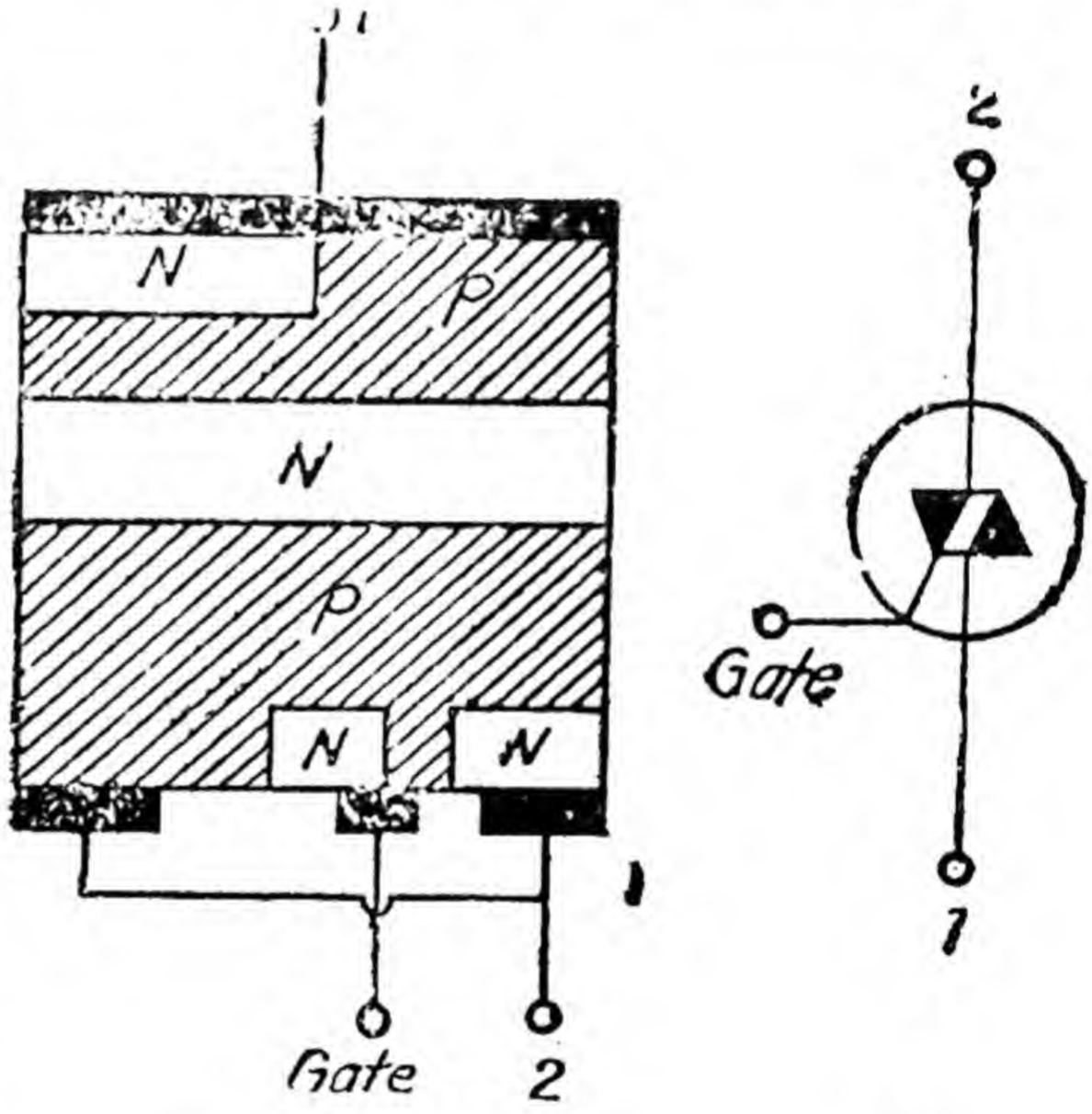


Fig. 56 (a) (b)
(a) structure of a triac (b) its symbol

The triac is said to be positively biased when terminal 2 is positive with respect to terminal 1 (operation in *I* quadrant). It is said to be negatively biased when terminal 2 is negative with respect to terminal 1 (operation in *III* quadrant). Thus the triac can be trigg-

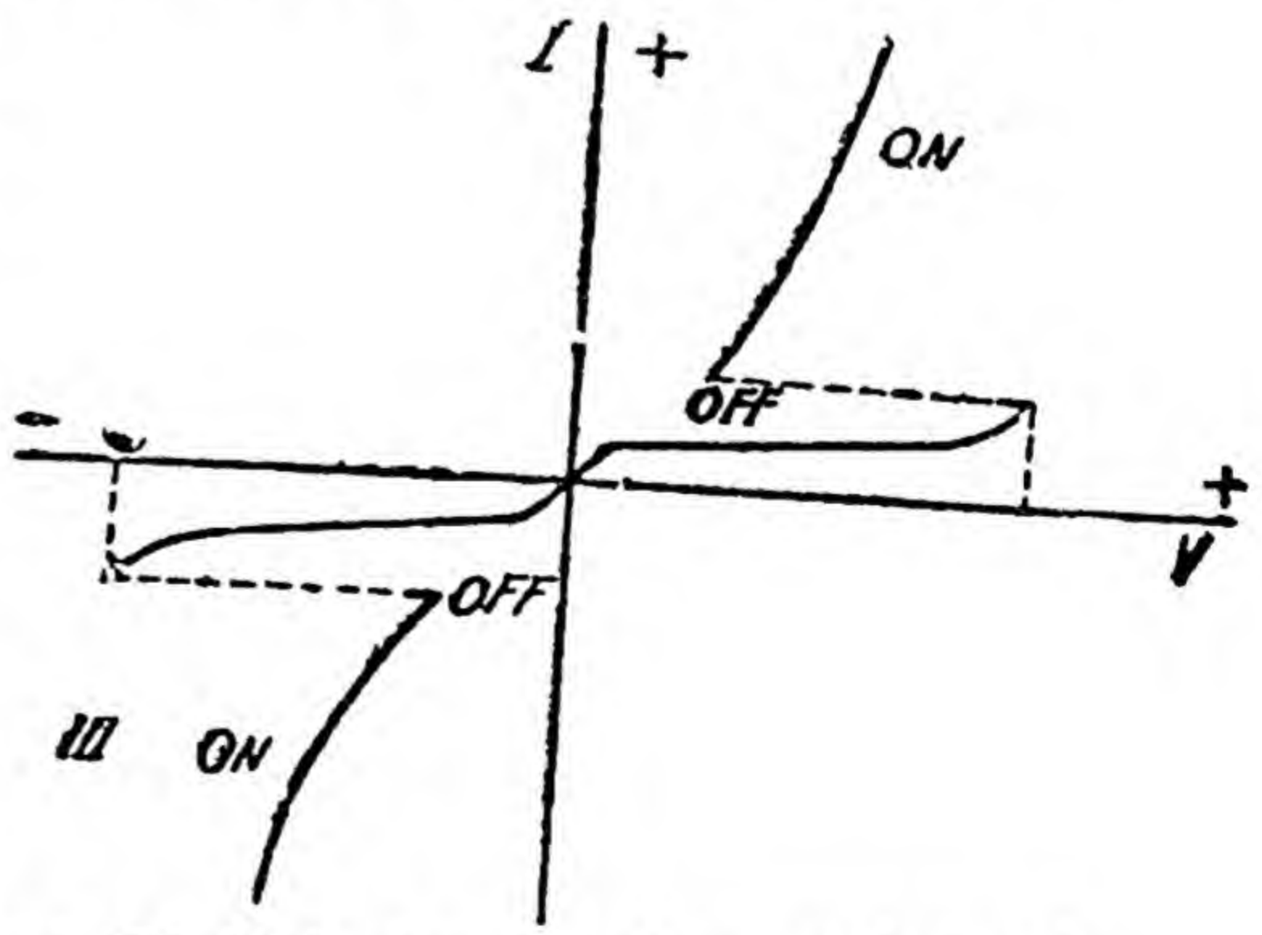


Fig. 57. Volt ampere characteristic of a triac.

ered with either positive or negative gate pulses when the anode potentials are positive or negative respectively *i.e.*, triac is an a.c. switch which can be made to conduct on both alterations (half cycles) of an a.c. voltage.

6·35. DIAC

Fig. 58 (*a*) and (*b*) show the structure and symbol of a diac. This is two terminal device having a three semiconductor layers *PNP*. It can be made to conduct in either direction and hence a bidirectional device. It has no gate terminal.

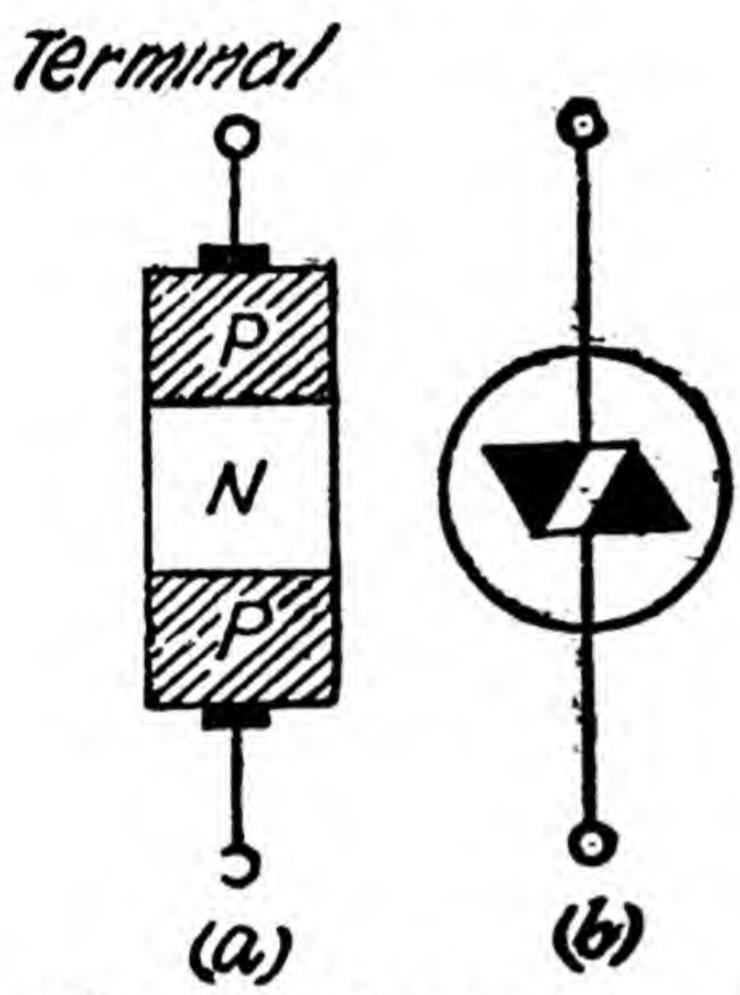


Fig. 58. (a) Structure of diac
(b) Its symbol

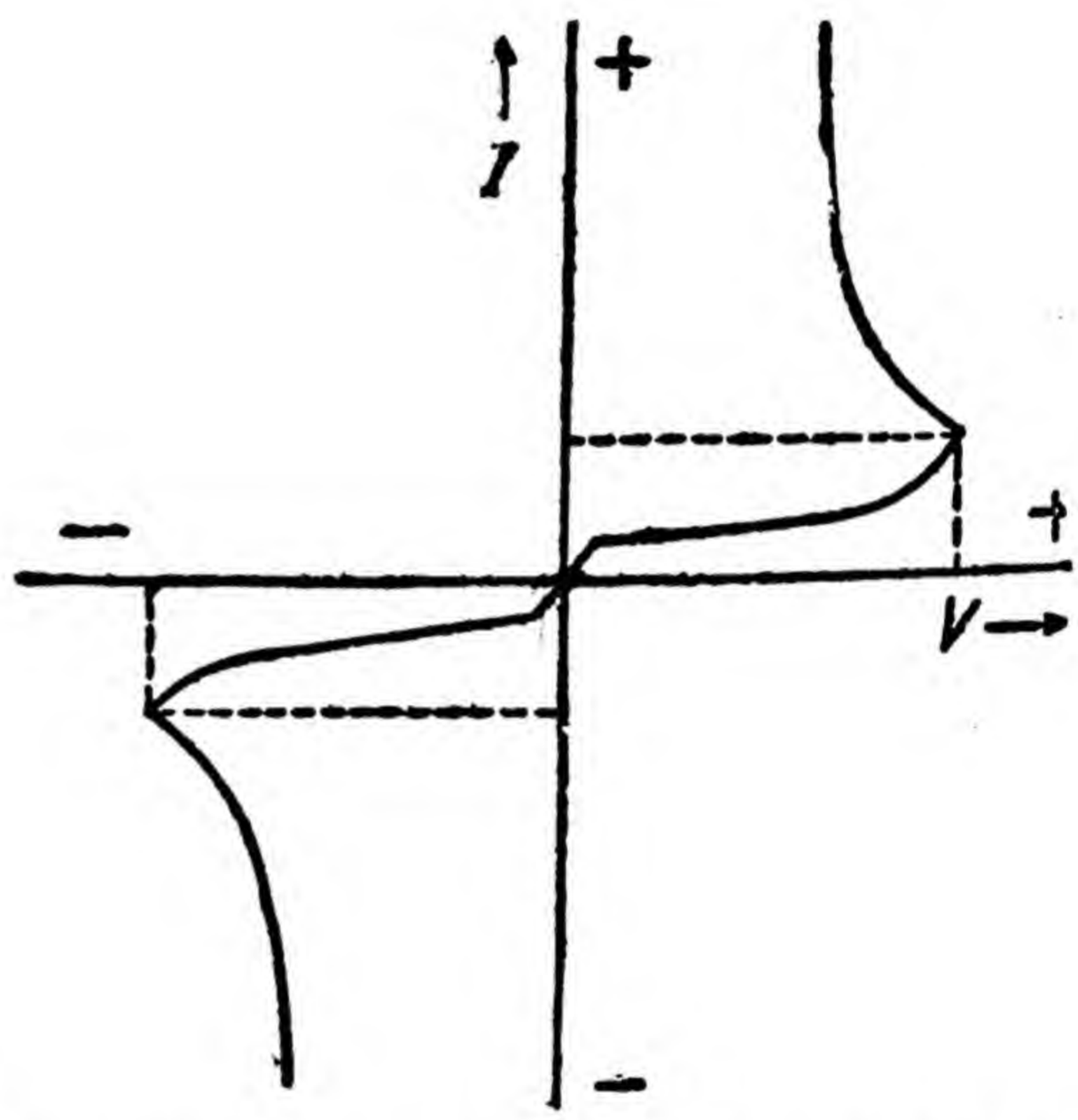


Fig. 59. Diac voltage current characteristic.

The voltage current characteristic of a diac is shown in fig. 59. The two *PN* junction of diac have equal doping characteristics. Therefore a diac produces symmetrical switching characteristics for both positive as well as negative voltages.

When the *PN* junction is reverse biased, then any applied voltage of either polarity causes a small saturation current across this junction. It is also obvious from the figure that the diac current rises sharply as the applied voltage exceeds the avalanche breakdown voltage. In this *ON* condition, the voltage across the diac decreases with increasing current showing thereby a negative resistance region. The diac is used as a trigger device in triac power control systems.

6·36. TRANSISTOR AS A FOUR POLE :

Whatever may be the circuit configuration, the transistor may be regarded as a four pole network. Considering a common base configuration, the emitter and base terminals may be regarded as input terminals, while the collector and base terminals as output

terminals. The four pole representation is advantageous because the four pole theory can be applied to the circuit analysis. This approach is more important for transistor than vacuum tube as the electrical properties of transistors vary appreciably with current, voltage and frequency; moreover, it is also difficult to find an equivalent electrical circuit, which will represent the transistor over a wide range of conditions. Figure 60 shows the transistor as a four pole. The currents are taken positive when they are entering the transistor and negative when leaving it. The dots show the assumed positive polarity of the voltage. The circuit involves two voltages v_1 and v_2 and two currents i_1 and i_2. These quantities (v_1, v_2, i_1 and i_2) can be related by equations with the circuit parameters of the transistor. For the analysis of a transistor circuit, three following sets of parameters are used :



Fig. 60. Showing the transistor as a four pole

(i) Impedance or Z parameters

(ii) Admittance or Y parameters, and

(iiii) Hybrid or H parameters.

6·37. IMPEDANCE PARAMETERS OR Z PARAMETERS :

The network behaviour is specified by v_1, i_1, v_2 and i_2 as external quantities. Any pair of these may be arbitrarily chosen as independent variables leading to two equations which can be solved for the other two dependent variables. In impedance parameters i_1 and i_2 are chosen as independent variables and v_1 and v_2 are chosen as the dependent variables. Thus

$$v_1 = f(i_1, i_2) \qquad \ldots(1)$$

$$v_2 = f(i_1, i_2) \qquad \ldots(2)$$

Writing the total differentials of equations (1) and (2), we have

$$dv_1 = \left(\frac{\partial v_1}{\partial i_1}\right)_{i_2} di_1 + \left(\frac{\partial v_1}{\partial i_2}\right)_{i_1} di_2 \qquad \ldots(3)$$

and

$$dv_2 = \left(\frac{\partial v_2}{\partial i_1}\right)_{i_2} di_1 + \left(\frac{\partial v_2}{\partial i_2}\right)_{i_1} di_2 \qquad \ldots(4)$$

Since we are interested in developing only the *a.c.* equivalent circuit, Δ quantities may be replaced by the symbols for the instantaneous values of these variable quantities ; thus

$$\Delta v_1 = V_1 = \left(\frac{\partial v_1}{\partial i_1}\right)_{i_2} di_1 + \left(\frac{\partial v_1}{\partial i_2}\right)_{i_1} di_2$$

or

$$V_1 = Z_{11} I_1 + Z_{12} I_2 \qquad \ldots(5)$$

and

$$\Delta v_2 = V_2 = \left(\frac{\partial v_2}{\partial i_1}\right)_{i_2} di_1 + \left(\frac{\partial v_2}{\partial i_2}\right)_{i_1} di_2$$

or $$V_2=Z_{21}\,I_1+Z_{22}\,I_2. \qquad \ldots(6)$$

Equations (5) and (6) may be represented in matrix form as

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix}=\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \qquad \ldots(7)$$

The quantities Z_{11}, Z_{12}, Z_{21} and Z_{22} are the open circuited impedances and may be defined as

$Z_i=Z_{11}=\left(\frac{\partial v_1}{\partial i_1}\right)_{i_2}=\frac{V_1}{I_1}$=open circuit input impedance, $I_2=0$

$Z_r=Z_{12}=\left(\frac{\partial v_1}{\partial i_2}\right)_{i_1}=\frac{V_1}{I_2}$=open circuit reverse transfer impedance $I_1=0$, ...(8)

$Z_f=Z_{21}=\left(\frac{\partial v_2}{\partial i_1}\right)_{i_2}=\frac{V_2}{I_1}$=open circuit forward transfer impedance, $I_2=0$,

$Z_0=Z_{22}=\left(\frac{\partial v_2}{\partial i_2}\right)_{i_1}=\frac{V_2}{I_2}$=open circuit output impedance, $I_1=0$.

From equation (5) it is clear tnat the voltage V_1 results due to the flow of current I_1 through the input impedance Z_{11} and due to flow of current I_1 through the reverse transfer impedance Z_{12}. Similarly, from equation (6), we may conclude that voltage V_2 results due to the flow of current I_1 through the forward transfer impedance Z_{21} and the flow of current I_2 through the output impedance Z_{22}. Figure 61 shows a two generator form of equivalent circuit for transistor derived from equations (5) and (6). In general, one generator equivalent circuit may be obtained by algebraic manipulation of equations (5) and (6). Adding and subtracting Z_{12}, I_1 in equation (5), we have

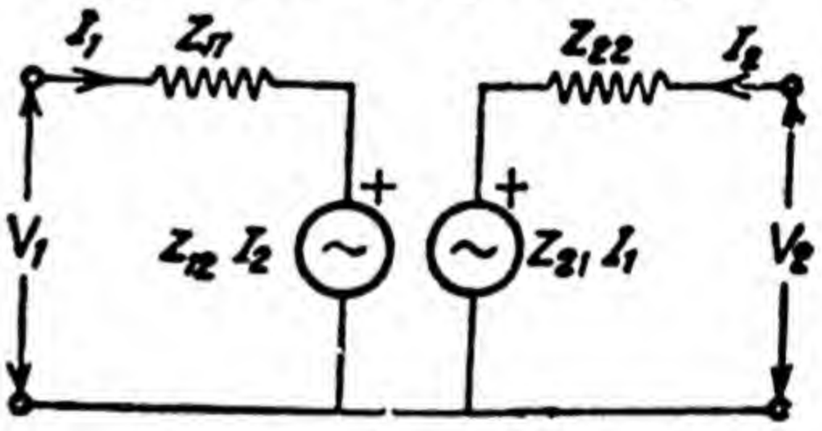


Fig. 61. The Z parameter equivalent circuit.

$$V_1=(Z_{11}-Z_{12})\,I_1+Z_{12}\,(I_1+I_2). \qquad \ldots(9)$$

Again adding and subtracting $(Z_{12}I_1+Z_{12}I_2)$ in equation (6), we have $V_2=(Z_{21}-Z_{12})I_1+(Z_{22}-Z_{12})I_2+Z_{12}(I_1+I_2)$ ..(10)

The one generator form of equivalent circuit may be obtained with the help of equations (9) and (10) which is shown in fig. 61.

The voltage generator of open circuit $(Z_{21}-Z_{12})\ I_1$ (fig. 62) may be replaced by a current generator of short circuit current $[(Z_{21}-Z_{12})/(Z_{22}-Z_{12})]\times I_1$ in shunt with the impedance $(Z_{22}-Z_{12})$ and then the equivalent circuit takes the form as shown in fig. 63.

With the help of equations (5) and (6), we have

Input impedance $Z_i=\frac{V_1}{I_1}=Z_{11}+Z_{12}\,\frac{I_2}{I_1}$ .. (11)

and output impedance $Z_0=\frac{V_2}{I_2}=Z_{22}+Z_{21}\,\frac{I_1}{I_2}$. ...(12)

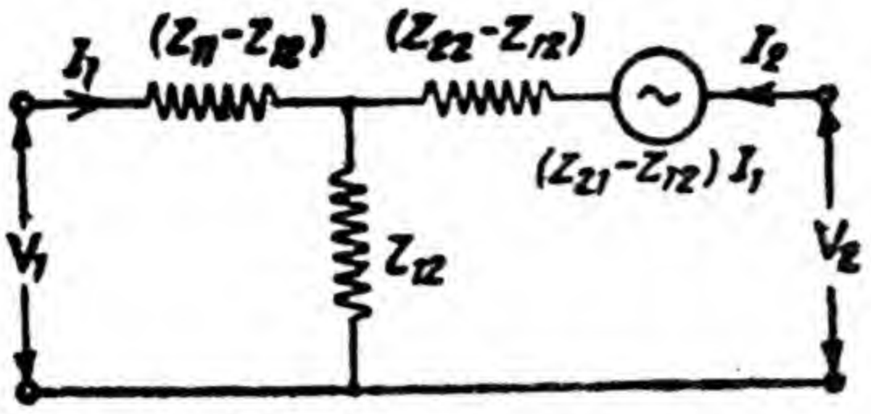


Fig. 62. One generator form of tra sistor.

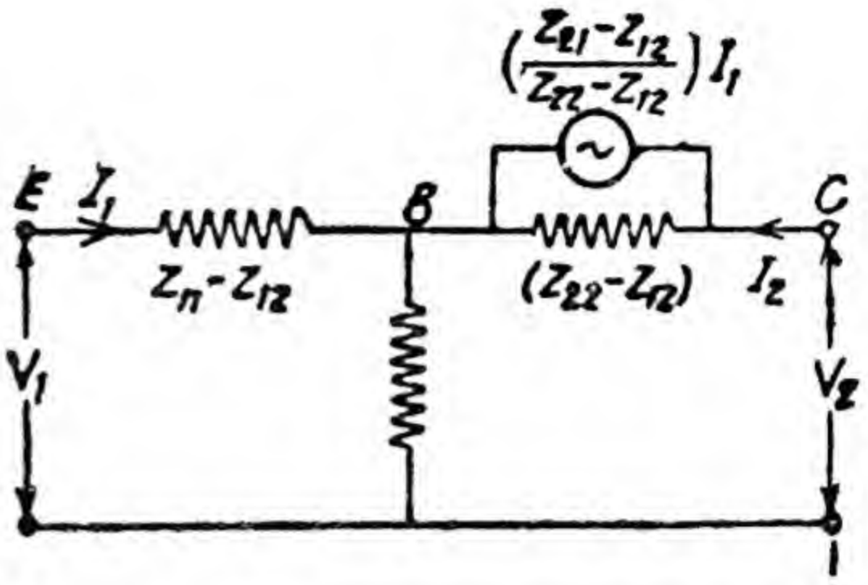


Fig. 63. Current generator form of equivalent circuit of trans.stor.

Let us consider that load impedance Z_1 is connected across the output terminals, then

$$V_2 = -I_2 Z_L \qquad \ldots(13)$$

or
$$\frac{V_2}{I_2} = -Z_L = Z_{22} + Z_{21}\,\frac{I_1}{I_2}.$$

Solving it we have

$$\frac{I_2}{I_1} = \frac{-Z_{21}}{Z_{22}+Z_L}. \qquad \ldots(14)$$

Thus the current gain A_I is given by

$$A_I = \frac{I_2}{I_1} = \frac{-Z_{21}}{Z_{22}+Z_L}. \qquad \ldots(15)$$

The negative sign appears due to the conventional direction of flow of I_2.

The input impedance in this case can be expressed as

$$Z_i = Z_{11} + Z_{12}\,\frac{I_2}{I_1} = Z_{21} - \frac{Z_{12}Z_{21}}{Z_{22}+Z_L}. \qquad \ldots(16)$$

The output impedance Z_0 can be obtained by looking from the output terminals into the circuit. If Z_S is the source impedance, then

$$\frac{V_1}{I_1} = -Z_S = Z_{11} + Z_{12}\,\frac{I_2}{I_1}$$

or
$$\frac{I_1}{I_2} = \frac{-Z_{12}}{Z_{11}+Z_S}.$$

$$\therefore \quad Z_0 = Z_{22} - \frac{Z_{12}Z_{21}}{Z_{22}+Z_S}. \qquad \ldots(17)$$

The voltage gain is given by

$$A_v = \frac{V_2}{V_1} = \frac{-Z_L}{Z_i}\,\frac{I_2}{I_1} = \frac{Z_L\,Z_{21}}{(Z_{22}+Z_L)\,Z_i}. \qquad \ldots(18)$$

Power gain $$A_P = \left(\frac{I_2}{I_1}\right)^2 \frac{Z_L}{Z_i} = A_I^2\,\frac{Z_L}{Z_i}. \qquad \ldots(19)$$

6·38. ADMITTANCE PARAMETERS OR Y-PARAMETERS :

In this type of parameters, v_1 and v_2 are taken as the independent variables, while the currents i_1 and i_2 are regarded as dependent variables. All the dimensions are those of admittance (mhos). The general functional relations are

$$i_1 = f(v_1, v_2),$$
$$i_2 = f(v_1, v_2).$$

Taking the total differentials as before

$$di_1 = \left(\frac{\partial i_1}{\partial v_1}\right)_{v_2} dv_1 + \left(\frac{\partial i_1}{\partial v_2}\right)_{v_1} dv_2 \quad \ldots(1)$$

and

$$di_2 = \left(\frac{\partial i_2}{\partial v_1}\right)_{v_2} dv_1 + \left(\frac{\partial i_2}{\partial v_2}\right)_{v_1} dv_2. \quad \ldots(2)$$

If a.c. currents are employed and operation is over a range in which the slopes of the volt-amp ere curves or the partial derivatives are constant, then

$$I_1 = Y_{11} V_1 + Y_{12} V_2, \quad \ldots(3)$$
$$I_2 = Y_{21} V_1 + Y_{22} V_2 \quad \ldots(4)$$

or

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}. \quad \ldots(5)$$

The Y parameters may be defined from measurement on the four terminal network or from the above equations, by using short circuit terminations. For example if 2, 2 terminals be short circuited, then $V_2 = 0$, and $Y_{11} = I_1/V_1$. Similarly, other parameters can be obtained. The definitions of admittance parameters may be summarized as :

$$Y_i = Y_{11} = \left(\frac{\partial i_1}{\partial v_1}\right)_{v_2} = \frac{I_1}{V_1} = \text{short circuit admittance, } V_2 = 0.$$

$$Y_r = Y_{12} = \left(\frac{\partial i_1}{\partial v_2}\right)_{v_1} = \frac{I_1}{V_2} = \text{short circuit reverse transfer admittance, } V_1 = 0,$$

$$Y_f = Y_{21} = \left(\frac{\partial i_2}{\partial v_1}\right)_{v_2} = \frac{I_2}{V_1} = \text{short circuit forward transfer admittance, } V_2 = 0, \quad \ldots(6)$$

$$Y_0 = Y_{22} = \left(\frac{\partial i_2}{\partial v_2}\right)_{v_1} = \frac{I_2}{V_2} = \text{short circuit output admittance, } V_1 = 0.$$

The two generator form of the equivalent circuit for a transistor derived by using Y parameters with the help of equations (3) and (4) is shown in figure 64.

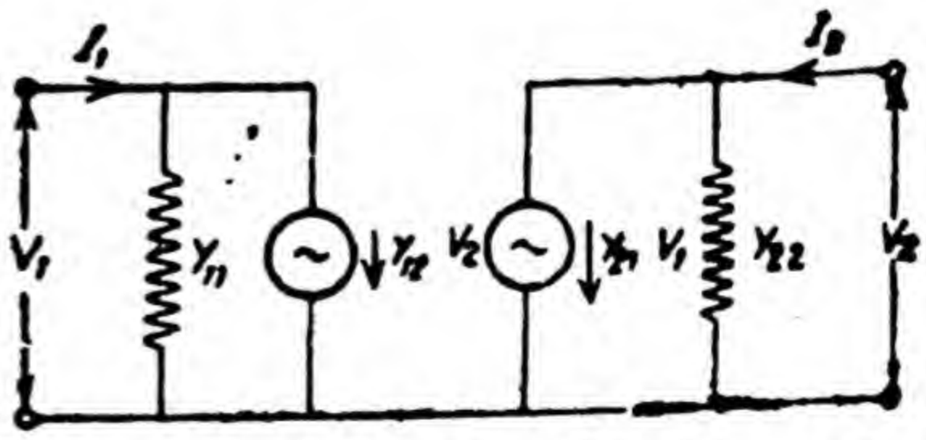


Fig. 64. Two generator form of equivalent circuit.

6·39. HYBRID PARAMETERS OR h PARAMETERS :

As the transistor has low input impedance and high output impedance, the use of Z and Y parameters becomes awkward at high frequencies. The hybrid h parameters (h for small signal device and H for large signal units) appear most useful for circuit analysis because they form a combination of impedance and admittance parameters, and are selected to suit the low input and high output impedances of the transistor.

In this case v_2 and i_1 are chosen as independent variables while v_1 and i_2 are taken as dependent variables. The different parameters have different dimensions and hence they are known as hybrid parameters,

$$v_1 = f(i_1, v_2),$$
$$i_2 = f(i_1, v_2).$$

Taking the total differentials :

$$dv_1 = \left(\frac{dv_1}{\partial i_1}\right)_{v_2} di_1 + \left(\frac{\partial v_1}{dv_2}\right)_{i_1} dv_2 \qquad \ldots(1)$$

and

$$di_2 = \left(\frac{\partial i_2}{\partial I_1}\right)_{v_2} di_1 + \left(\frac{\partial i_2}{\partial v_2}\right)_{i_1} dv_2. \qquad \ldots(2)$$

Again, we have

$$V_1 = h_{11}\, I_1 + h_{12}\, V_2, \qquad \ldots(3)$$
$$I_2 = h_{21}\, I_1 + h_{22}\, V_2 \qquad \ldots(4)$$

or

$$\begin{bmatrix} V_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2 \end{bmatrix}. \qquad \ldots(5)$$

The h parameters are defined as

$h_{11} = h_i = \left(\frac{\partial v_1}{\partial i_1}\right)_{v_2} = \frac{V_1}{I_1}$ = short circuit input impedance, $V_2 = 0$,

$h_{12} = h_r = \left(\frac{\partial v_1}{\partial v_2}\right)_{i_1} = \frac{V_1}{V_2}$ = open circuit reverse voltage ratio,

$I_1 = 0$, $\qquad \ldots(6)$

$h_{21} = h_f = \left(\frac{\partial i_2}{\partial i_1}\right)_{v_2} = \frac{I_2}{I_1}$ = short circuit forward current ratio,

$V_2 = 0$,

$h_{22} = h_0 = \left(\frac{\partial I_2}{\partial v_2}\right)_{i_1} = \frac{I_2}{V_2}$ = open circuit output admittance, $I_1 = 0$.

The subscripts are modified in this case. The first letter indicates whether the particular parameter is input, output, reverse transfer or forward transfer ; the second describe the transistor configuration. For example, in common base, following hybrid parameters are used h_{ib}, h_{rb}, h_{fb}, and h_{ob}. Similarly,

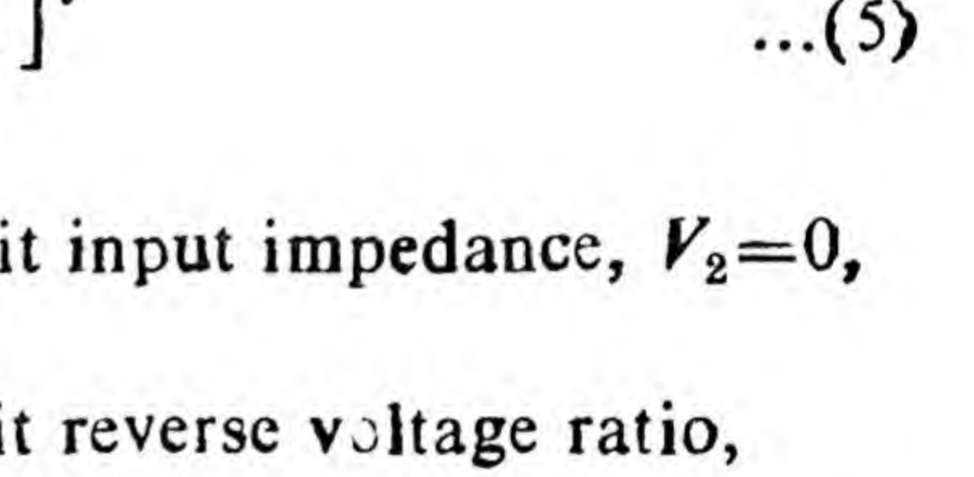


Fig. 65 Two generator-form of equivalent circuit using h parameters.

h parameters for common emitter configuration are h_{ie}, h_{re}, h_{fe} and h_{oe}. The two generator form of the equivalent circuit is shown in figure 65.

These parameters can be obtained with the help of the transistor characteristics as follows :

Figure 66 shows the circuit arrangement for plotting the different characteristics of a common emitter transistor. There are four variables *i.e.* V_{be}, V_{ce}, I_e and I_b and hence four families of characteristics can be drawn. These are :

(1) collector voltage V_{ce} versus collector current I_c with base current I_b as a parameter.

(2) collector voltage V_{ce} versus base voltage V_{be} with base current I_b as a parameter.

(3) base current I_b versus base voltage V_{be} with collector voltage V_{ce} as parameter.

(4) base current I_b versus collector current I_c with collector voltage V_{ce} as parameter.

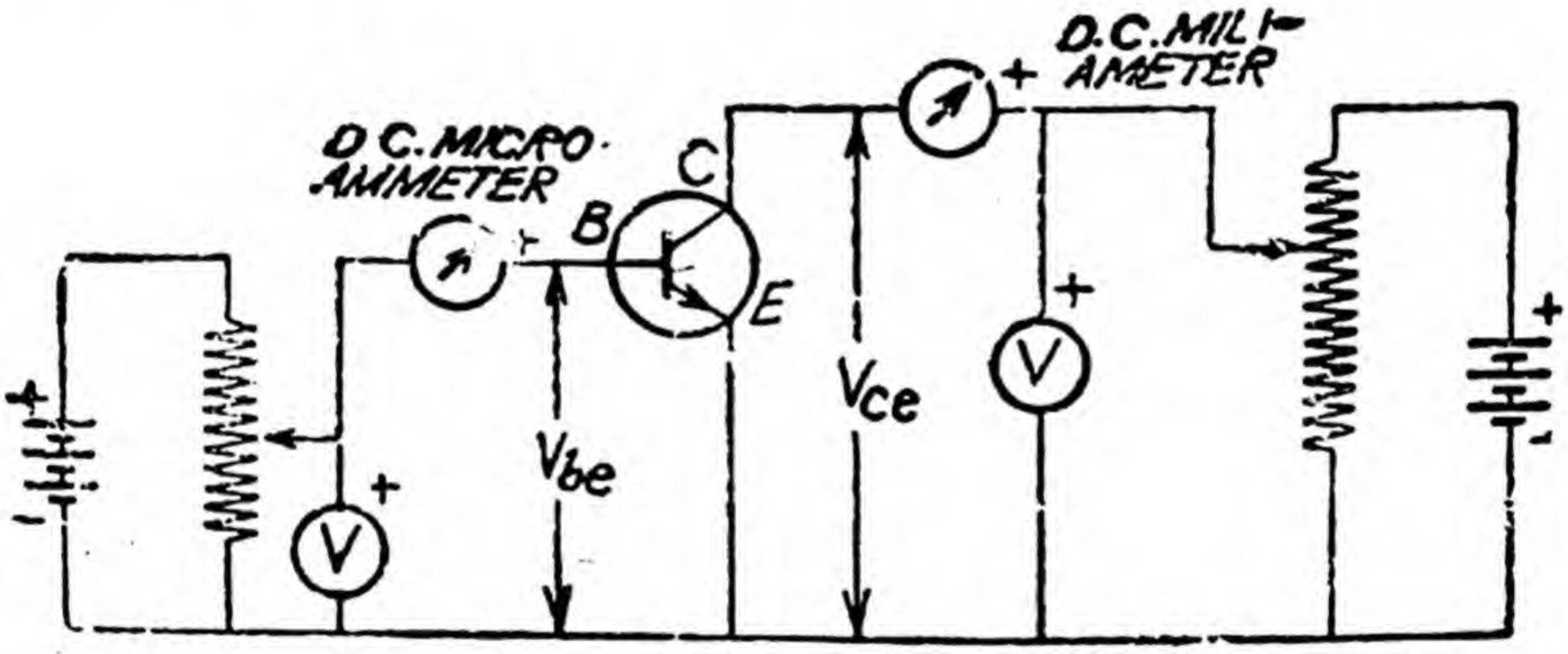


Fig. 66. Circuit for plotting the characteristics of common emitter transistor.

With the help of these characteristics, following hybrid parameters can be defined.

(1) Output admittance h_{oe} :

It is defined as the ratio between change in collector current and the corresponding change in collector voltage when base current is constant. Thus

$$h_{oe}=\left(\frac{\partial I_c}{\partial V_{ce}}\right)_{I_b}.$$

(2) Reverse voltage gain h_{re} :

It is defined as the ratio between the change in the base voltage and the corresponding change in collector voltage when base current is constant. Thus

$$h_{re}=\left(\frac{\partial V_{be}}{\partial V_{ce}}\right)_{I_b}.$$

(3) Forward current gain h_{fe} :

It is defined as the ratio between the change in collector current and corresponding change in base current, when collector voltage is constant. Thus

$$h_{fe}=\left(\frac{\partial I_c}{\partial I_b}\right)_{V_{ce}}.$$

(4) Input Impedance h_{ie} :

It is defined as the ratio between the change in the base voltage and the corresponding change in base current when collector voltage is constant. Thus

$$h_{ie}=\left(\frac{\partial V_{be}}{\partial I_b}\right)_{V_{ce}}.$$

6·40. LOAD LINE OF A DIODE CIRCUIT

DC Applied Voltage—Fig. (67 *a*) shows a basic diode circuit. The circuit consists of a diode D in series with a load resistor R_L and input battery. Fig (67 *b*) shows the static curve of the diode *i.e.* a graph drawn between V_D (voltage across diode D) and current in the circuit. Applying Kirchhoff's law to circuit (67 *a*), we have

$$V=V_D+V_o=V_D+I.R_L \quad ...(1)$$

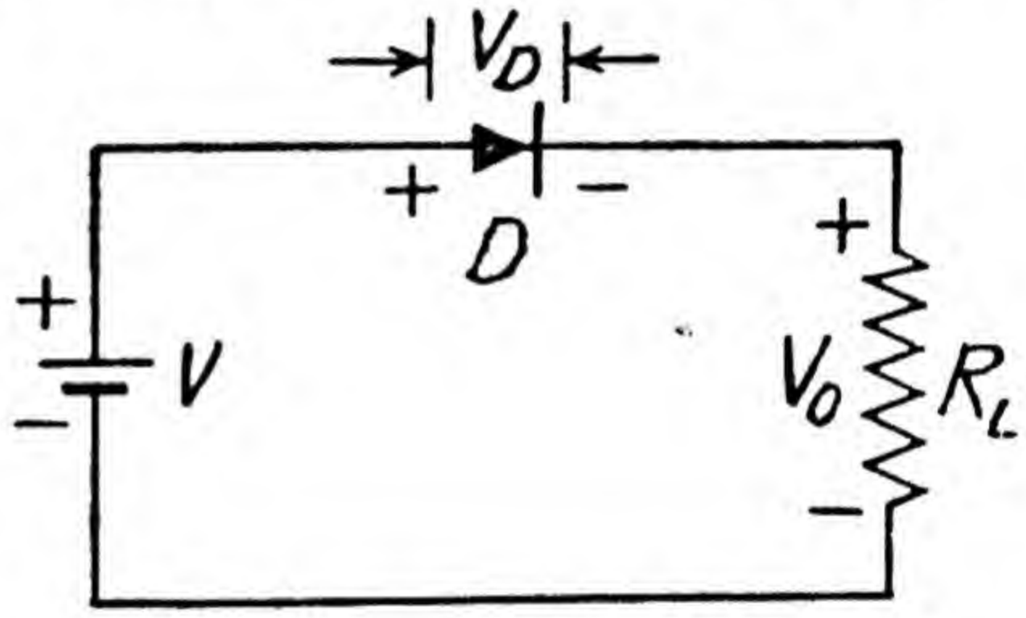


Fig. 67 (a) Basic diode circuit.

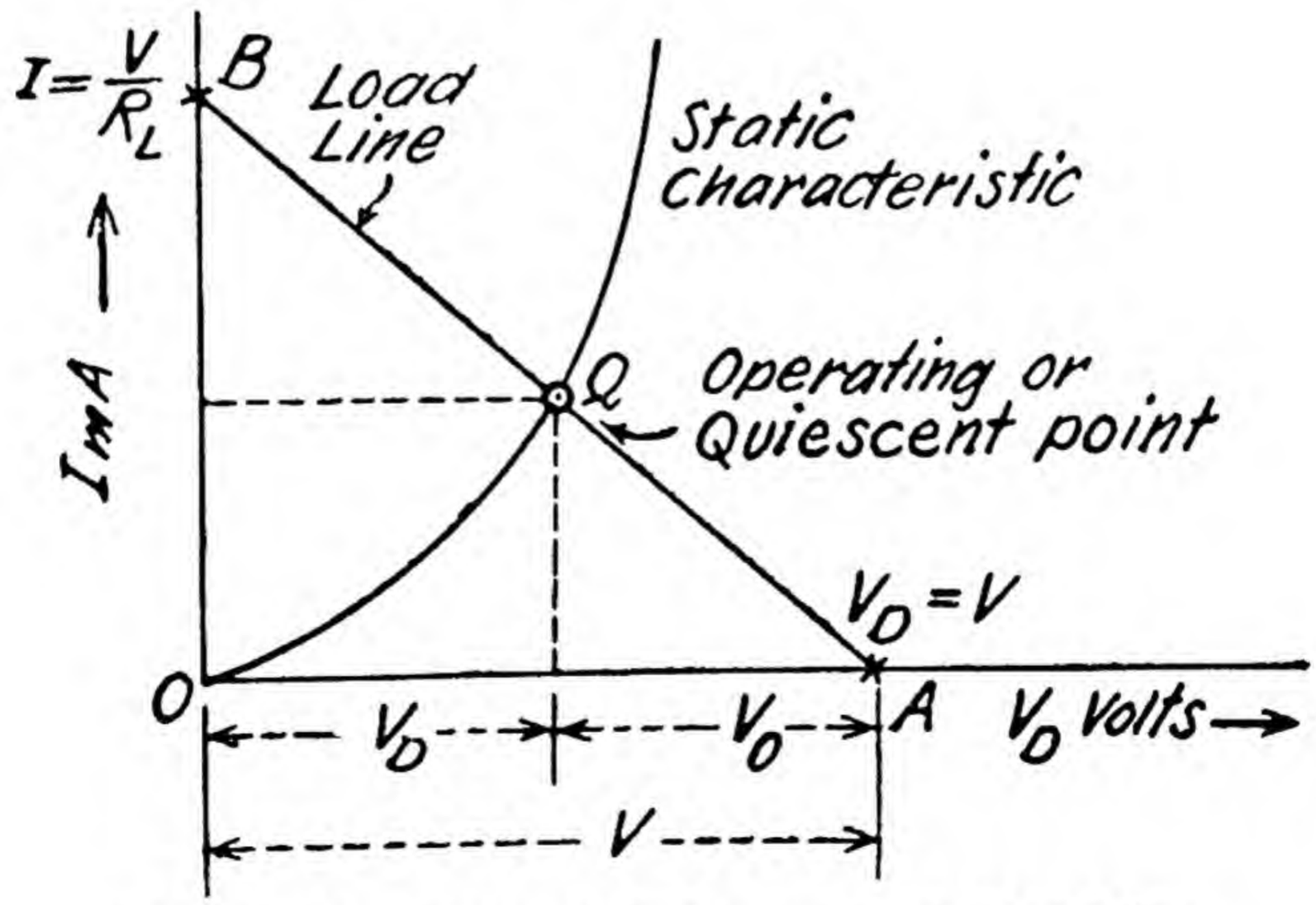


Fig. 67 (b) Characteristic curve with load line.

When $I=0$, then $V=V_D$. This gives a point $A\ (V_D, 0)$ on the

horizontal axis. Now consider $V_D=0$, then $I=V/R_L$ which gives another point B $(0, V/R_L)$ on vertical axis.

A line pas ing through the points A and B gives all possible combinations of I and V_D that satisfy eq. (1). The line is known as load line. The intersection of load line with static characteristic curve provides a point Q known as quiescent point or operating point.

From eq. (1),

$$I=\frac{V}{R_L}-\frac{V_D}{R_L} \quad \text{or} \quad I=-\frac{1}{R_L}V_D+\frac{V}{R_L} \qquad ...(2)$$

Comparing eq. (2) with the standard equation of a straight line *i.e.*

$$y=mx+c, \qquad ...(3)$$

we observe that the slope of the load line is $(-1/R_L)$. The load line intercept the y axis at V/R_L. Thus for a given input voltage, the effect of a higher load will be a steeper load line and hence a higher value of current at the operating point Q.

AC Applied voltage (Dynamic characteristic)—Fig. (68 *a*) shows the basic diode circuit in which the input voltage is a variable quantity. Our aim is to determine the output current waveform. Fig. (68 *b*) shows the characteristic curve for the diode. For a convenient interval of instantaneous voltage v_i, we draw different load lines. It is also obvious from the figure that the load lines are parallel to each other since R_L is constant. The intersection of different load lines with static curve of diode will give the current in the circuit corresponding to each value of input voltage. With the help of these currents, the output current waveform may be considered. But the following procedure is more convenient :

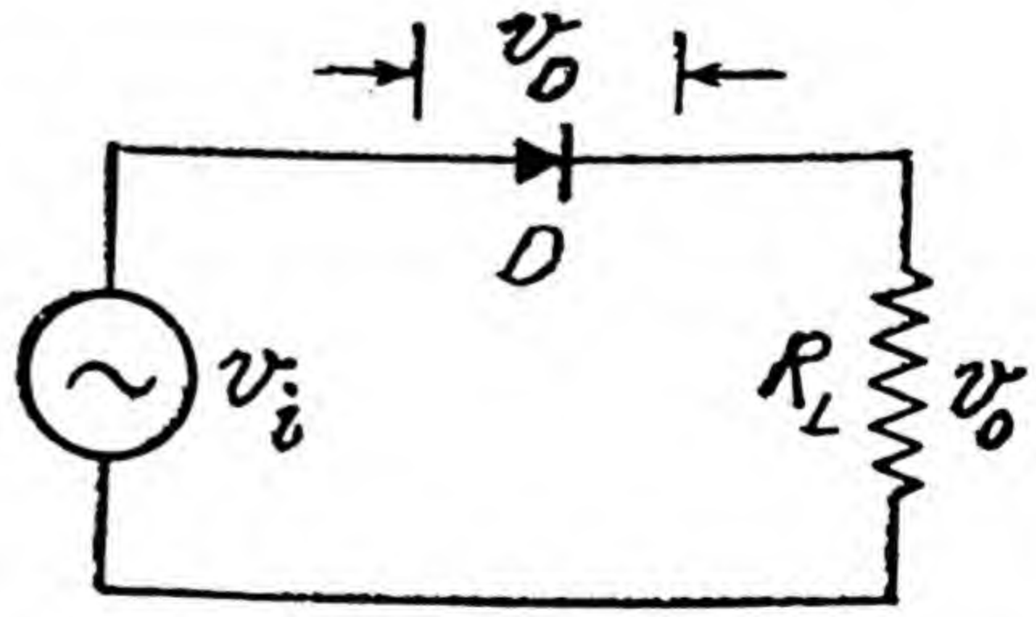


Fig. 68 (a) Basic diode circuit.

It is obvious from fig. (68 *b*) that the load lines for v_i and v_i' intersect the static characteristic at points B and A respectively. From v_i and v_i' we draw vertical lines and horizontal lines passing through B and A. The horizontal and vertical lines give the points B' and A'. Similarly other points may be considered. A curve passing through the points A', B', etc. is known as Dynamic curve or Dynamic characteristic. It should be remembered that the dynamic curve applies only to the circuit containing the value of load resistance for which it is drawn.

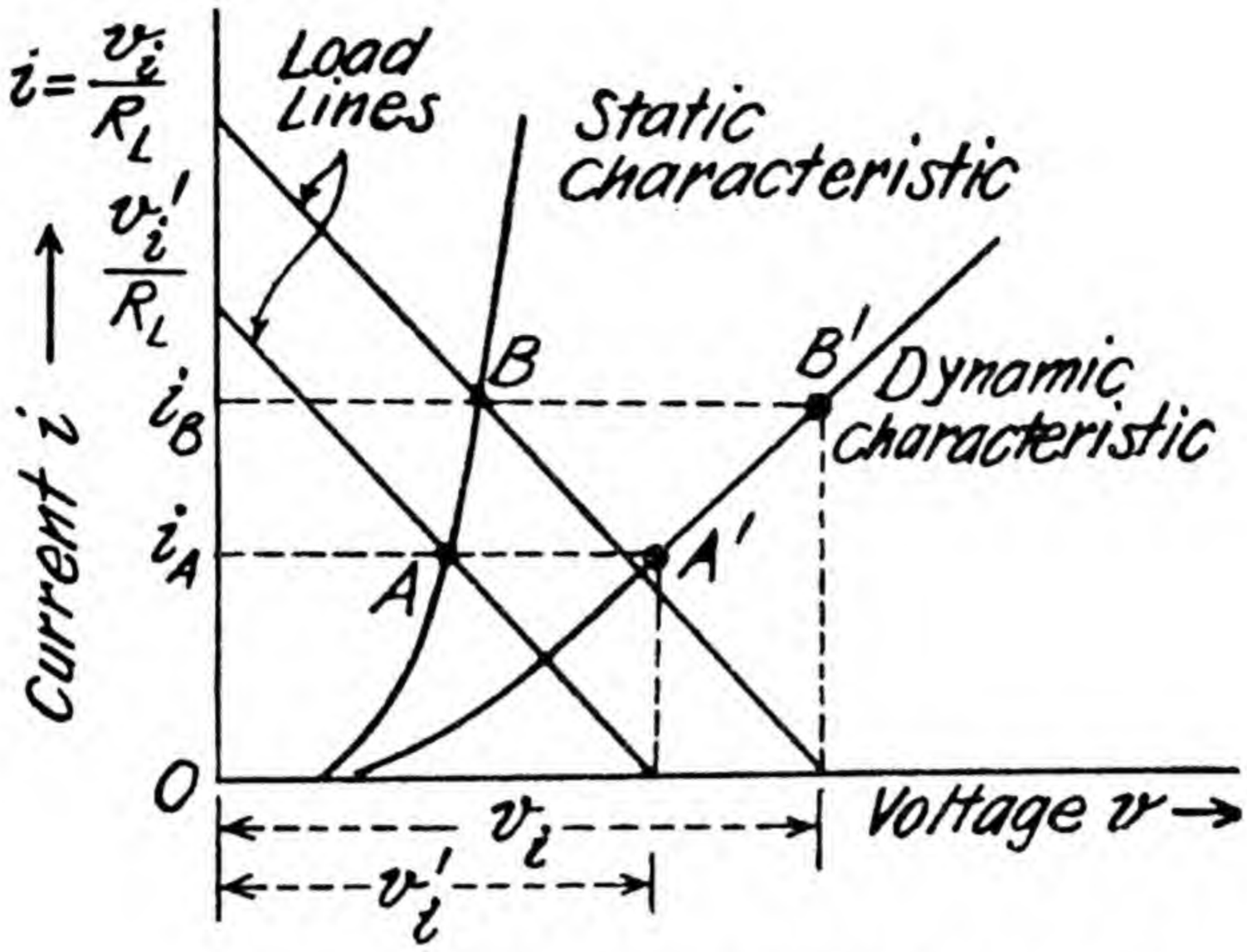


Fig. 68 (b). Dynamic characteristic.

Transfer characteristic—The curve which relates the output voltage to the input voltage is known as transfer characteristic. Fig. 69 gives a graphical method for finding the output voltage waveform for a given input waveform using dynamic characteristic.

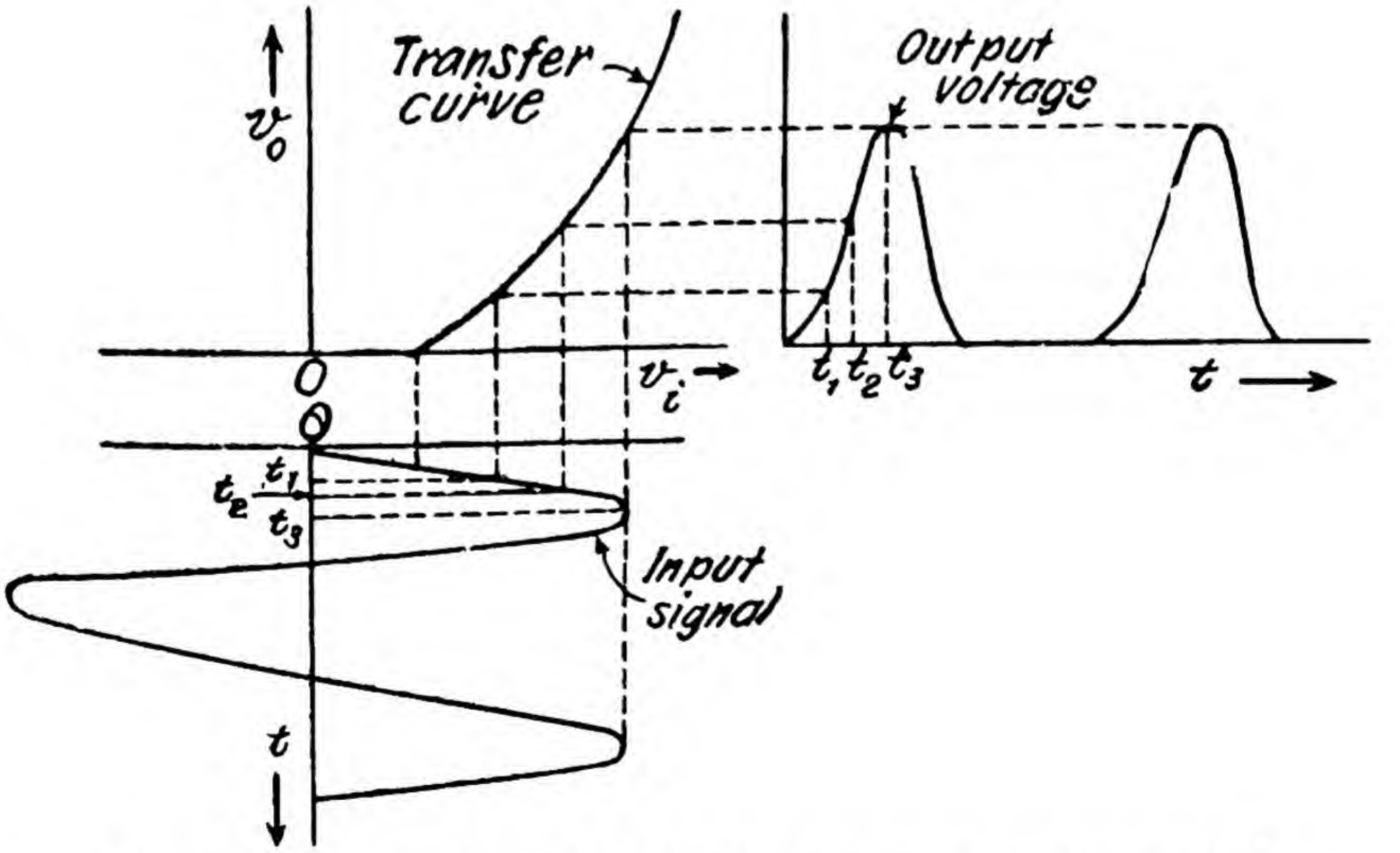


Fig. 69. Graphical method for obtaining output voltage waveform using transfer curve.

In fig. 69 the input voltage waveform has been drawn with time axis vertically and voltage axis horizontally. The different input voltages are considered at different times t_1, t_2, t_3 etc. and

corresponding vertical lines are drawn to meet the transfer curve. Intersections with the curve are drawn horizontally to give the values of output voltage corresponding to times t_1, t_2, t_3...etc. Next the output voltage waveform is plotted.

It is clear from the fig. 69 that for a part of the input signal, the output signal does not appear *i.e.*, the diode acts as a *clipper*. It is also obvious that the output voltage at the start is distorted due to nonlinear transfer characteristic.

EXERCISES AND PROBLEMS

1. Define the terms (a) conductor, (b) insulator, (c) semiconductor. Discuss what is meant by *P* and *N* type semiconductors. Explain the movement of electrons and holes in a semiconductor.
2. Explain the mechanism of electrical conduction in a typical semi-conductor like germanium or silicon. How the conductivity of pure semiconductor is affected by adding traces of trivalent and pentavalent impurities ?
3. Describe the characteristics of *PN* junction. Explain the meaning of (a) reverse bias, (b) potential hill, (c) reverse current.
4. Explain the difference between intrinsic and impurity conduction. Describe the manner in which the electrical conductivity of impurity in a semiconductor varies with temperature for different amounts of impurity contents.
5. Draw a typical current/voltage characteristics curve for a *p-n* junction diode and explain qualitatively its main features.
6. Describe the mechanism of rectification at a *p-n* junction. What factors determine the reverse current in such a rectifier ?
7. Explain what is meant by mobility of a charge carrier in a solid. Derive an expression for the conductivity of a semiconductor containing both free electrons and holes in terms of the concentrations *n* and *p* and the mobilities μ_n and μ_p.
8. Describe the static characteristics of transistor for (a) collector current versus collector voltage for a constant value of emitter current. (b) collector current versus collector voltage for a constant value of base current.
9. Explain the mechanism of current flow in a *PNP* and *NPN* transistors.
10. Give a classification of semiconductor devices. How grown junction diodes and transistors are manufactured ?
11. Describe in detail the manufacture of fused junction diodes and transistors.
12. Compare the performance of a transistor with a vacuum triode.
13. Describe briefly the principle of a field effect transistor and compare its performance with a junction transistor.

14. Write short notes on MOSFET.
15. Write short notes on (i) point contact diode and transistors, (ii) tunnel diode, (iii) field effect transistors and uni-junction transistor ?
16. What are the different parameters used for the analysis of a transistor.
17. Define (a) open circuit impedance (z) parameters (b) short circuit admittance (y) parameters and (c) hybrid (h) parameters of a four-terminal network. In what ways are the h-parameters useful than other parameters as far as transistors are concerned ?
18. What is meant by doping a semiconductor ? What basic properties must these doping elements possess to make a pure semiconductor a *P*-type or *N*-type ? Explain the physical principles involved.
19. Write short note on Gunn effect.
20. What are *DIAC* and *TRIAC* ? Discuss their characteristics.

7

ELECTRON EMISSION

The passage of electrons from a metal into a gas or vacuum is called electron emission. The mechanism of electron emission is classified into the following categories :

(1) Thermionic Emission. When energy applied to the conductor to emit electrons is in the form of heat energy, the process is called thermionic emission.

(2) Photoelectric Emission. When the energy applied to the metal is in the form of light energy, the process of electron emission is known as photoelectric emission.

(3) Secondary Emission. If a fast moving electron collides with the surface of the metal then the emission of electrons takes place. This is known as secondary emission.

(4) Field Emission. If the electrons are emitted from the surface of metal by maintaining a certain amount of electrostatic field between the two plates of a vacuum tube then emission is known as field emission.

For the study of emission process we shall first consider the structure of a crystal.

7·0. CRYSTAL STRUCTURE :

Whatever may be the method of application of energy to the crystal or metal, the electron in the crystal requires certain amount of energy to escape from the surface of the metal, known as potential barrier, and so it is necessary to study the structure of crystal.

A metal consists of a conglomeration of crystals. Every crystal consists of atoms which are held in position by mutual forces, forming a regular pattern. Each atom consists of a nucleus surrounded by a number of electrons which move in various orbits. The electrons in the outermost orbit are called valency electrons, on which chemical behaviour of the atom depends. In a metal, since the atoms are very densely packed in any crystal, the valency electrons move freely from one atom to another in a random way. It has been found that 10^{22} electrons/c.c. are free to move in this random manner. When an electric field is applied, the electrons move in

the definite direction depending upon the direction of field and thus constitute an electric current. Due to the heavy concentration of free electrons, average velocity of electrons is of the order of a few centimeters per second.

7·1. BOUND AND FREE ELECTRONS IN THE CRYSTAL LATTICE :

By Coulomb's law, an isolated nucleus will exert a force F on an electron distant r from nucleus equal to

$$E=\frac{Ze^2}{4\pi r^2\epsilon_0},$$

where ϵ_0 is the permittivity of free space and Z is atomic number of atom and potential energy

$$E=\int_{\infty}^{r} F\,dr=-\frac{Ze^2}{4\pi r\epsilon_0}.$$

In the diagrams 1 (a, b, c) the potential energy E is plotted against r with two or more nuclei being present.

The potential energy of the electron is largely affected by the presence of more nuclei. As shown in figures 1 (b, c) the force acting on any electron is the sum of the forces due to separate nuclei. The forces have opposite directions at any point in between two atoms and these two forces tend to cancel each other. The resultant force is small and so is the case with potential energy as shown in the figure 1 (c).

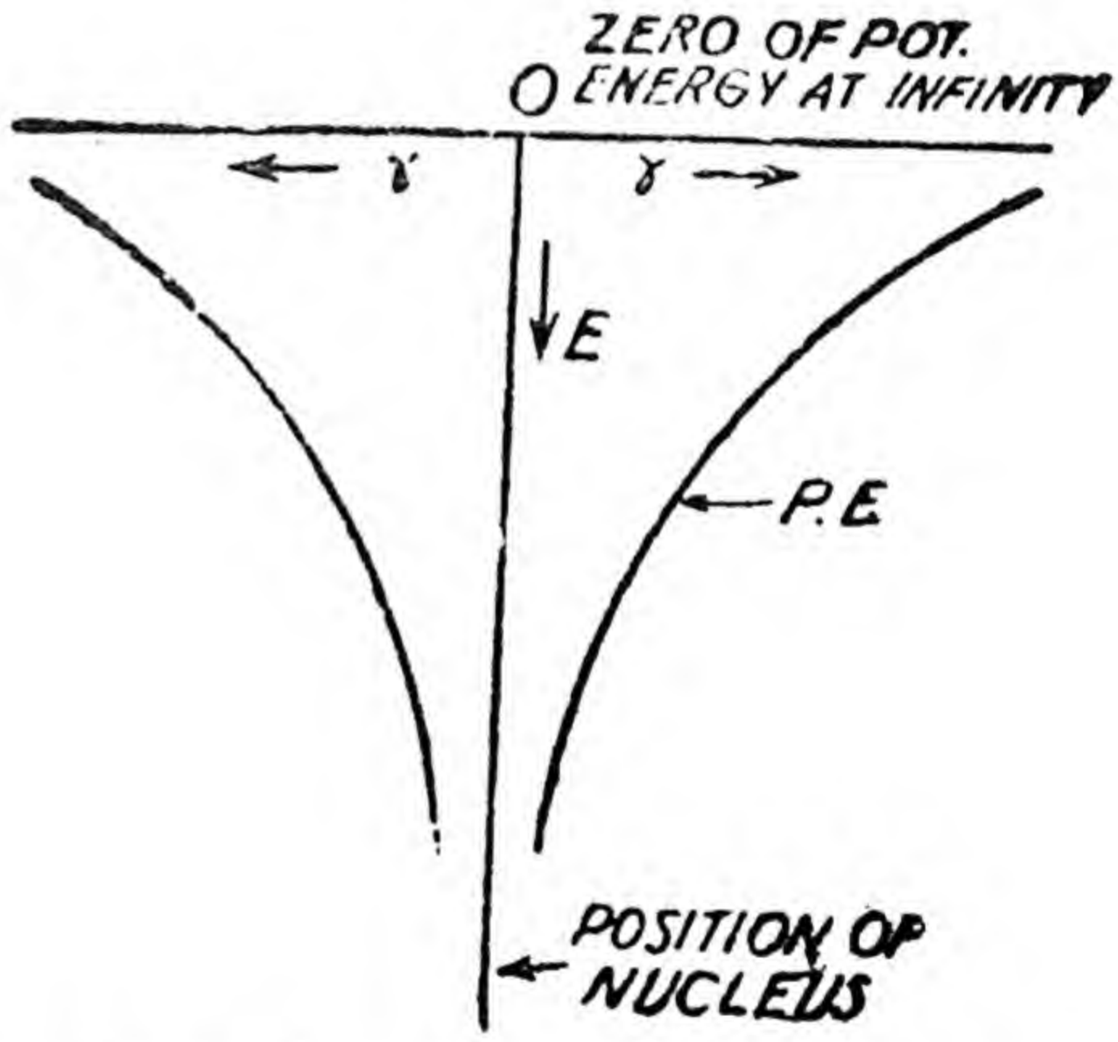


Fig. 1 (a) Variation of potential energy E with distance r for an isolated electron

Along with this potential energy, the electrons also possess kinetic energy. It is the tendency of an electron to occupy a position of minimum potential energy, *i.e.*, close to the nucleus, but only a certain number of electrons can exist in different shells. For example, the 'K' shell which is closest to nuclei can accommodate two electrons. After K shell, L shell comes which can accommodate eight electrons and in this way the other shells allow. The outermost shell is not often filled which consists of valence electrons. In between the atoms, highest potential energy levels are lowered due to combined effect of these several nuclei.

This allows some of the valence electrons to jump from their original nucleus. These electrons are called *free electrons*. The

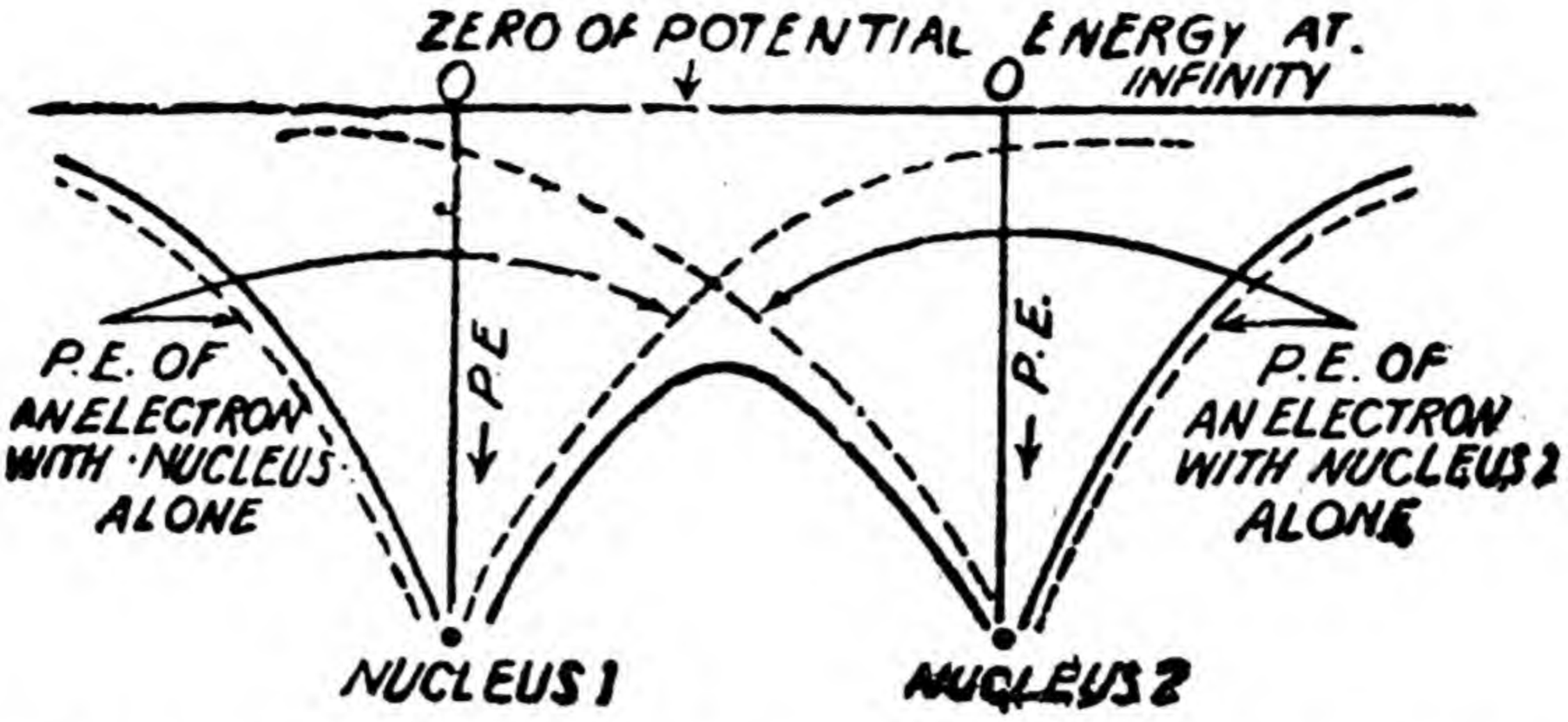


Fig. 1. (*b*) Potentia! energy of an electron along a line joining two nuclei.

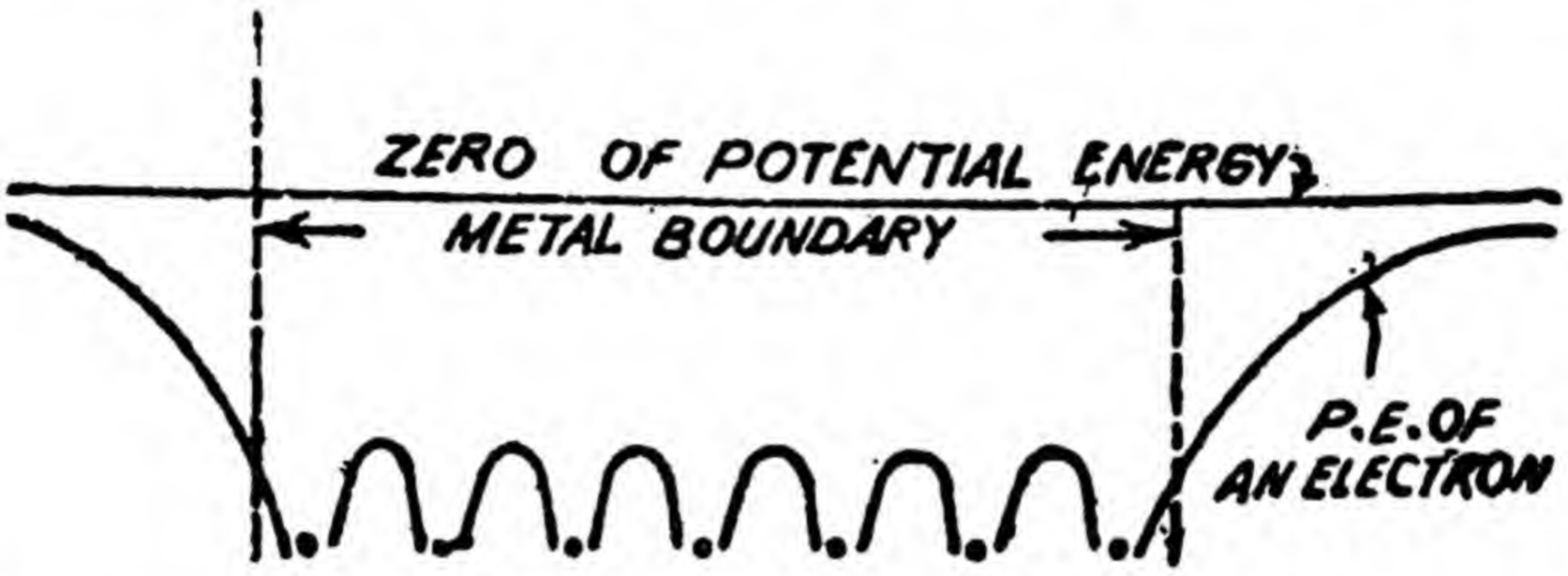


Fig. 1 (*c*) Potential energy of an electron along a line through a number of nuclei in a metal:

electrons having energy much less than the top of potential energy humps cannot leave their parent nucleus. Such electrons are called *bound electrons*.

7·2. SURFACE BARRIER :

The atoms in a metal are arranged in a particular geometrical form called space lattice structure. The valence electrons are free to move inside the metal but they are not allowed to leave the metal. When the electrons are detached from their parent atoms, the atoms acquire positive charges. The behaviour of electrons in two dimensions is shown in fig. 2. The removal of these electrons from the surface of the solid into the space is known as electron emission.

Let us consider a free electron in the position *C* or *B* in the interior of the metal. This electron is attracted by the surrounding positive ions and as the ions are symmetrically placed with respect to the electron, there will be no resultant force on the electron. This is true only for specific positions of electron and not for all positions. When this electron moves inside the metal, it is acted

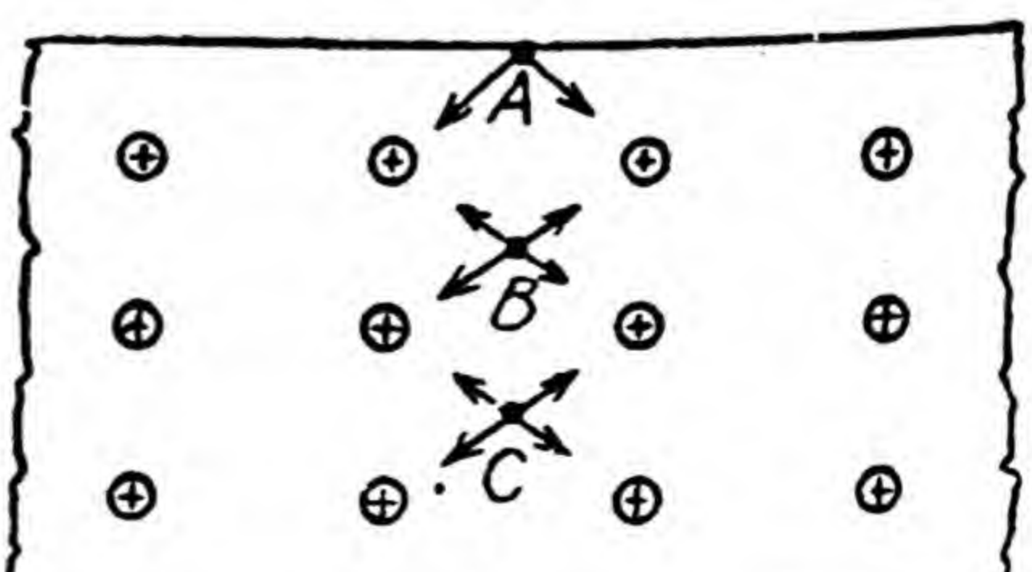


Fig. 2. Showing the arrangement of atoms of a metal in two dimensions.

upon by balanced and unbalanced forces at different instants, of course, the unbalancing forces are not too much. Now consider the electron *A* at the surface of the metal. This electron is attracted by downward forces and hence it can not go away from the surface. In this way we can say that as along as the electron is in the interior of the metal it moves freely from one position on the another but when it reaches near the surface, downward forces act upon it and it is prevented from crossing the surface. Thus under ordinary conditions, the electron can not pass through the surface or *the surface acts as a barrier.* When the kinetic energy of the electron is increased by external energy source, the electron crosses the barrier. *The barrier experiencea by an elec ·on at the surface of metal is known as surface barrier.*

7·2-1. POTENTIAL BARRIER DIAGRAM OF A METAL :

From the emission point of view, it can be assumed that the interior of the metal is an equipotential region *i.e.* with no potential barriers. This potential barrier level is regarded as zero. When an electron moves from the interior to exterior, first of all it is not being acted by any force *i e.* the barrier level is zero. When it reaches near the surface, it is acted upon by the downward force due to the nuclei of various atoms. The electron thus meets various barrier levels which correspond to a particular energy. If the energy of the electron is greater than the energy of a particular level, it will cross the level, and if the energy is less than the energy of a particular level, it will not cross it. The potential energy diagram is shown in fig. 3.

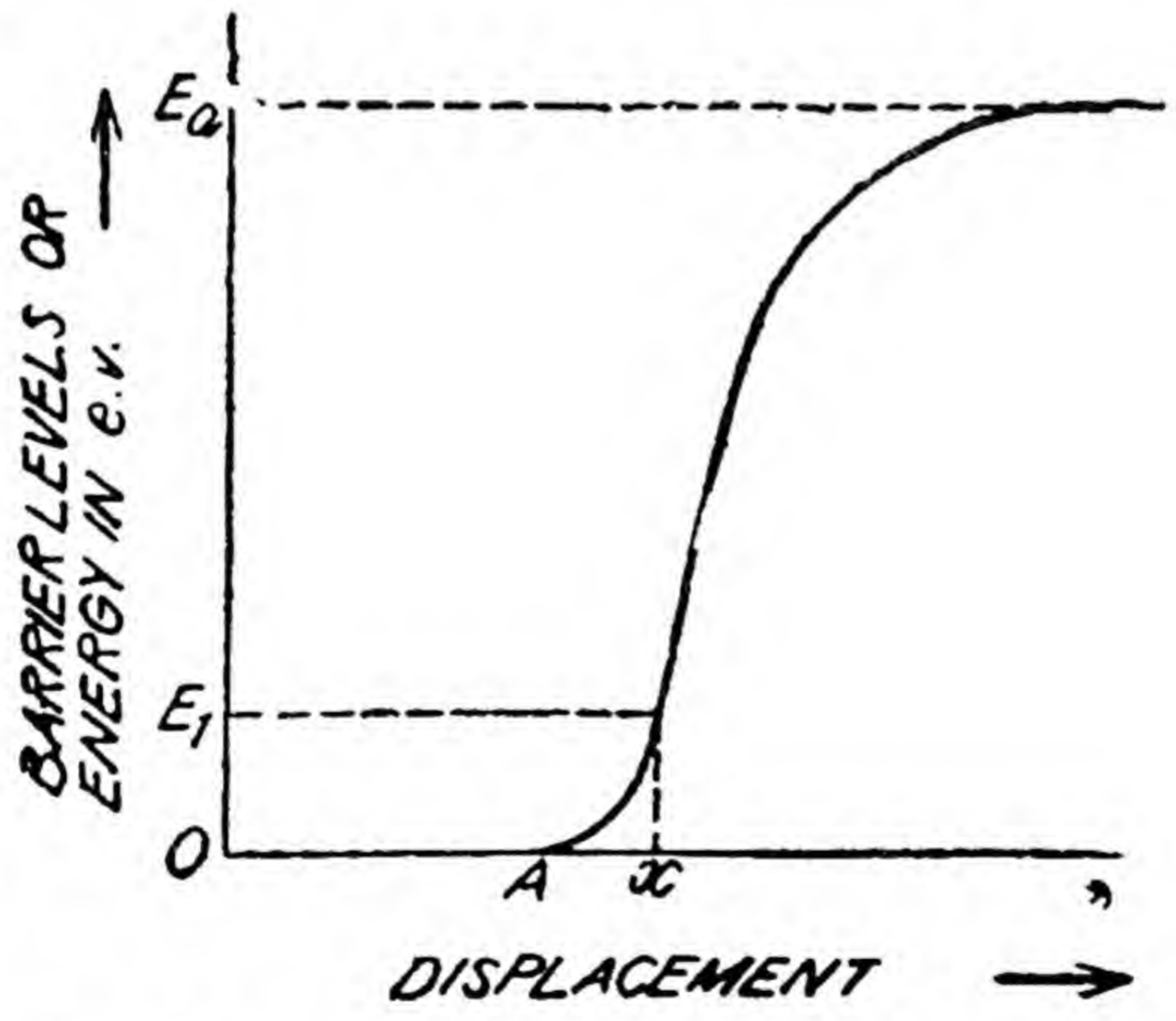


Fig. 3. The potential energy diagram.

It is observed from the figure that, (1) the distance *OA* corresponds to a region with no potential barrier, (2) if an electron having energy E_a is moving towards the surface, it will just reach the surface, and (3) if its energy is E_1 it will be confined within the metal. It moves a distance *OX*, measured from a point well within

the surface of the metal, (4) if the energy is greater than E_a, it will cross the surface and will appear as free electron outside the metal, (5) the surface energy barrier level E_a or simply barrier potential is different for different materials. This is measured in electron volts.

7·2-2. FERMI DIRAC ENERGY DISTRIBUTION CURVE :

To understand the thermionic emission of electrons from the metal it is necessary to have a knowledge of energies possessed by electrons in metals. The energies can be evaluated from the Fermi-Dirac energy distribution curve shown in fig. 4. The distribution is given as

$$dN = N(E)dE = \frac{8\pi\sqrt{(2m^3)}\sqrt{E}\,dE}{h^3\,[1+e^{(E-E_f)/kT)}}.$$

where E=kinetic energy of an electron,

E_f=Fermi level of energy,

$N(E)$=number of electrons per unit volume per unit change of energy.

m=mass of electron,

h=Planck constant,

T=absolute temperature of the electrons,

k=Boltzman's constant

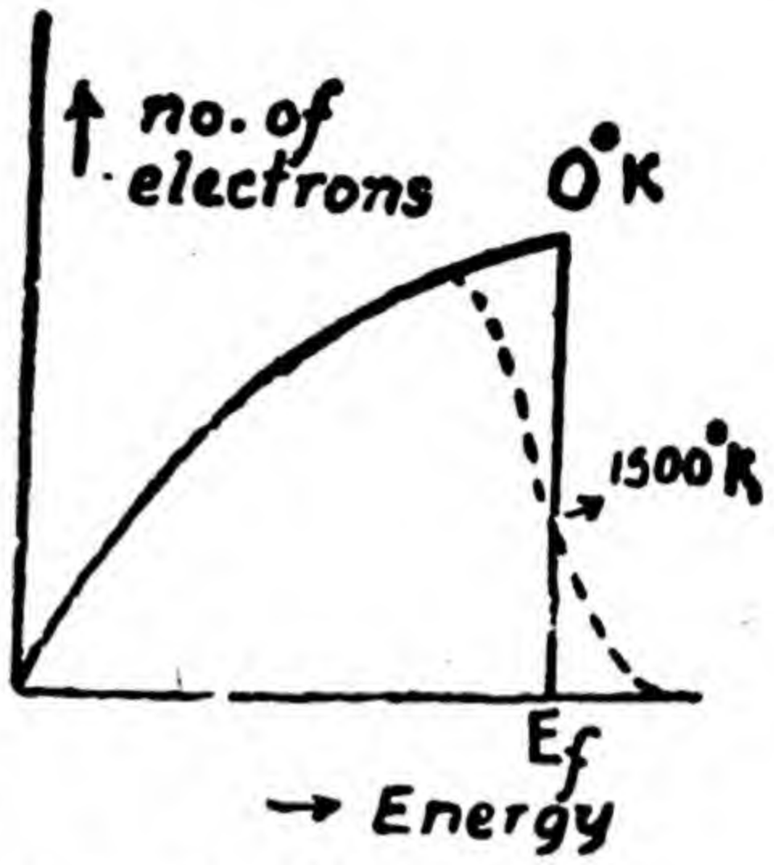


Fig 4. Distribution function for K E. of elements in a space of uniform potential.

When this is expressed in velocity range v and $v+dv$, it has the from

$$dN \cdot \frac{8\pi m^3 v^2 dv}{h^3\,[1-e^{(E-E_f)/kT}]}$$

$$= \frac{2(m/h)^3\,4\pi v^2\,dv}{1+e^{(E-E_f)/kT}}$$

The total area under the curve is equal to the total number of free electrons in the unit volume of the metal.

At $T=0$, E_f called the Fermi energy is greatest energy possessed by the electron and all states below E_e are filled, while all above E_f are vacant. The value of E_f is given by

$$E_f = \frac{h^2}{8m}\left(\frac{3n}{\pi}\right)^{2/3} \text{ Joules,}$$

where n is equal to the number of electrons per unit volume.

The Fermi level varies with iemperature but the variation is very small and Fermi equation holds good over a wide range of temperatures below the melting point of metals.

The quantity $N(E)\,dE$ represents the number of electrons in a unit volume having kinetic energies between the limits E and $E+dE$. The function $N(E)$ represents the number of electrons per unit volume per unit change of energy. Again the number of electrons in kinetic energy range ΔE is equal to $N(E)\times\Delta E$.

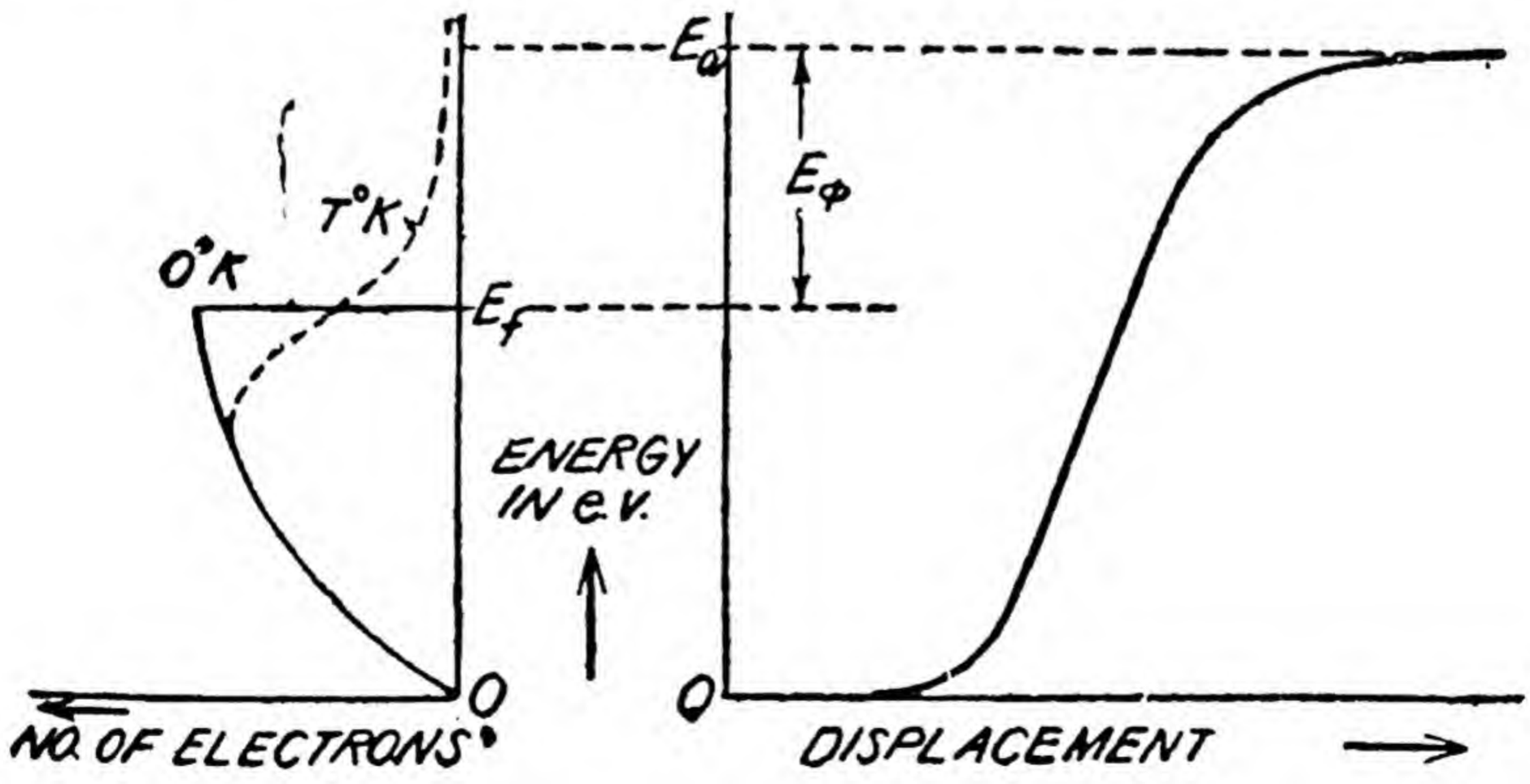
Thus according to Fermi-Dirac statistics, the electrons moving in a field free space have kinetic energies because of the velocities associated with their temperature and so distribution of energy among them. Even at absolute zero, the energy of the electron is not zero. At absolute zero, the maximum energy is E_f. At higher temperature, more heat energy is stored in the electrons and their average kinetic energy is increased. Some electrons attain energies even greater than E_f as shown in the curve at 1500°K.

7·2-3. ELECTRON EMISSION :

To understand the process of electron emission, the potential barrier curve and Fermi-Dirac distribution curve are plotted as shown in fig. 5 (*a*).

It is observed from fig 5 (*a*) that the Fermi level is less than the surface barrier E_a. Thus at 0°K, no electron can be emitted from the surface. The difference $(E_a - E_f)$ is known as work function. For the electron emission atleast this amount of energy $(E_\phi = E_a - E_f)$ must be given by some external source. This can be done by heating the metal.

When the metal is heated, the electrons that can acquire thermal energy are those which have the energies near the Fermi-level because these electrons have unfilled energy levels immediately above them. The energy distribution curve is modified as shown in fig. 5 *a* (1) as the temperature is increased from 0 K to T K. The curve is asymptotic to the energy axis so that there is a probability of some electrons having energies greater than the energy of potential barrier E_a, as the temperature is raised from 0 K. Appreciable emission occurs only at high temperatures. The emission of free electrons depends not only upon energy above barrier level but also on the direction of its velocity. Thus an electron that is not directed towards the surface will never be emitted whatever may be the energy.



(1) Fermi-Dirac curve. (2) Potential barrier curve

Fig. 5 (a)

The variation of distribution function $n_x(E_x)$ with E_x is sketched in fig. 5 (*b*) for temperatures of zero and 1500 degrees Kelvin.

At zero temperature, curve is straight line showing finite number of electrons at zero energy and zero electrons at the Fermi level, but for higher temperatare (T=1000K) the curve extends its toe and thus half of the electrons X-directed energies greater than the potential barrier energy will be emitted.

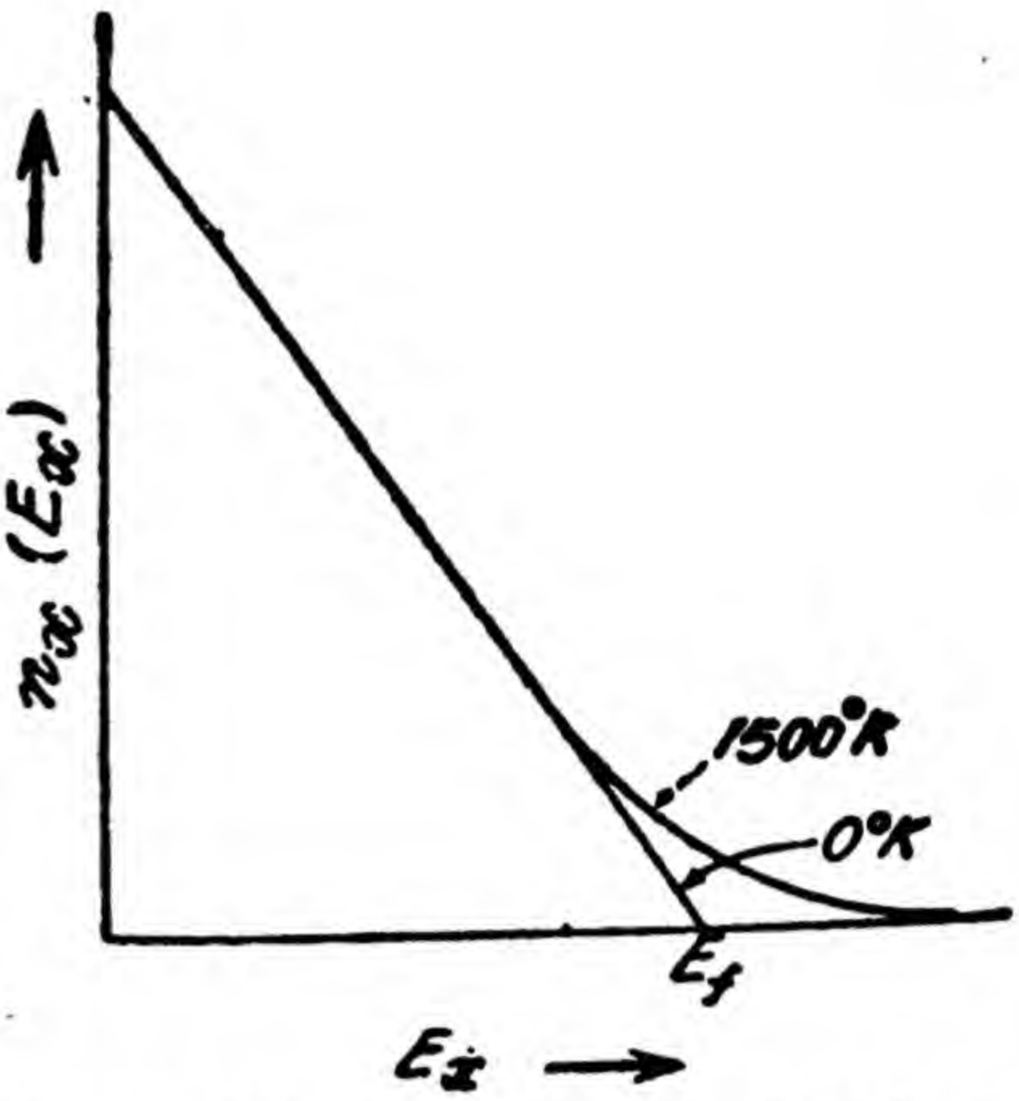


Fig. 5. (b). Distribution function for x-associated energy.

Now to have clear conception of the escape of the electrons from the metal, we shall plot together the fig. 5 (*b*) and fig. 1 (*c*), the latter figure gives the space distribution of the potential energy of electron in a metal or near the surface of a metal. To make this combination, a reference level of energy in the two concepts must be found. We know that at absolute zero all the energy levels in the band of valence electrons below a certain level are occupied while all higher levels are unoccupied which is the same as discussed above in Fermi-Dirac distribution that all states below E_f are filled while all states above E_f are vacant. The Fermi level of Fermi-Dirac distribution coincides with highest occupied level in band of allowed energies occupied by valence electrons and accordingly, we align the diagrams 5 (*b*) and 1 (*c*) as shown in fig. 5 (*c*).

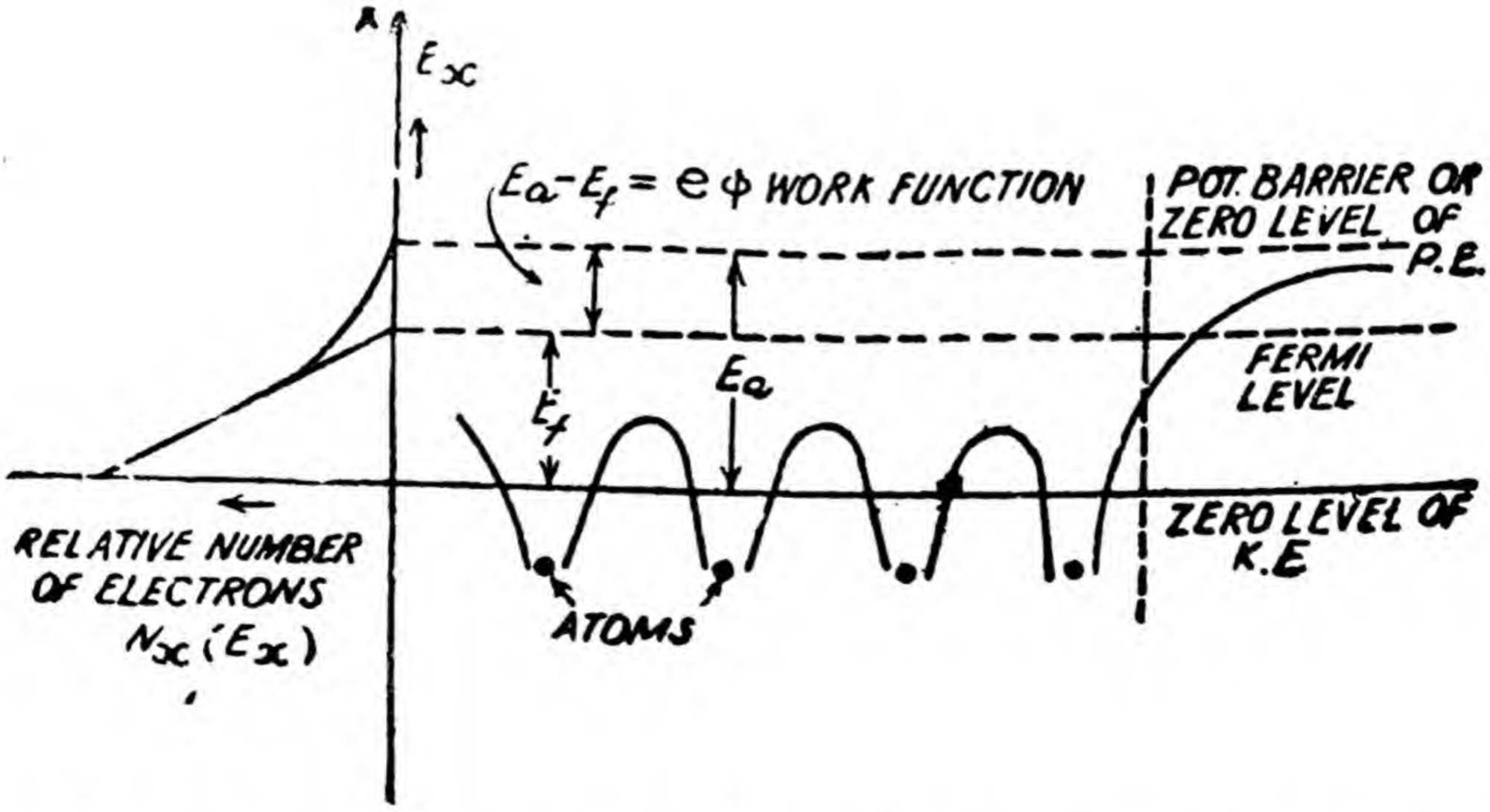


Fig. 5 (c). Energy level of P. E. and K. E , illustrating work function.

The energy difference between zero level of K.E. and potential barrier of the metal is called the surface energy E_a. The energy between zero level of K.E. and Fermi level is known as energy E_f. In order that an electron may escape from metal, it must have energy equal to E_a or greater than E_a. The difference of energy $(E_a - E_f)$ is called the *work function of metal*. It is a function of temperature although to a very small extent.

Therefore, an electron can be emitted from the metal only when the *energy greater than or equal to the work function is applied to* the electron and this energy may be supplied in the form of heat energy, light energy, electrostatic energy giving different types of emission.

Now we shall study different types of emission.

7·3. THERMIONIC EMISSION :

As discussed already, in the process of thermionic emission, the kinetic energy of electron is increased directly by raising the temperature. At sufficiently high temperatures an appreciable number of electrons will have the kinetic energy required to escape through the surface of the material. This results in thermionic emission of electrons.

Richardson, in 1901, derived a relation in number of electrons (hence current) emitted per unit area of metal surface and the temperature in degrees Kelvin of the emitting material. Later on the relation was corrected by Dushman.

Richardson, in his early attempt derived this relation on thermodynamic basis. Later on, it was derived on the basis of Fermi-Dirac statistics.

7·4. RICHARDSON EQUATION :

Thermodynamic Derivation : The emission of electrons from the surface of the metal resembles with the process of evaporation from the surface of liquid and the energy, that electron must give up in escaping, corresponding to the latent heat of evaporisation of the liquid.

The electrons in the metal, being densely packed, constitute the condensed phase, while electrons in space charge constitute the vapour phase. We assume that condensed and vapour phase are in dynamical equilibrium.

Let p be the pressure of the vapour phase when equilibrium exists between two phases, and let V and U be volume of one gram molecule of electrons in free space (vapour phase) and metal itself respectively. If W be the energy absorbed during the evaporation of N electrons (where N represents the Avogadro's number) then Clausius Clapeyron equation can be written as

$$W = T \frac{dp}{dT} (V - U).$$

where T is absolute temperature in K of the emitter.

Since $$V >> U.$$

$$\therefore \quad W = TV \frac{dp}{dT} \qquad \ldots(1)$$

Energy W is made up of two parts :

(*i*) the work function energy required to liberate electrons from metal surface which will be $N\phi$ if ϕ is work function.

(*ii*) the work done against the pressure p of electron cloud which will be pV.

First in his early derivation of $T^{1/2}$ relation, he assumed that the electrons in vapour phase (*i.e.* in free space cloud) behave as perfect gas particles; that is, he ignored the repulsive force between electrons in cloud due to their negative charges and purely applied the concepts of kinetic theory.

Therefore $pV = RT$, *i.e.* perfect gas equation, where R is gas constant per gram molecule, was applied in the derivation.

Energy W can be written as

$$W = N\phi + pV$$
$$= N\phi + RT = N\phi + NkT, \qquad \ldots(2)$$

where k is Boltzmann constant,

Thus equating (1) and (2), we get

$$TV \frac{dp}{dT} = N(\phi + kT).$$

Since $$pV = NkT. \quad \therefore \quad V = \frac{NkT}{p}$$

Putting in above equation, we get

$$\frac{NkT^2}{p} \frac{dp}{dT} = N\ (\phi + kT).$$

$$\therefore \quad k. \frac{dp}{dT} = \frac{\phi + kT}{T^2}\ dT.$$

Integrating the above equation, we get

$$k.\ \log_e p = \int \frac{\phi\ dT}{T^2} + \int k. \frac{dT}{T} + \text{constant},$$

$$\log_e p = \int \frac{\phi}{kT^2} .dT + \log_e T + \text{constant}.$$

$$\therefore \quad p = A_1 . T . e^{\int \frac{\phi}{kT^2}\ dT} \qquad \ldots(3)$$

where A_1 is a constant.

If in electron gas, the number of electrons per unit volume be n, then $p = nkT$ and thus eq. (3) becomes

$$n = \frac{A_1}{k}\ e^{\int \frac{\phi}{kT^2}\ dT} = A_2\ e^{\int \frac{\phi}{kT^2}\ dT} \qquad \ldots(4)$$

where A_2 is another constant.

If m_0 be the mass of an electron, then the mean square velocity v^2 of the electrons within the cloud is given by

$$\tfrac{1}{2}m_0v^2=\tfrac{3}{2}kT,$$

$$v^2=\frac{3kT}{m_0}$$

From Kundsen's cosine law, the number of electrons crossing unit area in the cloud per second, n_0, is given by

$$n_0=\frac{nv}{\sqrt{(6\pi)}}.$$

Putting the value of v,

$$n_0=\frac{n\sqrt{(k)}}{\sqrt{(6\pi m_0)}}.T^{1/2}=n.CT^{1/2},$$

where C is a constant.

Putting the value of n from equation (4), we get

$$n_0=A_2\,e^{\int \frac{\phi}{kT^2}\,dT}.\,CT^{1/2}.$$

Assuming ϕ to be independent of temperature, we can write after integration

$$n_0=AT^{1/2}e^{-\phi/kT}, \qquad \ldots(5)$$

where $A=A_2C=$constant.

The current density J of electron emission will be

$$\begin{aligned}J&=n_0e\\&=AeT^{1/2}e^{-\phi/kT}\\&=A_0T^{1/2}e^{-\phi/kT},\end{aligned} \qquad \ldots(6)$$

where $A_0=Ae$, a constant.

But when this relation was experimentally varified, then it gave the values A_0 and ϕ very different from experimental values and thus the relation was descarded.

Later on, Richardson and Dushman showed that ϕ is a function of temperature given by

$$\phi=\phi_0+\tfrac{3}{2}kT, \qquad \ldots(7)$$

where ϕ_0 is work function of metal at 0°K.

From (7) it is clear that they have assumed the non-existence of zero point energy which is contrary to modern quantum consideration and for this reason derivation is not favoured, although it gives an accurate relationship which is derived below :

Since $$\phi=\phi_0+\tfrac{3}{2}kT,$$

we have $$\int \frac{\phi}{kT^2}\,dT=\int \frac{\phi_0+\frac{3}{2}kT}{kT^2}.dT=\phi_0\int\frac{dT}{kT^2}+\frac{3}{2}\int\frac{dT}{T}$$

$$=-\frac{\phi_0}{kT}+\frac{3}{2}\log_e T.$$

Putting the exponential term value in equation (6), we get

$$J = A_0 T^{1/2} \text{ exp.} \left[-(\phi_0/kT) + \frac{3}{2} \log_e T \right]$$

$$= A_0 T^{1/2} \text{ exp.}(-\phi_0/kT).T^{3/2}$$

$$= A_0 T^2 \text{ exp.}(-\phi_0/kT).$$

When this expression was put to experimental test then it gave reliable values of work function. Dushman has shown that A_0 is a universal constant having a value of 120·4 amperes per square centimeter per degree Kelvin^2. The discrepancies between experimental and theoretical values of A_0 was explained by Dushman on the basis of internal reflection of electrons from the surface of metal.

Drawbacks of above method of derivation :

(i) In the derivation it is assumed that the metal contains 'free' electrons which behave like the molecules of perfect gas, that is, the different forces between the electrons have been ignored.

(ii) The classical theory fails to differentiate in forms of formulae (6) and (8) because ;

(*a*) it is not easy to measure the high temperature accurately at which electron emission takes place ;

(b) as the exponential term has a greater influence than other terms depending upon T, therefore a suitable adjustment of constants will fit both the formulae (6) and (8) to the experimental results ;

(*c*) if we ignore the assumption made *i.e.*, $\phi = \phi_0 + \frac{3}{2}kT$, then T^2 formula will result in $T^{1/2}$ formula. The assumption made is, however, not suitable.

Thus the thermodynamical derivation is not adopted these days. Thermionic emission is best explained by Fermi-Dirac statistics in which the concepts of Fermi energy and Fermi band explain the condition of emission.

7·5. DERIVATION OF RICHARDSON EQUATION BY FERMI DIRAC STATISTICS :

In the discussion of electron emission from metal, it was pointed out that only those free electrons will escape from the metal which have got energy equal to E_a or greater than E_a, where E_a denotes the surface energy.

According to the Fermi-Dirac distribution, the number of electrons per unit volume lying between momentum space p and $(p+dp)$ is

$$dN = \frac{8\pi p^2 \, dp}{h^3 [e^{(E-E_f)/kT} + 1]} \qquad \ldots(1)$$

As $p = mv$, $dp = m \, dv$, we get

$$dN = \frac{2\left(\frac{m}{h}\right)^3 4\pi v^2 \, dv}{e^{(E-E_f)/kT} + 1}$$

If v_x, v_y, v_z be the velocity components of the electron in x, y and z directions, then the phase-space volume between the velocity range v_x and v_x+dv_x v_y and v_y+dv_y, v_z and v_z+dv_z will be

$$4\pi v^2\, dv = dv_x\, dv_y\, dv_z$$

Therefore, $$dN = \frac{2\left(\frac{m}{h}\right)^3 dv_x\, dv_y\, dv_z}{e^{(E-E_f)/kT}+1} \qquad \ldots(2)$$

The value of $(E_a - E_f)$ is of the order of $20kT$, while kT at 300°K corresponds to an energy of about $0{\cdot}03eV$. Now in $(E-E_f)$ the least value of E required for emission of electrons is E_a; the quantity $e^{(E-E_f)/kT}$ is much greater than 1. Therefore in the denominator, we neglect 1, *i.e.*,

$$dN = \frac{2\left(\frac{m}{h}\right)^3 dv_x\, dv_y\, dv_z}{e^{(E-E_f)/kT}}$$

$$= 2\left(\frac{m}{h}\right)^3 e^{(E_f-E)/kT}\, dv_x\, dv_y\, dv_z$$

$$= 2\left(\frac{m}{h}\right)^3 e^{E_f/kT}.e^{-\frac{1}{2}m(v_x^2+v_y^2+v_z^2)/kT}\, dv_x\, dv_y\, dv_z$$

$$= 2\left(\frac{m}{h}\right)^3 .e^{E_f/kT}\, e^{-\frac{1}{2}mv_x^2/kT}.e^{-\frac{1}{2}mv_y^2/kT}$$

$$.e^{-\frac{1}{2}mv_z^2/kT}\, dv_x\, dv_y\, dv_z.$$

We consider that the surface of the metal is a plane surface perpendicular to the direction of x-axis.

Now the number of elcctrons with a velocity component between v_x and v_x+dv_x is given by

$$dN_x = N(v_x)\, dv_x$$

$$= 2\left(\frac{m}{h}\right)^3 e^{E_f/kT}\, e^{-\frac{1}{2}mv_x^2/kT}\, dv_x$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}m(v_y^2+v_z^2)}\, dv_y\, dv_z.$$

But only those electrons which are travelling in x-direction will come out in yz plane (metal surface) and have got energy greater than $(E_a - E_f)$, the work function $e\phi$. For finding out the number of electrons arriving at the metal surface per unit area per second (say n_x), the above must be multiplied by v_x; as in unit time there impinge on the surface all the electrons with the component v_x which were contained in the layer of breadth v_x infront of the metal surface. Therefore to calculate n_x, we multiply by v_x and integrate between the limits $\sqrt{\left(\frac{2E_a}{m}\right)}$ to ∞.

Thus
$$n_x=\int_{\sqrt{(2E_a/m)}}^{\infty} N(v_x)\, v_x\, dv_x$$
$$n_x=2\left(\frac{m}{h}\right)^3 e^{E_f/kT}\int_{\sqrt{(2E_a/m)}}^{\infty} v_x.e.^{-\frac{1}{2}mv_x^2/kT}\, dv_x$$
$$\times\int_{-\infty}^{\infty} e^{-\frac{1}{2}mv_y^2/kT}\, dv_y \int_{-\infty}^{\infty} e^{-\frac{1}{2}mv_z^2/kT}\, dv_z.$$

Solving this integral, we get

$$\dagger n_x=\frac{4\pi mk^2}{h^3}.T^2.e^{-(E_a-E_f)/kT} \qquad \ldots(3)$$

Therefore saturation emission current density per unit area

$$J_S=e.n_x$$
$$=\frac{4\pi mek^2}{h^3}.T^2e^{-(E_a-E_f)/kT} \qquad \ldots(4)$$

Putting $A=\frac{4\pi mek^2}{h^3}$ **and** $e\phi=(E_a-E_f)$,

ϕ=work function$=\frac{(E_a-E_f)}{e}$,

$$J_S=AT^2\, e^{-e\phi/kT}. \qquad \ldots(5)$$

Equation (4) can also be written as

$$J_S=AT^2\, e^{-b/T} \text{ where } b=\frac{(E_a-E_f)}{k}$$

The equation (4) is known as Richardson-Dushman equation. Dushman has shown that A is constant for different metals and theoretically substituting the values of m, k, e, h, we get $A=120{\cdot}4$ amperes/sq. cm./degree Kelvin^2. The value of work function varies from metal to metal.

7·6. EXPERIMENTAL VERIFICATION :

The value of constants can be calculated easily by putting the equation in the form

$$\frac{J_s}{T^2}=Ae^{-e\phi/kT},$$
$$\log_e\frac{J_S}{T^2}=\log_e A-\frac{e\phi}{kT}=\log_e A-\frac{b}{T}.$$

---

†**Hint :** $\int_{-\infty}^{+\infty} e^{-x^2}\, dx=\sqrt{\pi}$

$$\therefore \int_{-\infty}^{+\infty} e^{-\frac{1}{2}mv_y^2/kT}\, dv_y=\frac{\sqrt{\pi}\sqrt{(2kT)}}{\sqrt{m}}$$

To evaluate

$\int_{\sqrt{(2E_a/m)}}^{\infty} v_x e^{-\frac{1}{2}mv_x^2/kT}\, dv_x$, let $\frac{1}{2}mv_x^2/kT=t$

or $\frac{mv_x dv_x}{kT}=dt$ so that $\frac{kT}{m}\int_{E_a/kT}^{\infty} e^{-t}\, dt=\frac{kT}{m}\, e^{-E_a/kT}$

If we plot a graph in $\frac{1}{T}$ and $\log_e \frac{J_S}{T^2}$, a straight line is obtained. The intercept made on y axis gives the value of $\log_e A$. The slope of line with x-axis gives the value of b. From these data, values of A and ϕ can be calculated by taking different metals for different emitters.

We shall tabulate here the representative values of thermionic emission constants for some pure metals :

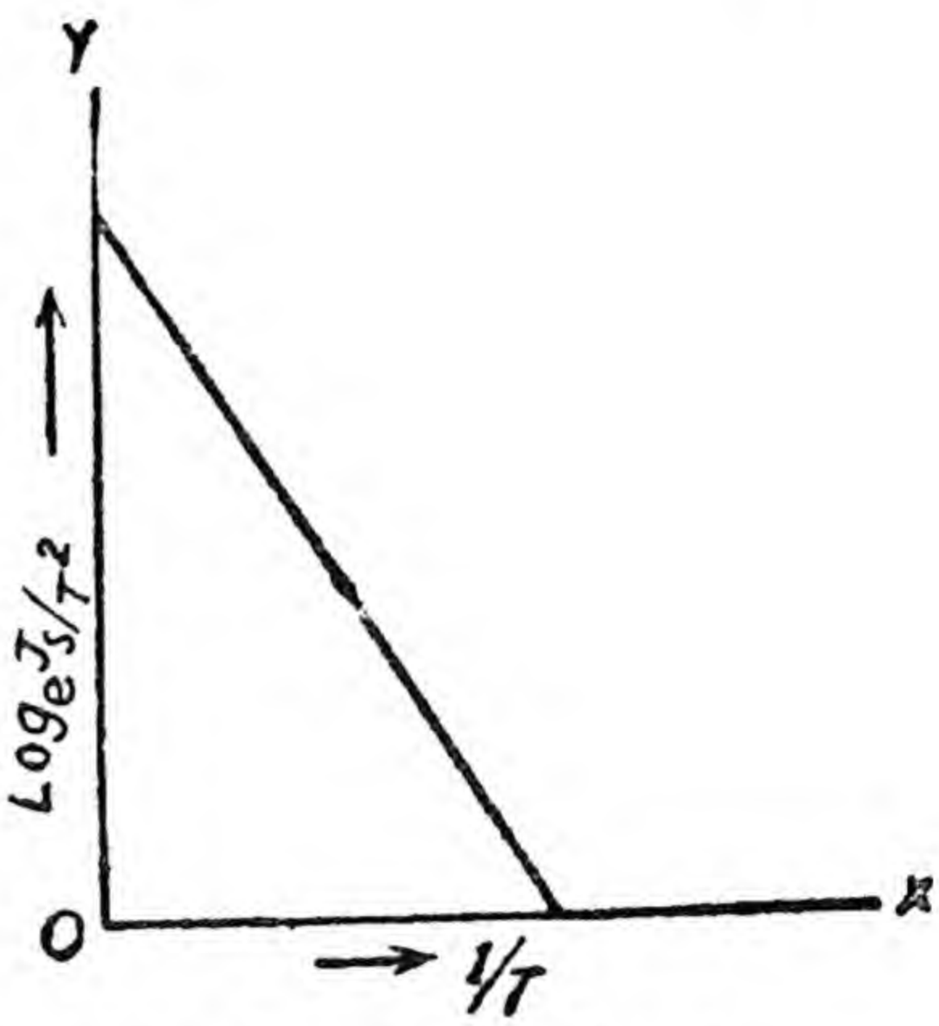


Fig. 6. Graph in $1/T$ and $\log_e J_S/T$.

Metal	A in amp/cm²/degree K^2	b in K	Melting point in K	Work function ϕ in eV
Calcium	60 2	26,000	1,083	2 24
Carbon	€0·2	46,000	3,773	5·34
Cesium	60·2	21,000	299	1·81
Platinum	60·2	59,000	2,028	5·32

7·7. EXPLANATION FOR THE DISCREPANCIES IN THEORETICAL AND EXPERIMENTAL RESULTS :

From the experimental data, we find that the values of A and ϕ do not tally with the theoretical results. This deviation from the experimental values was explained by Dushman on the basis of quantum mechanics as applied to the crystal structure. They are:

(*i*) In above discussion of thermionic emission we have taken the value of transmission coefficient equal to unity but according to quantum mechanics even an electron has got energy greater than the height of potential barrier; there is some probability of reflection inside the metal, *i.e.* the transmission coefficient cannot be equal to one. Dushman on this basis modified the formula as

$$J_S = A(1-R)T^2 e^{-e\phi/kT}$$
$$= A_r T^2 e^{-e\phi/kT},$$

where R is reflection coefficient. The value of R was determined experimentally and found to be $\frac{1}{2}$.

(*ii*) The variation in the value of work function from experimentally observed value is due to the fact that in the derivation of the formula we have taken the Sommerfeld structure of the crystal where it has been assumed that electron in crystal moves in a constant potential field. But actually the potential field inside the crystal structure is periodic. This assumption accounts for the discrepancy in the values of work function.

7·8. PRACTICAL EMITTERS :

They are of the two types :

(*A*) directly heated emitters, and

(*B*) indirectly heated emitters.

We shall first discuss the advantages and disadvantages of both the types as follows :

Disadvantages of the directly heated emitters : When alternating current is used to heat the emitter then the following defects arise :

(*i*) When alternating current flows in the filament, then an electrostatic field will be produced around the filament which affects the emission of electrons depending upon the periodicity of altrenating current and this results in alternating component of plate current.

(*ii*) Due to direct heating, the surface of filament is no more equipotential. The unequal distribution of potential also affects the amount of electron emission from different portions of the filament and thus it also produces a.c. component of plate current *i.e.* hum is produced.

Due to above two drawbacks, directly heated emitters are not preferred in receiving sets.

The advantages of indirectly heated emitters :

(*i*) Higher emission efficiency is obtained by the use of oxide-coated emitters. The emission efficiency is defined as the ratio of emission current to the power consumed in heating.

(*ii*) Hum, due to supply frequency component of plate current, is very much reduced.

(*iii*) Because heater element is insulated from the emitter, therefore a single a.c. source can be used to heat different cathodes at different potentials.

(A) Directly Heated Emitters :

(i) Tungsten Emitters : For pure tungsten emitter,

$A = 60{\cdot}2 \times 10^4$ amperes/sq. meter/K^2,

$k = 52400K$.

Emission efficiency is 200 mA./watt at $2400K$ and 4·2 mA/watt at $200K$.

The emission efficiency of pure tunsgten emitter is small as compared with oxide-coated and thoriated tungsten emitter or in otherwords tungsten has relatively high work function.

The main advantage of having pure tungsten as emitter is that it can be separated at a sufficiently high temperature between $2600K$ and $2800K$.

(ii) Thoriated Tungsten : A thoriated tungsten filament is prepared by adding a small amount of thorium oxide to the tung-

sten. The very thin layer of thorium oxide reduces the work function of the tungsten filament.

For the activated thoriated tungsten emitter, the value of A varies from 3.10^4 to 59.10^4 amperes/sq. meter/K^2 depending upon the electric field between anode and cathode. Value of b varies between 30500 to 36500K.

This can be operated at about 1800°K. In order to reduce the rate of evaporation of thorium layer from the filament the emitter is carbonized.

The practical emission efficiency is of the order of 100 milli-amperes/watt.

(B) Indirectly Heated Oxide Coated Emitters : Oxide-coated emitters are about twenty times as efficient as tungsten. They provide a high emission current even at low temperatures.

Such emitters are generally made by a coating of barrium or strontium oxide on a base or core material like pure nickel.

For oxide coated emitters the value of constant are

$A = 0{\cdot}01 \times 10^1$ amp./square meter/K^2,

$b = 1,16,000$ degree Kelvin,

work function $\phi = 1$ volt.

The usual operating temperature is 1000 degree Kelvin. At higher temperatures than 1000K, the life of oxide coated cathode is considerably reduced. The emission efficiency is of the order of 100·200 milliamperes per watt.

7·9· COMPARATIVE STUDY :

A comparative study of above emitters can be had from the following table :

Types of Cathode	Application	A in amp./sq. cm./K^2	b in K	Maximum steady anode potential (volts)	Anode Dissi-pation (Watts)
Oxide Coated	Radio-receiving small amplify-ing, transmitting tubes; small high vacuum rectifiers, gas filled recti-fiers.	$0{\cdot}01 \times 10^4$	116,000	200	300
Thoriated tungsten	Medium amplify-ing and transmit-ting valves.	3×10^4	30,500	5,000	1,000
Pure tungsten	Large transmit-ing valves; high vacuum rectifiers.	$60{\cdot}2 \times 10^4$	52,400	20,000	100,000

7·10. PHOTOELECTRIC EMISSION :

The process in which the energy required to emit an electron from the metal is supplied by means of light is known as photo-electric emission.

According to quantum theory, one quantum of light, called a 'photon', has energy equal to $h\nu$ where h is Planck's constant and ν is frequency of light radiation. When a photon strikes the metal surface, then some portion of energy $h\nu$ is supplied to the electron to cross over the barrier.

There are certain laws of photo-electric emission. For photo-electric emission, it is essential that the frequency of impinging light must exceed the 'threshold frequency' which depends upon the work function ϕ of the surface or, in other words,

$$h\nu \geqslant e\phi$$

or $$\nu_t = \frac{e\phi}{h}$$ where ν_t is threshold frequency.

The amount of photo-electric emission depends upon the intensity of light radiation falling on the surface. Alkali metals have low work functions and hence are used as common photo-tube emitters as discussed in chapter 5.

7·11. SECONDARY EMISSION :

Secondary emission takes place when high velocity electrons strike the metal surface. In this impact, they deliver sufficient energy to some of the electrons within the metal to enable them to cross over the barrier

The number of secondary electrons emitted per primary electron is called secondary emission ratio. This depends upon the following factors :

(*i*) material of the target,
(*ii*) energy of incident particle,
(*iii*) type of incident particle, and
(*iv*) the angle at which the target is struck.

The secondary emission ratio increases with velocity of incident electron and hence with the energy but it falls when kinetic energy is too large.

Experimental study of secondary emission :

When a moving particle strikes a solid surface, it enables the electron in the solid to escape through the potential energy barrier at the surface. This process is known as *secondary emission*. The experimental arrangement for the study of secondary emission is shown in figure 6 (*a*).

A beam of primary electrons having energy E_1 electron volts is directed from electron gun towards the target of the material to be studied for secondary emisson. The secondary electrons emitted

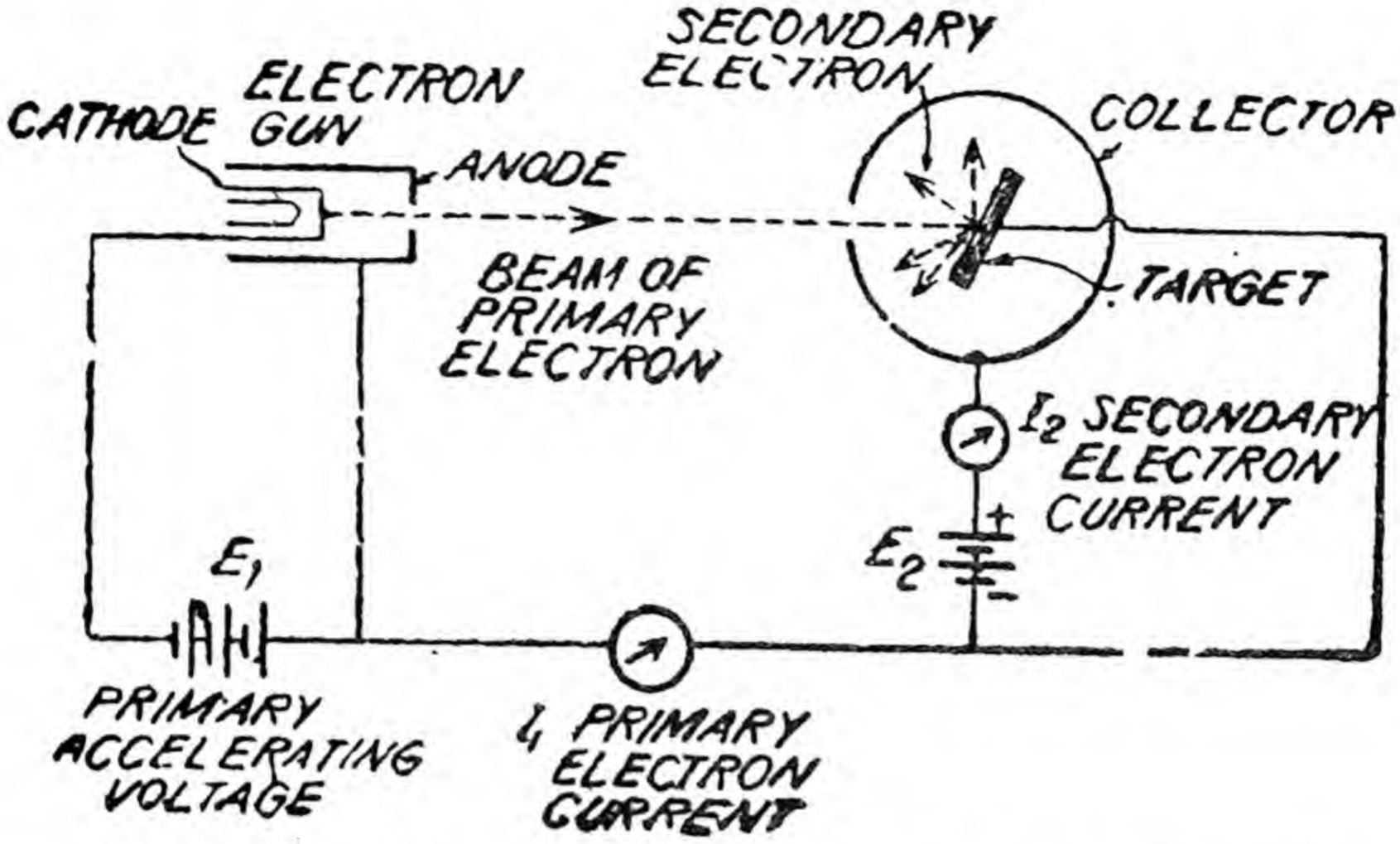


Fig. 6. (a) Measurement of secondary electron emission.

from the target are drawn to the collector by a voltage E_2. The secondary electrons produce a current I_2. I_1 is the the current due to the primary electrons. The secondary emission ratio is defined as

$$\frac{I_2}{I_1} = \frac{\text{No. of secondary electrons}}{\text{No. of primary electrons}}$$

Thus the number of secondary electrons emitted per primary electron is defined as secondary emission ratio. This ratio depends upon the target material, and its surface conditions and the type of primary particle, its energy, and the angle of incidence at the target.

Using different metals, the experiment is repeated and a graph is drawn between emission ratio and energy of primary electrons fig. 6 (*b*). From the figure it is observed that for pure metallic

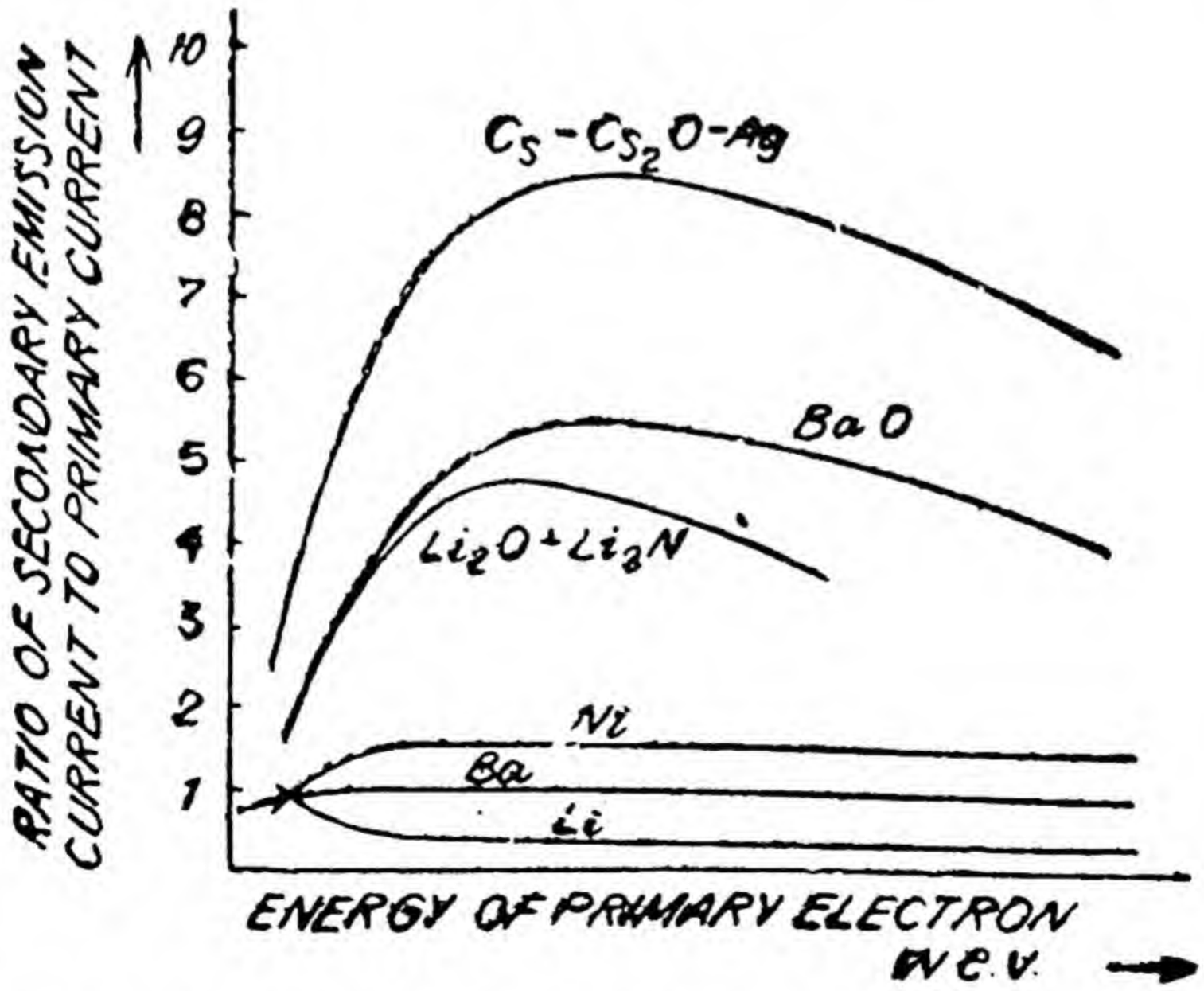


Fig. 6 (b) Secondary emission ratio as a function of energy of the primary electrons for various surfaces.

surfaces, the value of secondary emission ratio is less than two. For alloys and composite surface it is as high as 10. In this case the emitter with thin coating of cesium and cesium oxide on silver base has the highest value of secondary emission ratio.

7·12. FIELD EMISSION :

When the energy applied for electron emission is in the form of electrostatic energy, the process is known as field emission.

The emission is explained due to two effects :

(1) Schottky Effect : It will be discussed in article 7·13 that for low anode voltages and high temperatures, the current in a diode is space-charged limited, since at such voltages anode is not able to draw whole of the electrons emitted from cathode, but some of them return to the cathode and thus set up a negative gradient on it. If however, the voltage is increased, this negative gradient decreases and with suitable voltage value, the current becomes temperature limited as discussed in articles 7·3 and 7·4. But it is seen that even in the temperature limited region, current from a thermionic emitter never truly saturates and its value increases with anode potential. The effect is called *Schottky effect*. The reason is quite obvious. When the valve operates under temperature-limited current condition, then space charge is not dense enough to make the electric field zero at the cathode; but, on the other hand if anode potential is high, a rising potential gradient would exist at the cathode which tends to draw electrons from it (hence field emission). Such a gradient alters the potential energy barrier at the surface. We shall, therefore, in order to account for Schottky effect, calculate the decrease in work function due to the alteration in potential barrier height (or energy).

The force acting on an electron, when it is removed from metal surface by a distance large compared with the spacing of the atoms in the crystal lattice, can be calculated by assuming metal surface to be a conducting plane surface. The force, so accounted, very well explains the shape of the potential energy curve at the right of figure 1 (*c*) except in a region close to the metal boundary.

Now the forces between an electron, charge '$-e$' and plane conducting surface (since metal surface is assumed to be so) can be computed by assuming an electric image $+e$ at the same distance, but on other side of the metal surface as shown in figure 7. According to Coulomb's law, the attractive force or inward force is

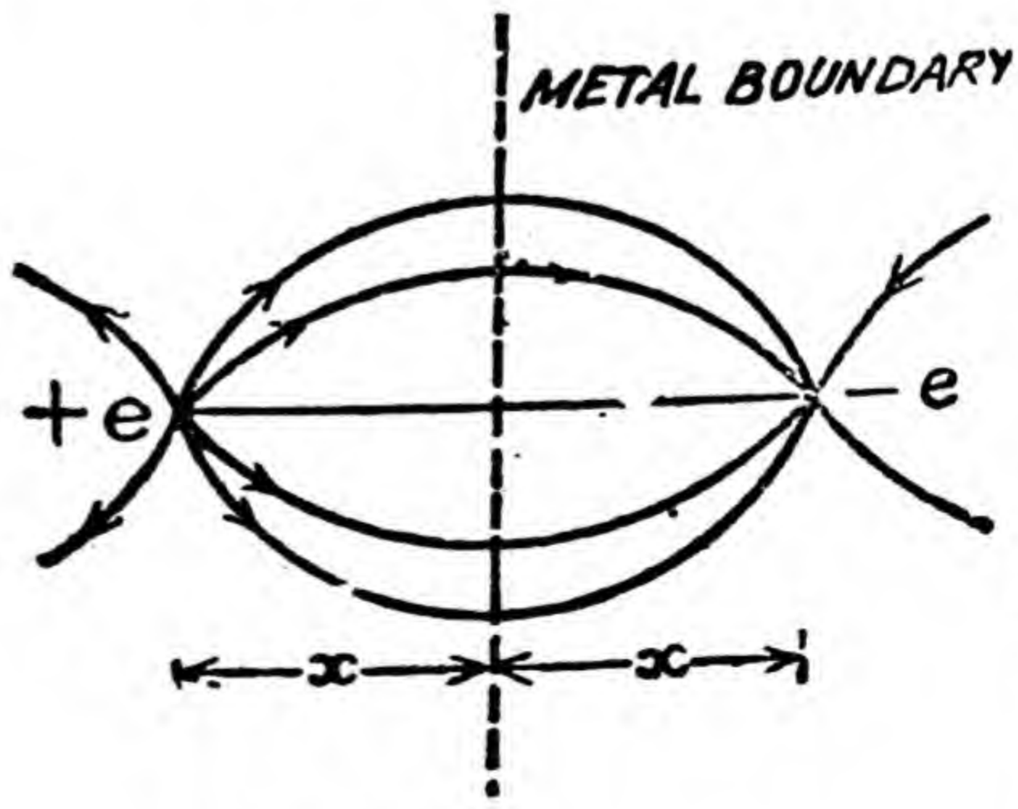


Fig. 7.

$$P = \frac{e^2}{16\pi\epsilon_0 x^2}$$

where x is the distance from the surface to the electron. In fig. 8 this force function is plotted over the range of x in which it applies.

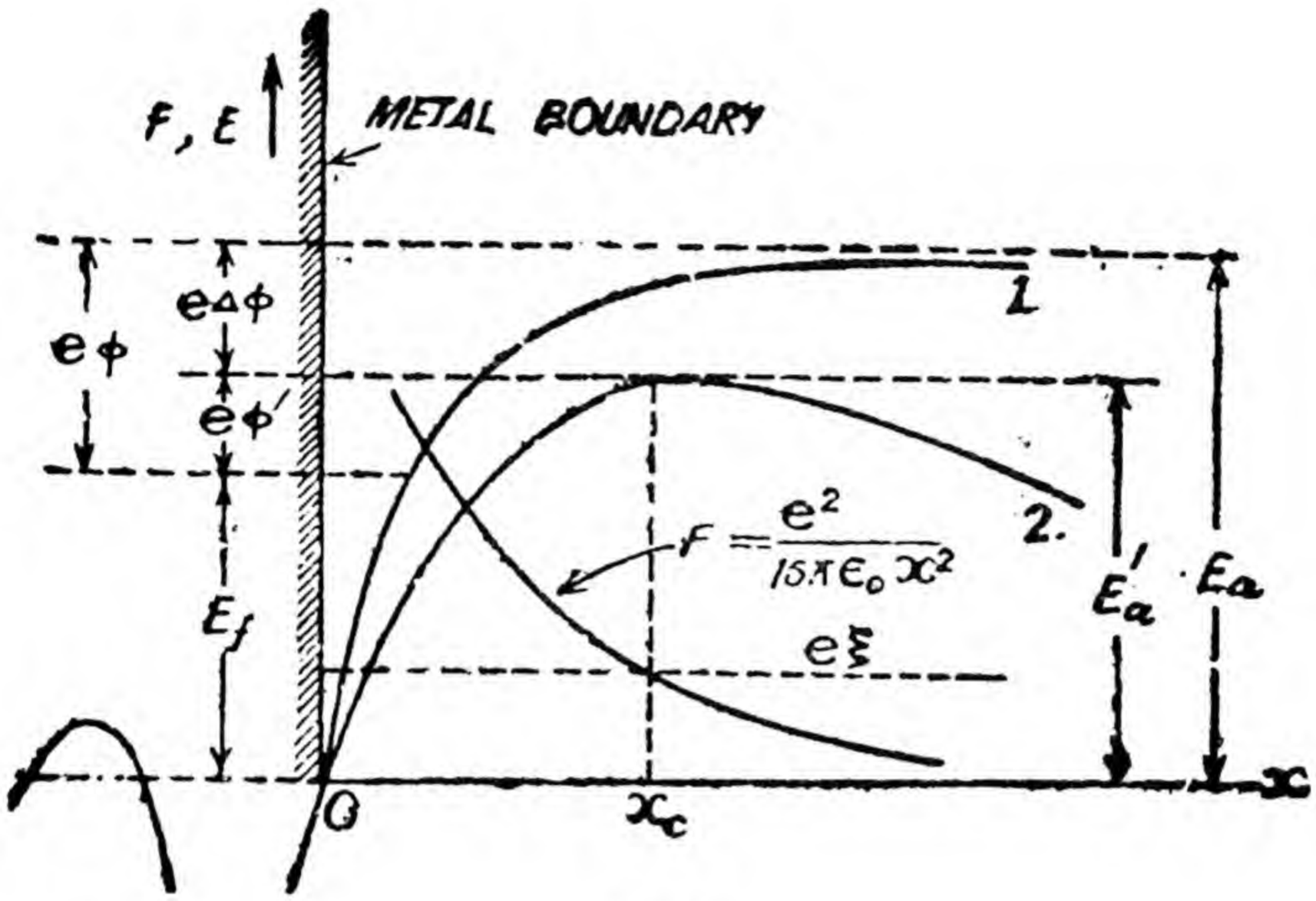


Fig. 8.

When an electric field ξ is set up at the cathode surface (due to high anode potential), then force $e\xi$ tends to draw electrons away from metal surface and thus acts in opposite direction to inward force. Now at a critical distance x_c, the two components of force equal each other and net force on the electron is zero. For larger distance, the force due to electric field predominates and the etectron is drawn away from the surface. Thus *the electron escapes in the presence of electric field if it has sufficient kinetic-energy on leaving the metal to reach the point where* x *equals* x_c. Except at very small distance, the inward force is given by the theory of images,

i.e.,
$$F=\frac{e^2}{16\pi\epsilon_0 x^2}.$$

At critical point $x=x_c$.

$$\therefore \quad F=\frac{e^2}{16\pi\epsilon_0 x_c^2}=e\xi$$

or
$$x_c=\frac{1}{2}\sqrt{\left(\frac{e}{\pi\epsilon_0\xi}\right)}$$

where ξ is the electric field intensity (equal to $\frac{dV}{dx}$, the applied potential gradient, V being the potential due to electrostatic field at a point distant x from cathode).

In fig. 8, curve 1 represents the potential energy barrier which exists in the absence of an external field, while curve 2 shows the potential energy barrier modified by the presence of external field. It is obvious that the height of the potential barrier

is reduced which clearly means that less work is done due to the presence of an external field in removing an electron from the metal, *i.e* value of work function is reduced.

We shall calculate the reduction in work done in two parts : First reduction in work done in taking an electron from inside the metal to the point x_c. This can be calculated as follows :

Over small distances x_o from the cathode to the crest of the barrier (point x_o), we can take ξ to be constant, *i.e.*, force $e\xi$ to be constant. The inward force is reduced by an amount $e\xi$ and hence the reduction in work, required to bring an electron from inside the metal to point x_o, will be

$$e\xi x_o.$$

Secondly, since at x_o electron is quite free to escape, the complete work done by the external field from point x_o to ∞ (escape) will be the net reduction in the work necessary to move an electron out from the critical point. The reduction for this case, in the work done is, therefore,

$$\int_{x_o}^{\infty} \frac{e^2}{16\pi\epsilon_0 x^2}.\ dx = \frac{e^2}{16\pi\epsilon_0 x_c}.$$

The total reduction in the work done due to the presence of an external field, is obviously the total reduction in kinetic energy for the escape of an electron from the metal, *i.e.*

$$E_a - E_a' = e\xi x_o + \frac{e^2}{16\pi\epsilon_0 x_o},$$

where E_a is the heignt of potential barrier in the absence of an external field and E_a' is the height in the presence of the field. Putting the value of x_c in above expression, we get

$$E_a - E_a' = \tfrac{1}{2}e\sqrt{\left(\frac{e\xi}{4\pi\epsilon_0}\right)} + \tfrac{1}{2}e\sqrt{\left(\frac{e\xi}{4\pi\epsilon_0}\right)}$$

$$= e\sqrt{\left(\frac{e\xi}{4\pi\epsilon_0}\right)}.$$

The kinetic energy of the electron, within the metal, is unaffected by the external field. Therefore, the Fermi energy E_f which depends upon the kinetic energy of electrons remains unaffected and hence the decrease in the height of potential barrier should directly represent the decrease in the function (since the common reference level E_f is unchanged) *i e.*

$$\text{decrease in work function} = e\,\Delta\phi = e.\sqrt{\left(\frac{e\xi}{4\pi\epsilon_0}\right)}.$$

The expression shows that the decrease in work function is directly proportional to the square root of external electric field and, therefore, requires the attention to be paid while studying

thermionic emission under high electrostatic field conditions*.

Now true work function $e\phi'$ will be

$$e\phi' = e\phi - e\sqrt{\left(\frac{e\xi}{4\pi\epsilon_0}\right)},$$

where $e\phi$ is the work function in the absence of an external field.

Richardson equation should, therefore, be modified in view of the Schottky effect. Richardson-Dushman equation for temperature-limited current per unit area is

$$J_S = AT^2 e^{-e\phi/kT}$$

Let $(J_S)_f$ be the saturation current density per unit area in the presence of the field, then

$$(J_S)_f = AT^2 e^{-e\phi/kT}$$

$$= AT^2 e^{-\left\{e\phi - e\sqrt{\left(\frac{e\xi}{4\pi\epsilon_0}\right)}\right\}/kT}$$

$$= AT^2 e^{-e\phi/kT} \cdot e^{e\sqrt{\left(\frac{e\xi}{4\pi\epsilon_0}\right)}/kT}$$

$$= J_S \, e^{\frac{e^{3/2}}{2kT}\cdot\sqrt{(\xi/\pi\epsilon_0)}}$$

On substituting the numerical values of e, ϵ_0 and k, the expression takes the form

$$(J_S)_f = J_S . e^{0\cdot 44\sqrt{\xi}/T},$$

where T is temperature of cathode in degree Kelvin and ξ is in volts per meter. If α be the area of cathode, then

$$\alpha.(J_S)_f = \alpha\, J_S . e^{0\cdot 44\sqrt{\xi}/T},$$

$$(I_S)_f = I_S . e^{0\cdot 44\sqrt{\xi}/T},$$

where $(I_S)_f$ and I_S represents the currents corresponding to $(J_S)_f$ and J_S current densities.

Now to calculate the true saturation current, we can put the expression

$$(J_S)_f = J_S . e^{\frac{e^{3/2}}{kT}\cdot\sqrt{(\xi/\pi\epsilon_0)}}$$

---

*For a diode having axial symmetry, if V is the anode potential, r_a and r_k are the anode and cathode radii respectively, then field ξ is given by the formula

$$\xi = -\left(\frac{V}{r_k}\right).\log_e\left(\frac{r_a}{r_k}\right).$$

Suppose V=20,000 volts, r_a=0·5 cm. and r_k=0·001 cm., then

$$\xi = -3\cdot 22\times 10^8 \text{ volts/cm.}$$

Corresponding to this value of ξ the decrease in the work function $\Delta\phi$, becomes to be 0·68 volts which, of course, is appreciable.

in the form

$$(I_S)_f = I_S . e^{\frac{e^{3/2}}{2kT}\cdot\sqrt{(\xi/\pi\epsilon_0)}}$$

or $$\log_e (I_S)_f = \log_e I_S + \frac{e^{3/2}}{2kT}\cdot\sqrt{\left(\frac{\xi}{\pi\epsilon_0}\right)}$$

or $$2{\cdot}3026.\log_{10}(I_S)_f = 2{\cdot}3026.\log_{10} I_S + \frac{e^{3/2}}{2kT}\cdot\sqrt{\left(\frac{\xi}{\pi\epsilon_0}\right)}$$

or $$\log_{10}(I_S)_f = \frac{e^{3/2}}{2{\cdot}3026.2kT}\cdot\sqrt{\left(\frac{\xi}{\pi\epsilon_0}\right)} + \log_{10} I_S,$$

which is of the form $y=mx+c$, representing a straight line. Therefore if a graph in $\log_{10}(I_S)_f$ and $\sqrt{\xi}$ is plotted, a straight line will be obtained with slope of

$$\frac{e^{3/2}}{4{\cdot}6052kT}\cdot\frac{1}{\sqrt{(\pi\epsilon_0)}}$$

and an intercept on the $\log_{10}(I_S)_f$ axis of $\log_{10} I_S$. From the intercept value, I_S, can be computed.

The value of $(I_S)_f$ depends upon the cathode temperature T as shown by the expression.

$$(I_S)_f = I_S e^{0{\cdot}44\sqrt{\xi}/T},$$

but it was found experimentally that even at low temperatures, sufficient current can be drawn with high field application which cannot be explained by Schottky effect.

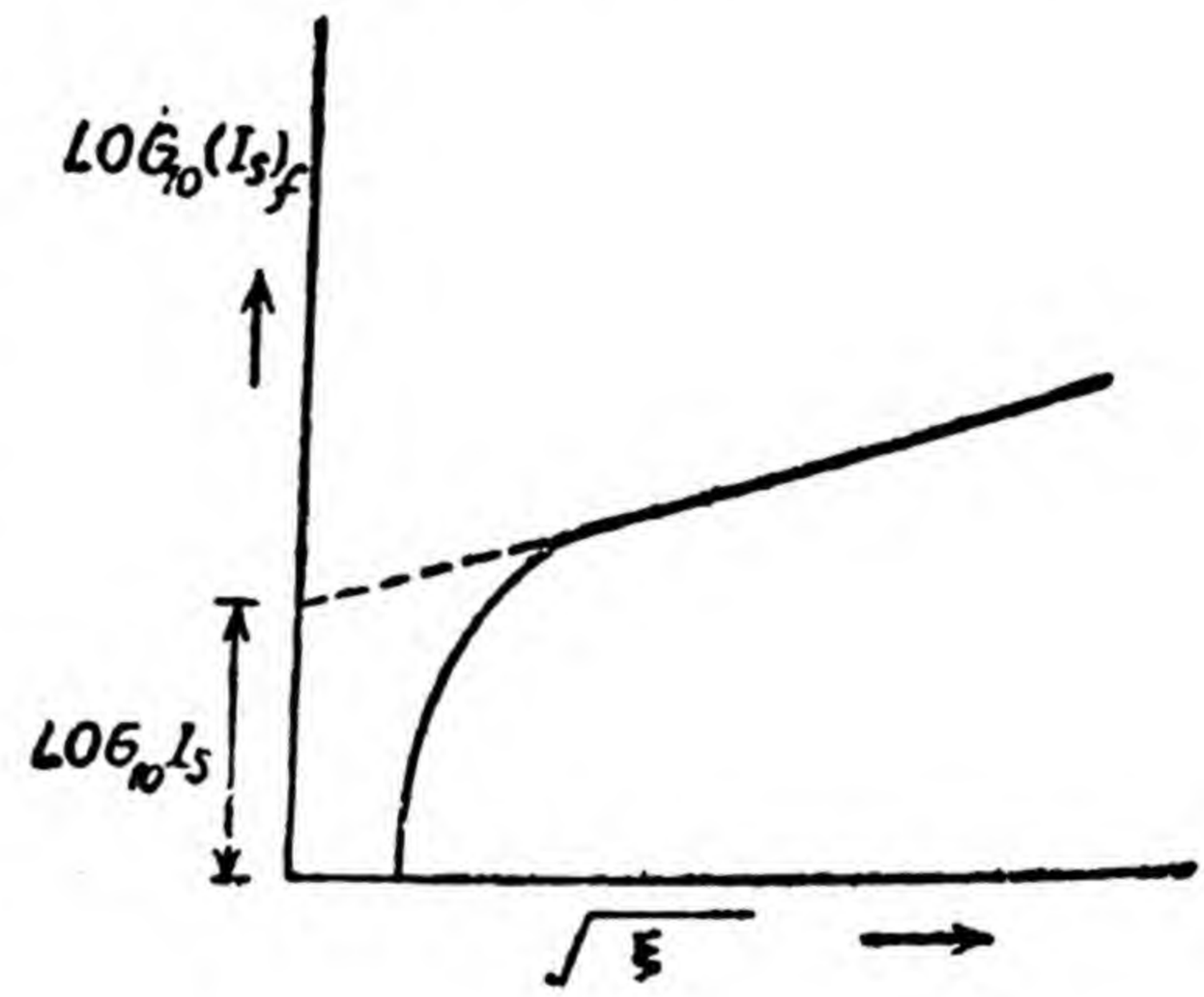


Fig. 9.

(2) Cold Cathode Effect : This effectaccounts for th e above experimental fact. Cold Cathode effect deals with the wave nature of electrons. Then some of the arriving electrons are reflected back, while others are liberated.

In the absence of electric field, the barrier extends to infinity but when electric field is applied then height of potential barrier is reduced. At the time, the electrons having energy slightly less than potential barrier may penetrate it. This number increases with the increase of field without raising temperature. Thus this type of emission is almost independent of temperature. The relation of Schottky was modified as

$$I_f = A_f \xi^2 e^{-b_f/\xi},$$

where $I_f \rightarrow$ density of field emission current,
$b_f \rightarrow$ constant of metal,
$A_f \rightarrow$ constant coefficient,
$\xi \rightarrow$ electric field intensity at cathode surface.

The above relation is similar to Richardson equation for thermionic emission.

7·13. SPACE CHARGE AND TEMPERATURE LIMITED CURRENTS :

Consider a diode valve with plane parallel electrodes or cylindrical electrodes. Suppose the cathode temperature T_1 is constant ; then as shown in figure 10, at first the plate current increases as the plate potential is increased from zero. But after a certain value of plate potential V_{p_1} the plate current becomes constant. The diode then operates at a saturated plate current I_{p_1} or in other words all the electrons emitted from the cathode are attracted by the plate.

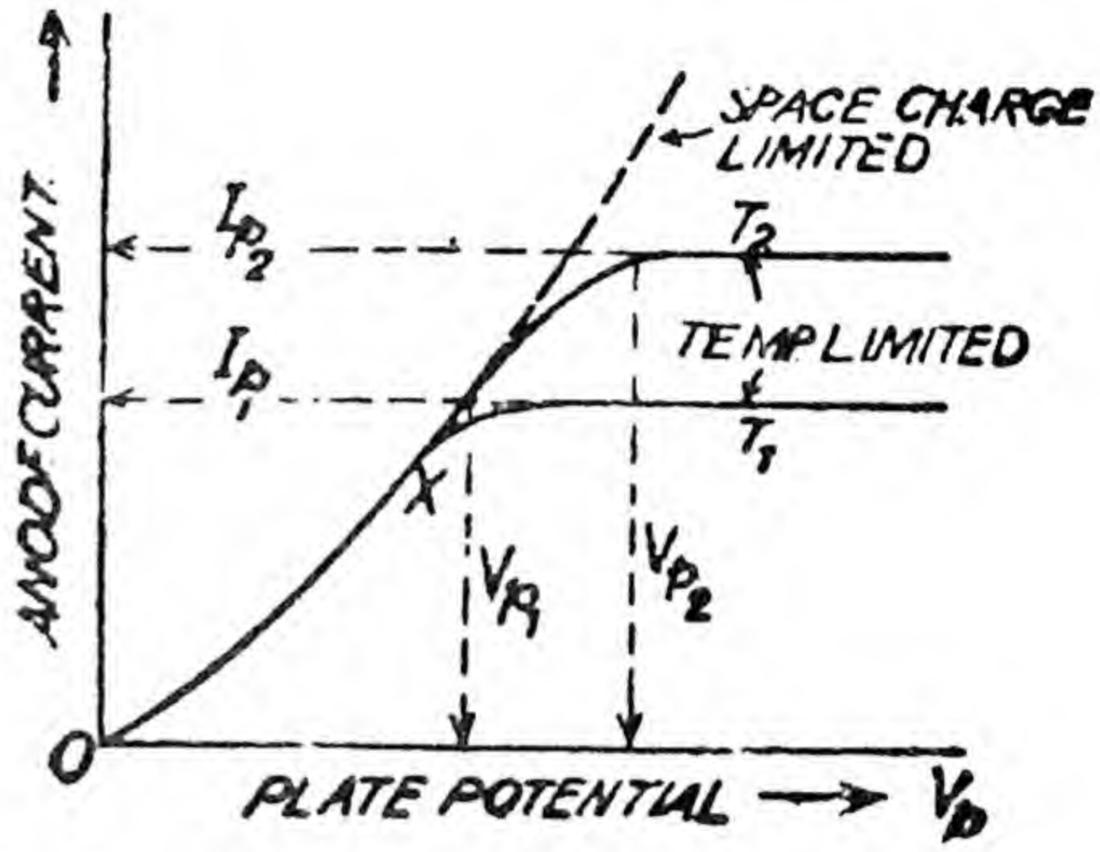


Fig. 10. Plate characteristic of a diode.

Now if cathode temperature is raised to T_2, the *saturated* current value also increases because of the fact that at higher temperature of cathode emission of electrons will be greater than at temperature T_1 and then there will be more electrons to reach the plate for saturated current. This saturation current is therefore the total cathode emission current given by the Richardson's formula.

When the plate potential is sufficiently high, electrons are attracted by the plate as rapidly as emitted from the cathode. The plate current is then limited by the electron emission of the cathode and so depends upon cathode temperature rather than upon the plate voltage as shown in fig. 11. A high plate potential V_{pH} is applied but it can be seen from fig. 11 that saturation current is different depending upon cathode temperature. Thus this saturation current portion which depends upon cathode temperature is said to be '*temperature limited*'.

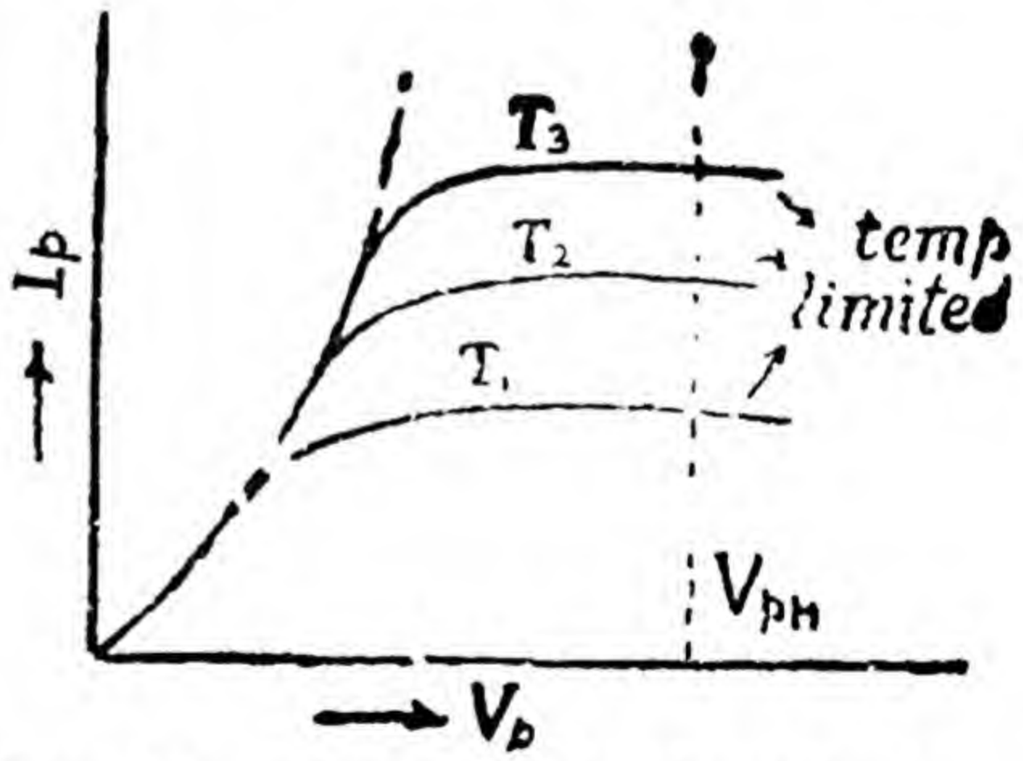


Fig. 11 Volt-ampere characteristic of diode

At plate potential lower than the saturation plate potential, *i.e.* over the region OX of fig. 10, the plate current is less than the total emission current. The number of electrons emitted by cathode is the same but due to space charge effect, some of these electrons do not reach the anode and are returned to cathode.

At low or moderate plate potentials, the electrons are emitte

from the cathode at a greater rate than they are collected by the plate. Thus the emission rate depends upon Richardson's law but the collection rate depends on the acceleration of electric field due to the plate potential. Consequently, electrons form a cloud of negative charge near the cathode called the space charge region. This negative space charge, which reduces the attraction of positive plate on the electrons just leaving cathode, repels these electrons (leaving the cathode) so that some of the electrons depending upon the density of the space charge are returned to the cathode instead of reaching the plate. At negative plate potential the density of space charge region is very much increased due to which all the electrons, emitted from cathode, are returned to the latter and the density of space charge decreases due to which more and more electrons approach the plate until at the saturation plate potential all the electrons are attracted by the plate as soon as they are emitted and hence a saturation current flows in the plate diode. Therefore OX portion of the curve is same at the same plate potential whatever be the temperature. It is supposed that the electrons that are emitted from the cathode have zero initial velocities.

The density of the electrons and the potential at any point in the interelectrode space are related by the Poission's equation

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\epsilon_0}$$

where V is potential in volts,
ρ is the magnitude of the electrode charge density,
and ϵ_0 is permittivity of free space.

At temperature T at which there is no emission, *i.e.* $\rho=0$, then

$$\frac{d^2V}{dx^2} = 0.$$

Integrating twice, we get

$$V=ax+b,$$

where a and b are arbitrary constants which are to be determined by boundary conditions *i.e.*

(*i*) when $x=0$, $V=0$ (potential of cathode=0)
(*ii*) when $x=d$, $V=V_p$ (potential of the plate=V_p)
(distance between cathode and plate=d)

Applying these conditions, we have

$$b=0 \text{ and } a=\frac{V_p}{d}.$$

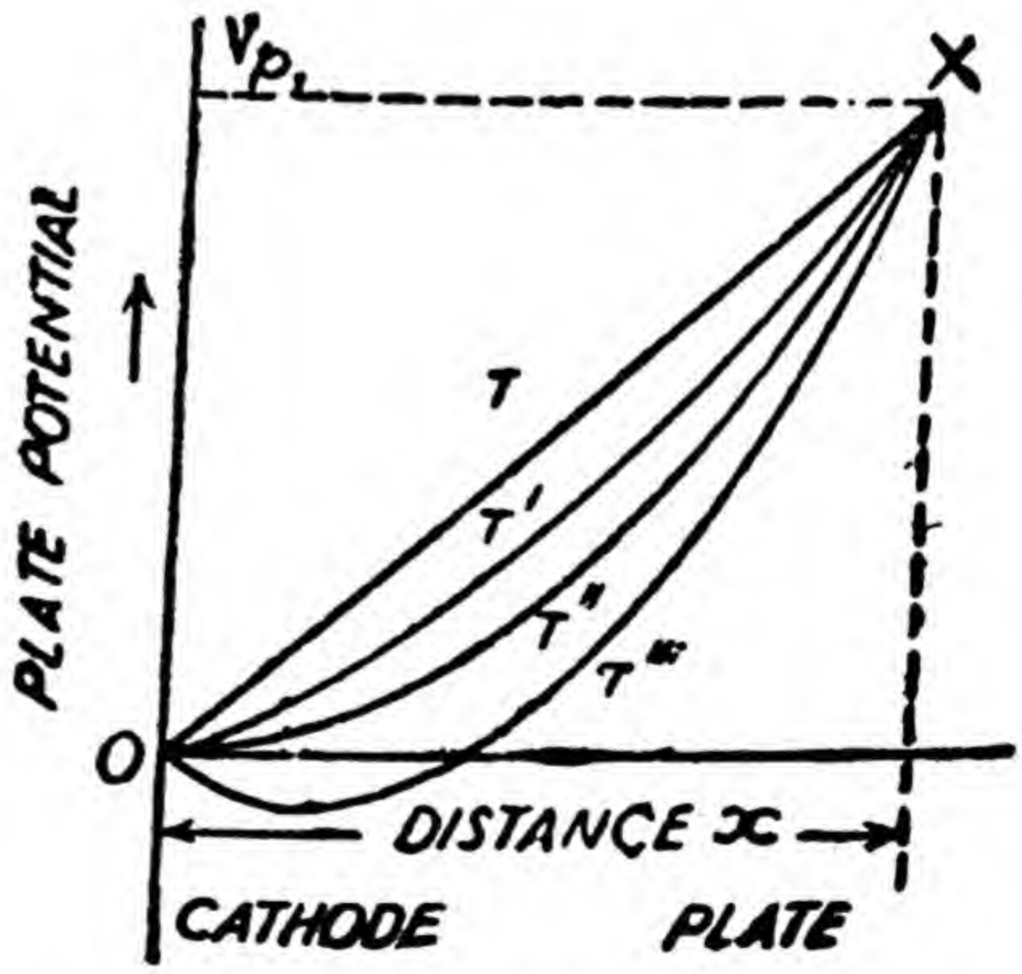


Fig. 12. Potential distribution curves of a parallel plate electrode.

Substituting these values, we get

$$V=\frac{V_p x}{d},$$

which is the equation of a straight line.

Therefore, the potential distribution is a linear function of the distance from cathode to the plate as shown in fig. 12.

At higher temperature T, some electrons are emitted, *i.e.* ρ is not zero, but has a value equal to $-\rho_1$. The negative sign is used because the charge density is due to negative charge space electrons. The potential distribution is represented by corresponding curve in fig. 12. When ρ is not zero $\frac{d^2V}{dx^2}=\frac{\rho_1}{\epsilon_0}$ is positive which is also a measure of curvature. In this case

$$\frac{d^2V}{dx^2}=\frac{\rho_1}{\epsilon_0}.$$

Integrating twice, we get

$$V=\frac{\rho_1 x^2}{2\epsilon_0}+a_1x+b_1$$

where a_1 and b_1 are arbitrary constants. The boundary conditions are

(*i*) when $x=0$, $V=0$

(*ii*) when $x=d$, $V=V_p$,

Applying these conditions, we get

$$b_1=0 \text{ and } a_1=\frac{V_p}{d}-\frac{\rho_1 d}{2\epsilon_0}.$$

Substituting values, we get

$$V=\frac{\rho_1 x^2}{2\epsilon_0}+\left(\frac{V_p}{d}-\frac{\rho_1 d}{2\epsilon_0}\right)x=\frac{V_1 x}{d}-\frac{\rho_1 x}{2\epsilon_0}(d-x),$$

which is the equation of parabola.

A positive curvature means that the change in slope $\frac{dV}{dx}$ between two adjacent points must be positive and therefore all the curves must be concave upwards. Since anode-cathode potential is independent of cathode emission, all curves must pass through fixed end points.

With increasing values of cathode temperature, the potential distribution curves are depressed more and more until a temperature T'' is reached at which slope of curve is zero at the cathode, *i.e.* $\frac{dV}{dx}=0$ at $x=0$ or in other words, the electric field on the electrons emitted from the cathode is just zero. This constitutes the limit of electrons attracting capacity of the plate at potential V_{p_1}.

If temperature is increased further then slope of the curve

$\frac{dV}{dx}$ at cathode will be negative represented by a curve corresponding to T'' temperature. This condition exists momentarily, because $\frac{dV}{dx}$ is now negative which will exert a repeling force on all electrons tending to be emitted from the cathode with the result that these excess electrons re-enter the cathode and then

$$\left(\frac{dV}{dx}\right)_{x=0} = 0$$

condition is reached which is almost stable stage for all temperatures near or greater than T''.

The condition $\left(\frac{dV}{dx}\right)_{x=0} = 0$ is valid under the assumption of zero initial velocities of emission of electrons.

7·14-1. CHILD-LANGMUIR LAW FOR PLANE PARALLEL ELECTRODES

The equation relating the anode current with anode voltage for plane parallel electrodes, under space charge limited condition, was first derived by Child in 1911. Later on, in 1913, the analysis was extended to cylindrical electrodes by Langmuir.

Derivation : Let us consider two plane parallel electrodes of diode. For this derivation we make certain assumptions :

(*i*) The spacing between cathode and plate which are plane parallel plates is small such that the electrostatic field between them is uniform and normal to the electrostatic field between them is uniform and normal to the electrode surface.

(*ii*) The surfaces of anode and cathode are equipotential. If we take cathode as the plane $x=0$, then there is no variation of potential in y and z directions and electron moves parallel to the axis of x.

(*iii*) Electrons have zero initial velocity after emission from cathode.

(*iv*) The cathode provides an infinite supply of electrons.

(*v*) When the current is space charge limited, the electric field at the cathode is zero.

The distribution of potential is governed by Poisson's equation

$$\frac{d^2V}{dx^2} + \frac{d^2V}{dy^2} + \frac{d^2V}{dz^2} = -\frac{\rho^*}{\epsilon_0}$$

where ρ is the space charge density in coulombs per cubic metre and ϵ_0 is permittivity of free space.

As described earlier, the potential varies only in x-direction. Therefore,

$$\frac{d^2V}{dy^2} = \frac{d^2V}{dz^2} = 0.$$

In M K S system of units

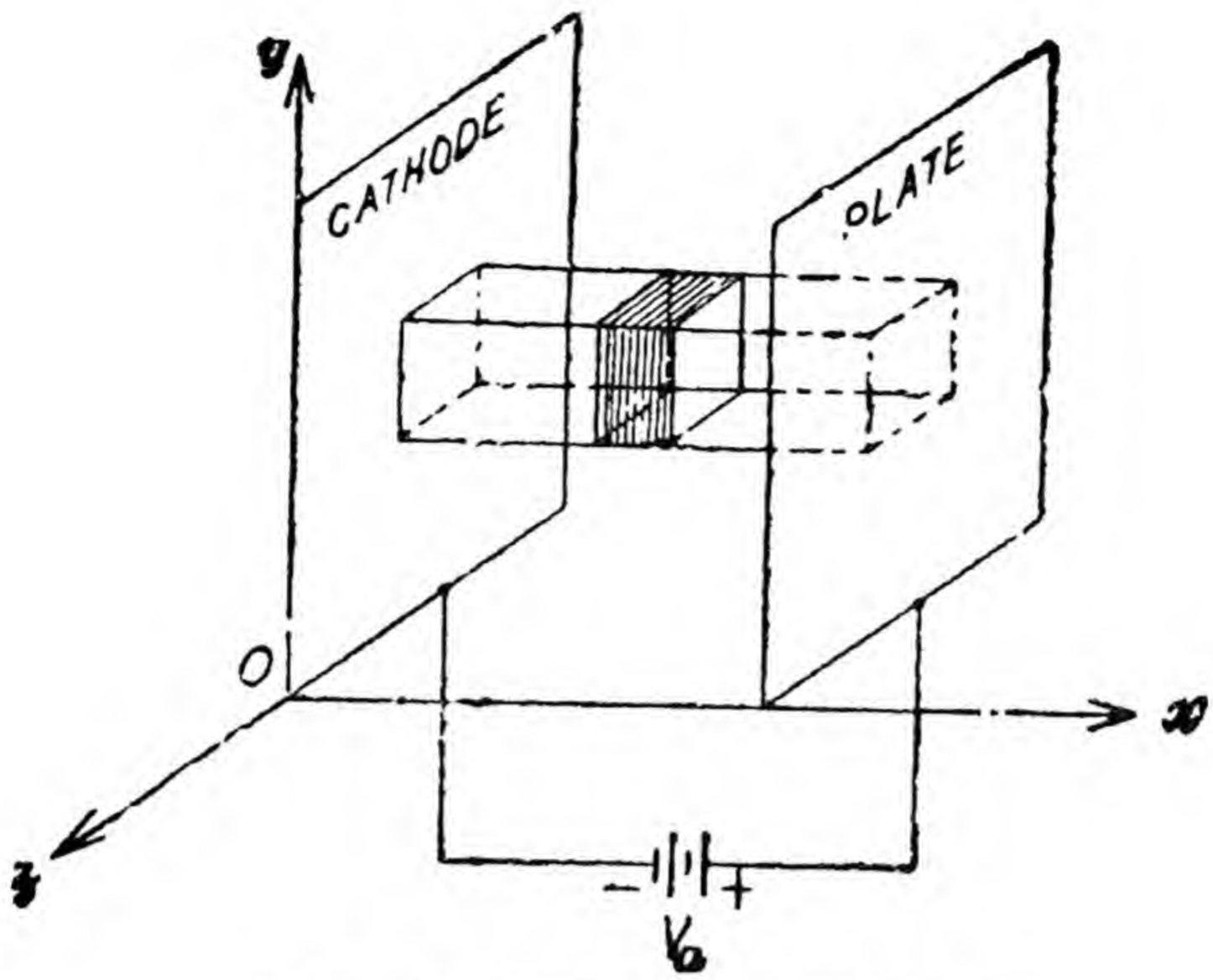


Fig. 13. A plane parallel diode.

Now we can write the Poisson's equation as

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\epsilon_0}.$$

We know the following facts :

(*a*) The velocity v of an electron at any point of potential V is related by

$$\tfrac{1}{2}mv^2 = eV, \qquad \ldots(1)$$

where e = charge on electron = $1{\cdot}6 \times 10^{-19}$ coulomb,

m = mass of electron.

V is plate-cathode potential difference.

(*b*) The current density due to the flow of electrons may be calculated by the rate at which electrons move through unit area of cross section per second in the direction of electric field. Let

v = velocity of electrons in metre per sec.,

e = charge on an electron in coulombs.

N = electron density per cubic metre.

Then the current density per ampere square metre is given by $J = Nev$ ...(2)

We know that space charge is due to negatively charged electrons ; therefore $\rho = -Ne$.

Putting above value of space charge density, the Poisson's equation becomes

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\epsilon_0} = \frac{Ne}{\epsilon_0} \qquad \ldots(3)$$

From (2), we have $Ne = J/v$,

$$\frac{d^2V}{dx^2} = \frac{J}{e_0 v}. \qquad \ldots(4)$$

From (1), we see that

$$v = \sqrt{\left(\frac{2eV}{m}\right)}.$$

Putting in equation (4), we have

$$\frac{d^2V}{dx^2}=\frac{J}{\epsilon_0}\sqrt{\left(\frac{m}{2eV}\right)}. \qquad ...(5)$$

For integration of eq. (5), we multiply each side by $\frac{dV}{dx}$,

$$\frac{dV}{dx}\cdot\frac{d^2V}{dx^2}=\frac{J}{\epsilon_0}\sqrt{\left(\frac{m}{2e}\right)}.V^{-1/2}.\frac{dV}{dx}$$

$$\frac{1}{2}.\frac{d}{dx}\left(\frac{dV}{dx}\right)^2=\frac{J}{\epsilon_0}\sqrt{\left(\frac{m}{2e}\right)}.V^{-1/2}.\frac{dV}{dx},$$

$$\frac{d}{dx}\left(\frac{dV}{dx}\right)^2=2.\frac{J}{\epsilon_0}\sqrt{\left(\frac{m}{2e}\right)}.2\frac{d}{dx}(V^{1/2}).$$

$$=4\frac{J}{\epsilon_0}\sqrt{\left(\frac{m}{2e}\right)}.\frac{d}{dx}(V^{1/2}).$$

Integrating it we get

$$\left(\frac{dV}{dx}\right)^2=\frac{4J}{\epsilon_0}\sqrt{\left(\frac{m}{2e}\right)}.V^{1/2}+C_1.$$

where C_1 is constant of integration.

But $dV/dx=0$ when $V=0$ since both vanish at $x=0$,

$$C_1=0.$$

$$\therefore \quad (dV/dx)^2=4J/\epsilon_0.\sqrt{(m/2e)}.V^{1/2},$$

$$(dV/dx)=2.\{J/\epsilon_0\sqrt{(m/2e)}\}^{1/2}V^{1/4},$$

$$V^{-1/4}\,dV=2.\{J/\epsilon_0.\sqrt{(m/2e)}\}^{1/2}\,dx.$$

Again integrating,

$$\tfrac{4}{3}V^{3/4}=2\{J/\epsilon_0.\sqrt{(m/2e)}\}^{1/2}.x+C_2.$$

At $x=0$, $V=0$, therefore $C_2=0$.

$$\therefore \quad V^{3/4}=\tfrac{3}{4}\{J/\epsilon_0\sqrt{(m/2e)}\}^{1/2}x.$$

Squaring, we get

$$V^{3/2}=\tfrac{9}{4}\{J/\epsilon_0.\sqrt{(m/2e)}\}.x^2. \qquad ...(6)$$

At anode, *i.e.* $x=x_a$.

$$\therefore \quad V=V_a$$

$$\therefore \quad V_a^{3/2}=\tfrac{9}{4}\{J/\epsilon_0.\sqrt{(m/2e}\}\,x_a^2. \qquad ...(7)$$

From (7), we find for current density, the expression

$$J=4\epsilon_0/9.\sqrt{(2e/m)}.V_a^{3/2}/x_a^2. \qquad ...(8)$$

This is Child-Langmuir equation for plane parallel electrodes.

In order to obtain the total space current we multiply the above equation by the cross-section S of the electron stream, so

$$I=SJ=S.4\epsilon_0/9.\sqrt{(2e/m)}.V_a^{3/2}/x_a^2. \qquad ...(9)$$

For a given electrode system,

$$I \propto V_a^{3/2}.$$

The relation (9) is also called the *three halves power* law.

The constant factor of equations (8) and (9) can be calculated since $\epsilon_0=8{\cdot}85.10^{-12}$ farad/metre, and $e/m=8{\cdot}66.10^{11}$ coulombs/kilogram.

Equations (8) and (9) can be rewritten as

$$J = 2{\cdot}34.10^{-6}\ \frac{V_a^{3/2}}{x_a^2} \qquad \ldots(10)$$

and

$$I = 2{\cdot}34.10^{-6}\ S\ \frac{V_a^{3/2}}{x_a^2}. \qquad \ldots(11)$$

If V_a is in volts, I will be in amperes. S and x_a should be measured in the same units.

Distribution of Potential and Space Charge Density :

Dividing eq. (6) by (7), we have

$$V/V_a = (x/x_a)^{4/3}. \qquad \ldots(12)$$

Also $\rho = -\epsilon_0 \dfrac{d^2V}{dx^2}$

$$= -\epsilon_0 \frac{d^2}{dx^2}\left\{V_a \left(\frac{x}{x_a}\right)^{4/3}\right\}$$

from (12)

$$= -\frac{4}{9\epsilon_0}\cdot\frac{V_a}{x_a^{4/3}}\ x^{-2/3} \qquad \ldots(13)$$

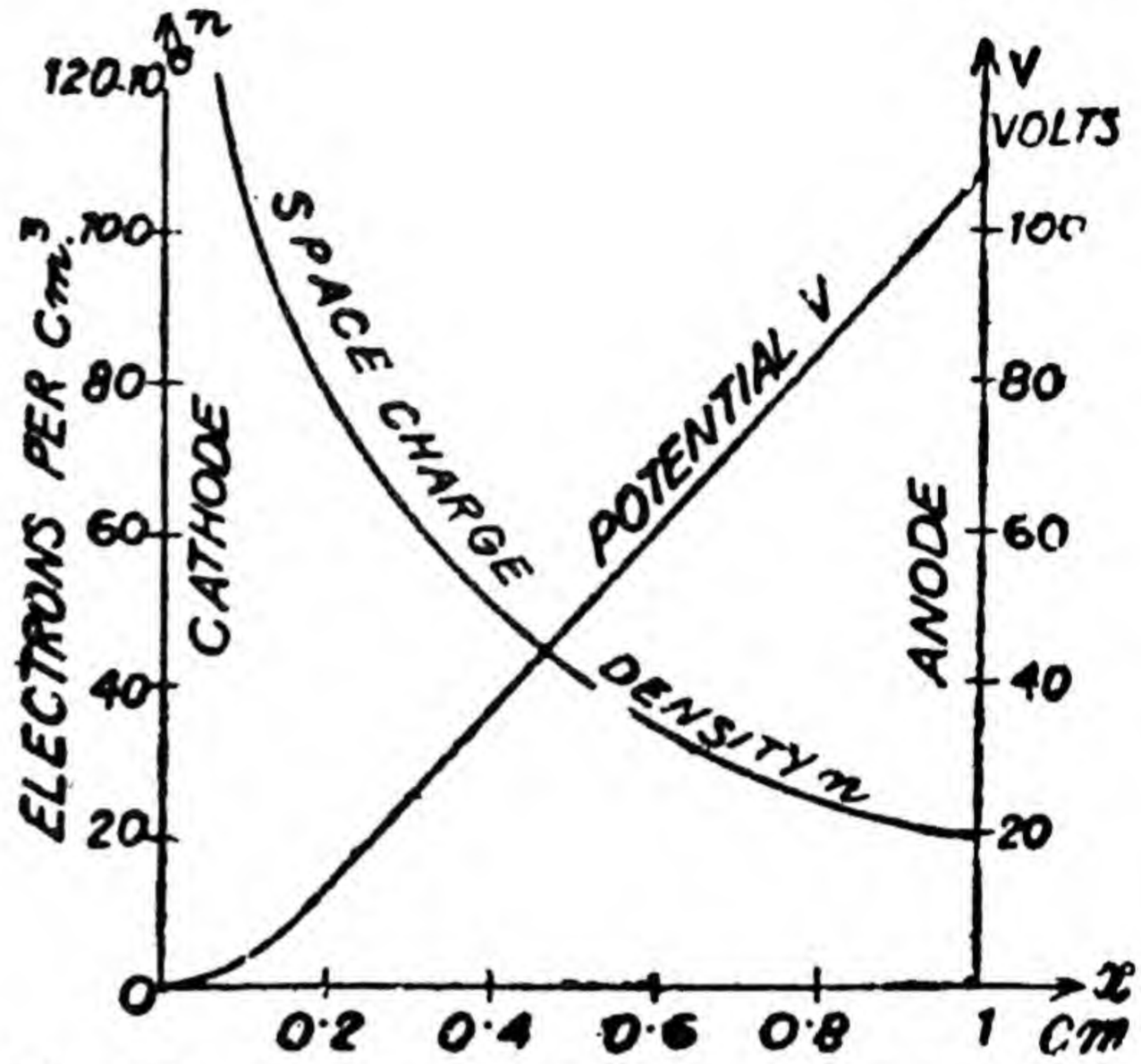


Fig. 14. Representation of potential and space charge distribution.

From equations (12) and (13 we can plot the distribution of potential and space charge density as shown in figure 14.

We see that potential distribution is almost linear except near cathode. Space charge density tends to infinity at cathode which is inaccordance with our assumption that cathode supplies an infinite number of electrons.

7·14-2. CHILD LANGMUIR EQUATION FOR CYLINDRICAL ELECTRODES :

The behaviour of cylindrical diode, *i.e.*, consisting cylindrical cathode and anode. can be studied in the same way as the plane electrode diode, but for the analysis of this case, cylindrical coordinates should be used. The problem is somewhat complicated due to the effect of the electrode dimensions.

Due to the symmetry of the electrode system, the problem will be one dimensional. The Poisson's equation for this case reduces to

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dV}{dr}\right) = -\frac{\rho}{\epsilon_0}. \qquad \ldots(1)$$

The same energy equation as for plane electrode case holds

i.e. $\frac{1}{2}mv^2 = eV.$ $\ldots(2)$

Current density at any radial distance r is given by

$$J(r)=\rho v. \qquad \ldots(3)$$

Since the current density varies with radial distance, it is more convenient to express the space charge density interms of total current per unit length of axis. That is

$$\frac{I}{l}=2\pi r J(r)$$

$$=2\pi r\rho v, \qquad \ldots(4)$$

where I is the total current passing at right angles through an imaginary cylinder of length l and radius r

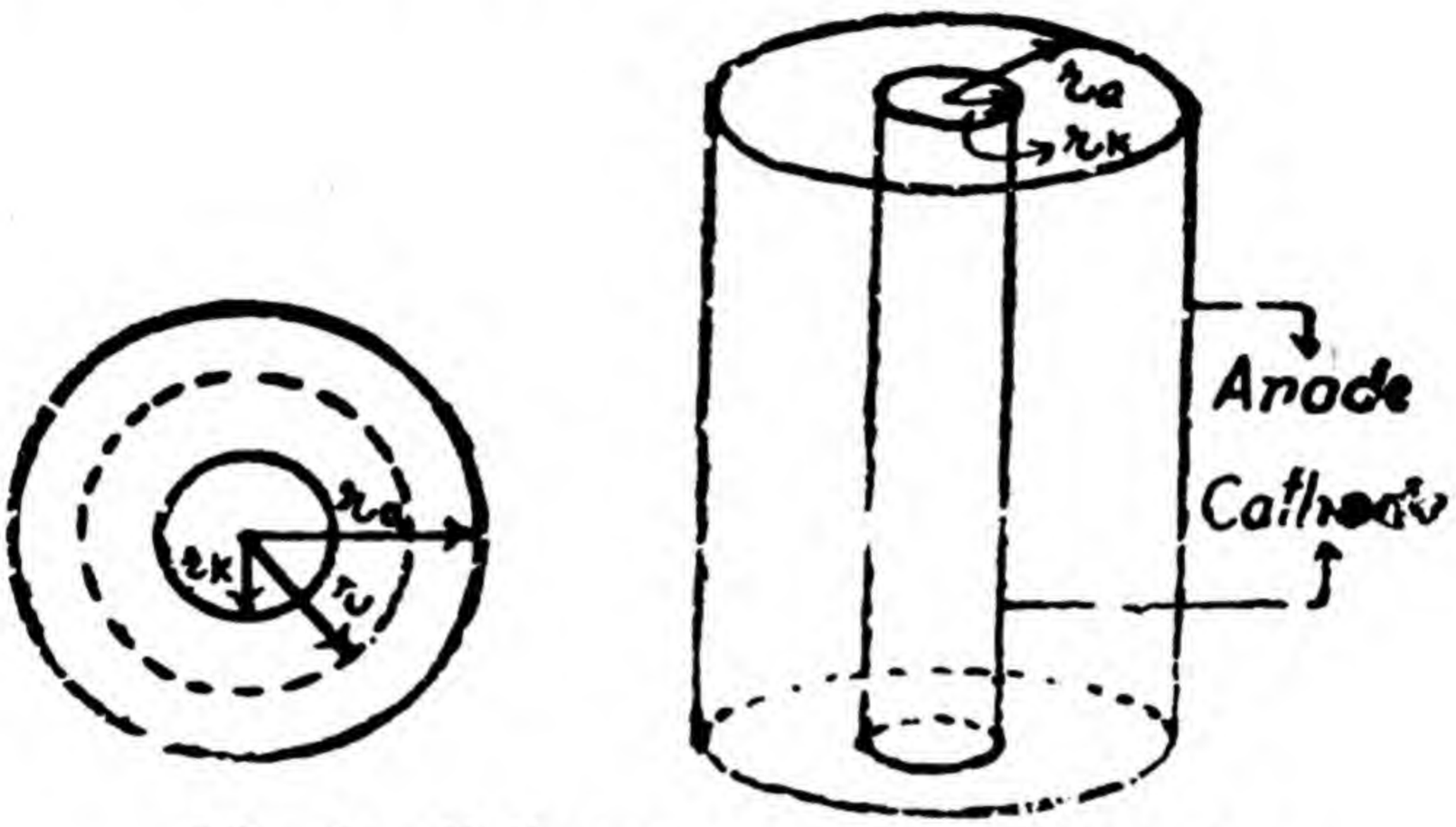


Fig. 1. Cylindrical electrodes of diode.

Putting these values of ρ from (4) in (1), we get

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dV}{dr}\right)=\frac{I}{2\pi rvl}\frac{1}{\epsilon_0}.$$

From (2), $$v=\sqrt{\left(\frac{2eV}{m}\right)}.$$

$$\frac{d}{dr}\left(r\frac{dV}{dr}\right)=\frac{I}{2\pi\epsilon_0 l}\sqrt{\left(\frac{m}{2eV}\right)}, \qquad \ldots(5)$$

$$r\frac{d^2V}{dr^2}+\frac{dV}{dr}=\frac{I}{2\pi\epsilon_0}\sqrt{\left(\frac{m}{2eV}\right)},$$

$$\frac{d^2V}{dr^2}+\frac{1}{r}\frac{dV}{dr}=\frac{J}{\epsilon_0}\sqrt{\left(\frac{m}{2eV}\right)}.$$

Putting $V=Ar^{3/2}$, the above equation can be solved to give

$$I=\frac{8\pi\epsilon_0 l}{9}\sqrt{\left(\frac{2e}{m}\right)}\cdot\frac{V_a^{3/2}}{r_a},$$

where V_a=anode potential
r_a=anode radius,

$$I=2\pi l.\frac{4\epsilon_0}{9}\sqrt{\left(\frac{2e}{m}\right)}\frac{V_a^{3/2}}{r_a}. \qquad \ldots(6)$$

We know from the analysis of the piane parallel electrodes that

$$\frac{4\epsilon_0}{9}\sqrt{\left(\frac{2e}{m}\right)}=2{\cdot}335.10^{-6}.$$

$$\therefore \quad \frac{I}{l}=\frac{2\pi.2{\cdot}335.10^{-6}}{r_a}.V_a^{3/2}. \qquad ...(7)$$

Equation (7) is approximate solution of the problem which holds within 10 percent for $r_a < 10\ r_k$, where r_k is cathode radius.

For an exact solution, we take the expression for the current

$$\frac{I}{l}=\frac{KV^{3/2}}{r\beta^2}, \qquad ...(8)$$

where β^2 is a function of the ratio of the radius at any point to the cathode radius, *i.e.* $\frac{r_a}{r_k}$. At anode $\beta^2=f\left(\frac{r_a}{r_k}\right)$.

Substituting equation (8) in (5), we shall have

$$2\beta r^2\frac{d^2\beta}{dr^2}+r^2\left(\frac{d\beta}{dr}\right)^2+7\beta r\frac{d\beta}{dr}+\beta^2-1=0. \qquad ...(9)$$

The above equation can be simplified slightly by letting $\alpha=\log_e(r/r_k)$. With this change of variable, equation (9) will take the form

$$3\beta\frac{d^2\beta}{d\alpha^2}+\left(\frac{d\beta}{d\alpha}\right)^2+4\beta\frac{d\beta}{d\alpha}+\beta^2-1=0. \qquad ...(10)$$

This may be solved by series method, the solution being

$$\beta=3\alpha-\frac{2\alpha^2}{5}+\frac{11\alpha^3}{120}-\frac{47\alpha^4}{3300}+\cdots$$

This expression is valid either for an internal or an external cathode.

From equation (8), it is obvious that the expression for current in practical units becomes

$$\frac{I}{l}=\frac{14{\cdot}66.10^{-6}\ V^{3/2}}{r_a\beta^2} \text{ amperes/unit length of axis.} \qquad ...(11)$$

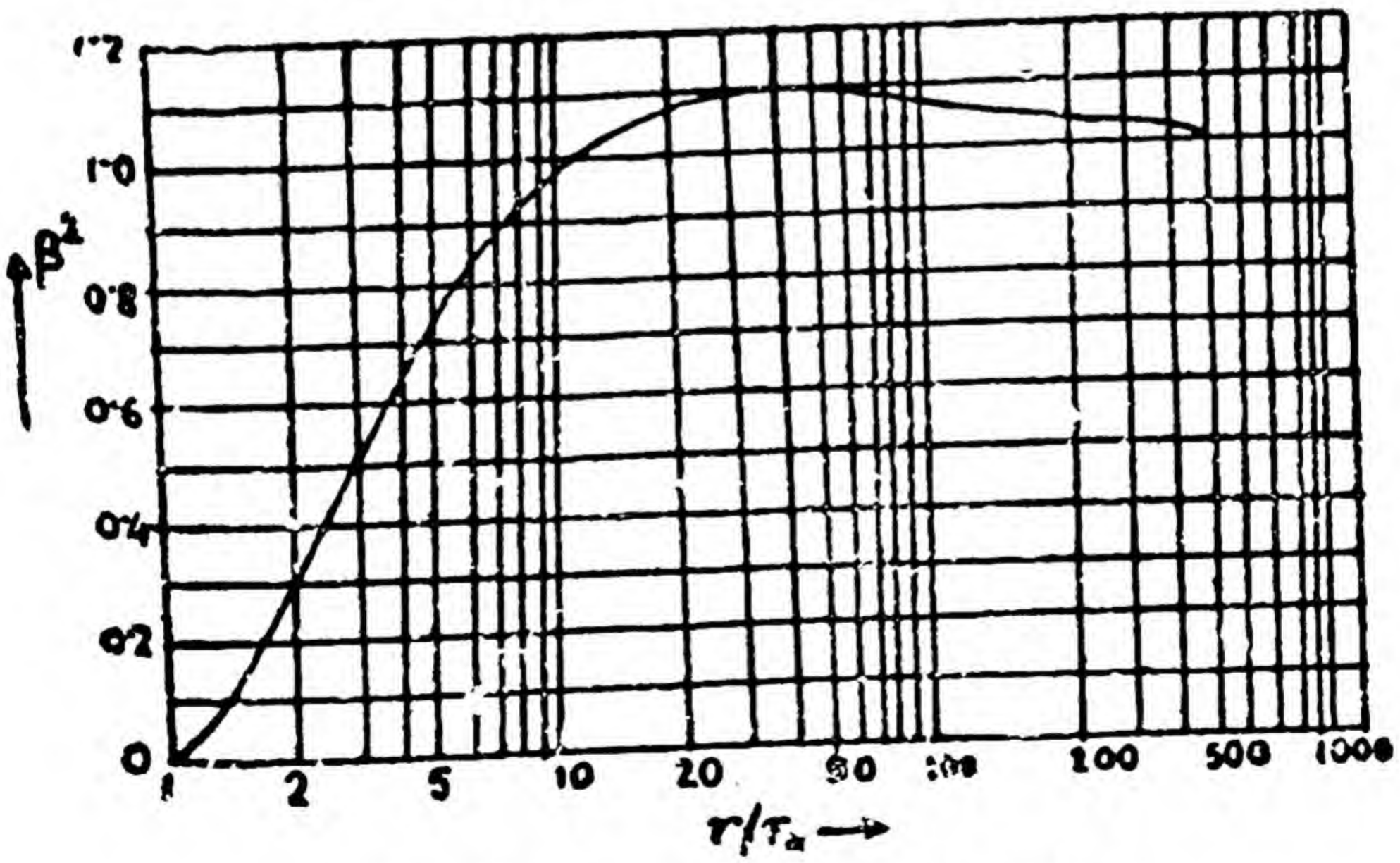


Fig. 16. Curve β^2 as a function (r/r_k).

Values of β^2 have been calculated by Langmuir by means of series for β. In fig. 16, values of β^2 are plotted against r/r_k.

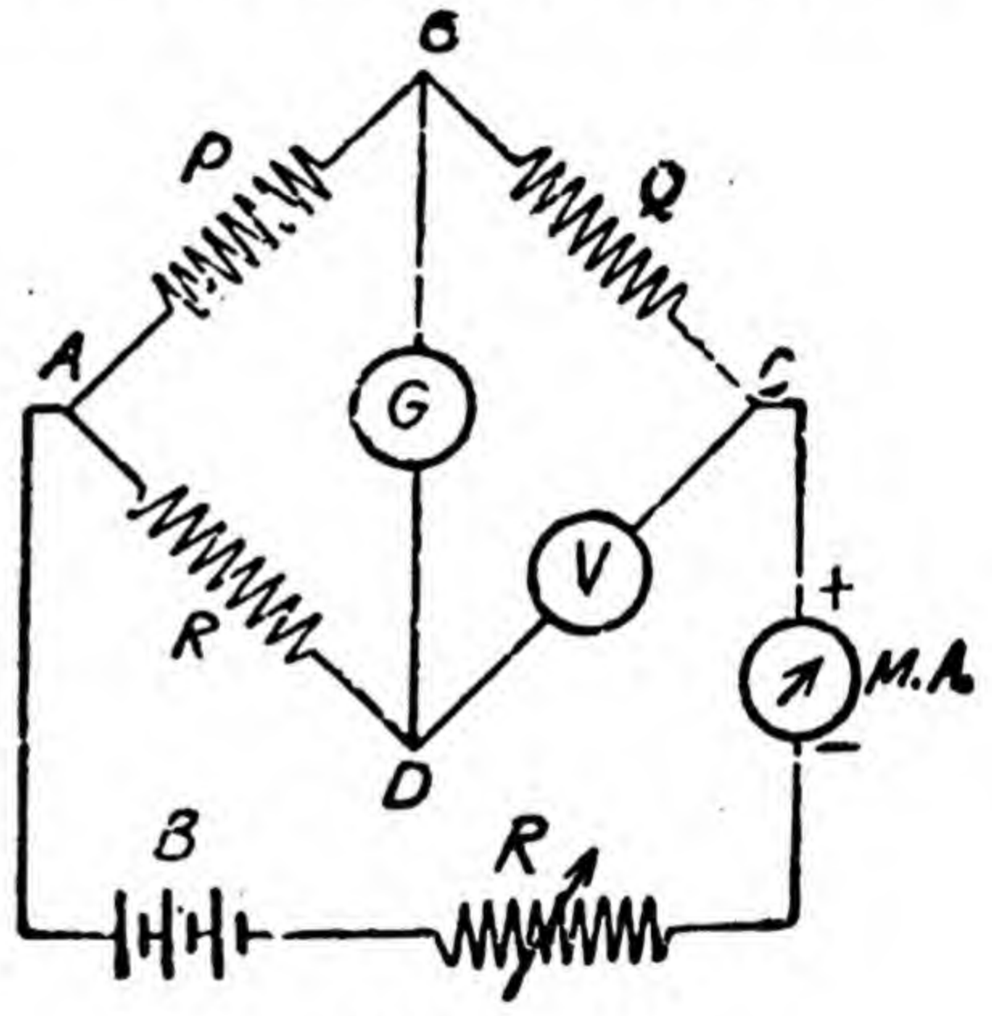
It is seen that for ratio $\frac{r}{r_k} > 7$, the value of β^2 differs from 1 by less than 10 percent and thus for this case the equation (7) can be adopted.

EXPERIMENT

To determine the work function of the emitter (pure metal) of vacuum tube.

Procedure : The work function of a pure emitter can be determined by means of Richardson equation. Circuit arrangement is shown in fig. 17.

(*i*) *Measurement of resistance of the filament at room temperature* : Connect the filament of the valve (1*G*4 *GT*) in the unknown arm of wheatstone bridge and complete the connections as shown in fig. 17*a*. Note down the room temperature. Determine the resistance of the filament for a particular value of current say 5 amp. Again decreasing the current in steps by rheostat *R*, determine the filament resistance for each current. Plot a graph in filament resistance and filament current as shown in figure 18. From this graph resistance of the filament R_0 at zero filament current is calculated. This is the value of filament resistance at room temperature T_R. The absolute temperature corresponding to R_0 will be given by



(*a*)

$$T_0 = 273 + T_R.$$

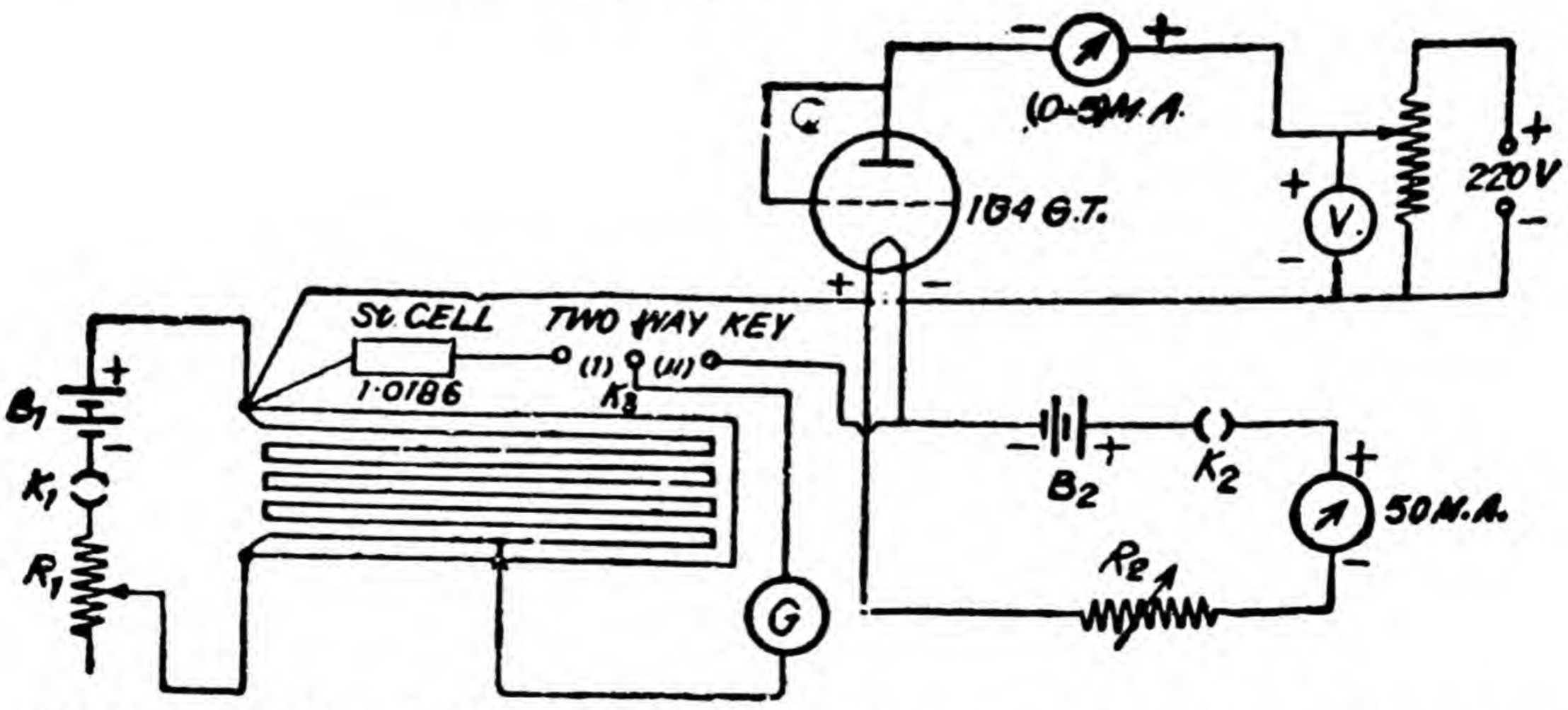


Fig. 17. (*a*) Circuit diagram for the measurement of filament resistance, (*b*) Circuit diagram for measurement of work function.

(*ii*) ***Calibration of potentiometer wire*** : Make the electrical circuit as shown in fig. 17 (*b*). Close K_3 (*i*) and K_1 and place the Jockey on a length of 101·8 cm. of the potentiometer wire. Adjust the rheostat such that there is no deflection in the galvanometer *G*. In this way potential gradient is 10^{-3} volts per cm.

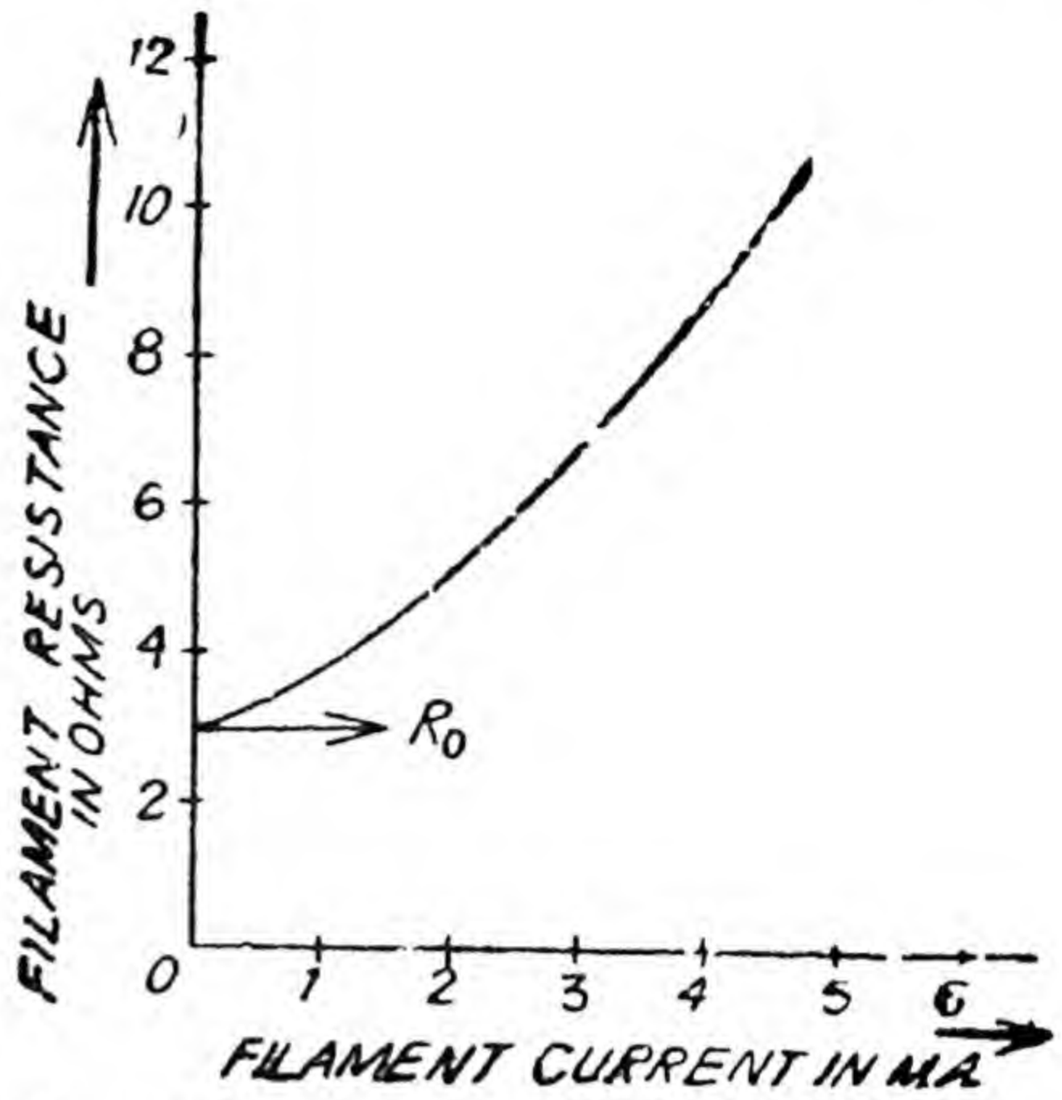


Fig. 18. Calculation of filament resistance at zero filament current.

(*iii*) ***Determination of saturated current*** : Open K_3 (*i*) and close K_3 (*ii*). Apply 160 to 1·0 volts to the plate of the valve. Adjust a low value of the filament current by means of circuit rheostat R_2. Obtain the condition of saturated plate current. Note down the filament current and saturated plate current. Note down the voltage across the filament by means of potentiometer *i.e.*, obtain the balancing length and multiply it by potential gradient which gives the filament voltage. Repeat the whole procedure for different filament current to obtain the saturated plate current J_s. For each value of filament current, obtain the filament resistance $R\ (=V/I)$. The corresponding absolute temperature T is calculated from the relation

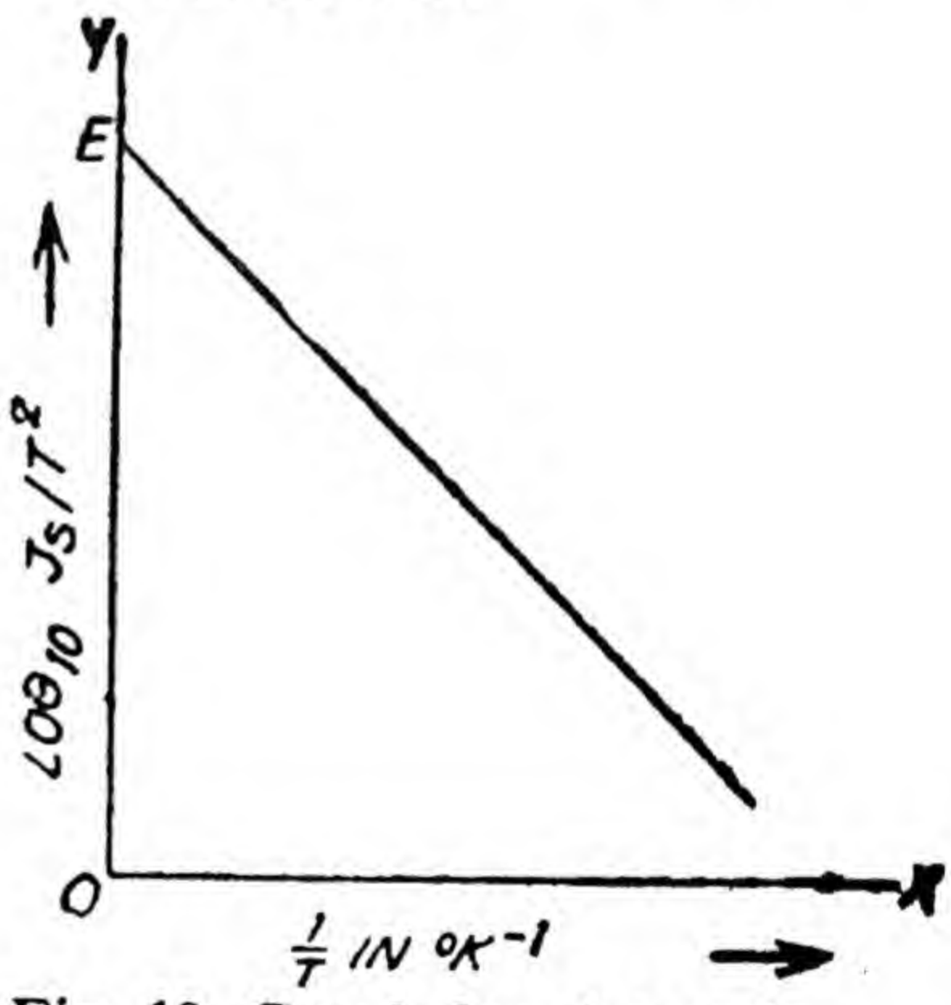


Fig. 19. Graph for determining A and ϕ.

$$\left(\frac{R}{R_0}\right)=\left(\frac{T}{T_0}\right)^{1\cdot2}$$

Now obtain the value of $\frac{1}{T}$ and $\log_e \frac{J_S}{T^2}$. Draw a graph between $\frac{1}{T}$ and $\log_e \frac{J_S}{T^2}$. The graph (fig. 19) is a straight line which proves the validity of Richardson law.

(*iv*) ***Calculation of work function*** :

The value of work function can be obtained by the slope of the graph (19).

$$\text{The work function } \phi=\frac{\text{Slope}}{5040}$$

NUMERICAL EXAMPLES

1. ***(i) Calculate the velocity of the emerging photo electrons when a light of wavelength 4000Å falls upon a surface of potassium with work function 2·5 eV.***

(ii) What is the threshold wavelength in the above case?

(*i*) $\lambda = 4000\text{Å}$ ∴ Frequency of radiation, $\nu = \frac{c}{\lambda} = \frac{3\times10^{8}}{4000\times10^{-10}}$

$= 7{\cdot}5\times10^{14}$ *c/s* or *Hz*.

From Einstein's equation

$$h\nu = e\phi + \tfrac{1}{2}mv^{2}$$

where h = Planck's constant $= 6{\cdot}62\times10^{-34}$ joules. sec.

m = mass of the electron $= 9{\cdot}11\times10^{-31}$ kg.

v = velocity of photo-electrons in *m*/sec.

e = charge of electron $= 1{\cdot}60\times10^{-19}$ coulombs.

Now $\tfrac{1}{2}mv^{2} = h\nu - e\phi$

$= 6{\cdot}62\times10^{-34}\times7{\cdot}5\times10^{14} - 1{\cdot}60\times10^{-19}\times2{\cdot}5$

$= 49{\cdot}6\times10^{-20} - 4\times10^{-19}$

$= (4{\cdot}96-4)\times10^{-19}$

$$\therefore\quad v^{2} = \frac{2\,(4{\cdot}96-4)\times10^{-19}}{9{\cdot}11\times10^{-31}} = \frac{2\times0{\cdot}96}{9{\cdot}11}\times10^{12}$$

$$v = \sqrt{\left(\frac{1{\cdot}92}{9{\cdot}11}\right)}\times10^{6} = 0{\cdot}446\times10^{6}\ \text{m/sec}.$$

(*ii*) The threshold frequency

$$\nu_0 = \frac{e\phi}{h}\ c/s.$$

∴ Threshold wavelength $\lambda_0 = \frac{c}{\nu_0} = \frac{ch}{e\phi}$ meters

$$= \frac{3\times10^{8}\times6{\cdot}62\times10^{-34}}{1{\cdot}60\times10^{-19}\times2{\cdot}5} = \frac{12{\cdot}4\times10^{-7}}{2{\cdot}5}\ \text{metre}$$

$$= \frac{12{\cdot}4\times10^{-7}}{2{\cdot}5\times10^{-10}}\ \text{Å} = 4960\text{Å}.$$

2. ***A cylindrical oxide-coated emitter is 2·5 cm long and 0·1 cm in radius. The emission constant $A_0 = 0{\cdot}01\times10^{4}$ amp/m^2K^2 and $\phi_0/k = 11960K$. Find the current at temperature 1000K.***

According to Richardson's equation, the emission current density is given by

$$J = A_0T^{2}\ \text{exp.}\ (-\phi_0/kT)$$

Here $A_0 = 0{\cdot}01\times10^{4}$ amp./m^2K^2, $\phi_0/k = 11960K$

and $T = 1000K$

∴ $J = (0{\cdot}01\times10^{4})\ (1000)^{2}\ \text{exp.}\ (-11960/1000)$

$= 10^{6}\ \text{exp.}\ (-11{\cdot}96)\ \text{amp}/m^2$

Taking log of both sides, we have

$\log J = 8\log 10 - 11{\cdot}96\log e$

$= 8 - 11{\cdot}96\times(0{\cdot}4343)$ (∵ $\log e = 0{\cdot}4343$)

$=8-5\cdot1942=2\cdot8058$

$\therefore \quad J=$ Antilog $2\cdot8058=639\cdot5$ amp/m^2

Surface area of emitter

$2\pi rl=2\times3\cdot14\times(0\cdot1\text{ cm})\times(2\cdot5\text{ cm})$

$=1\cdot57\text{ cm}^2=1\cdot57\times10^{-4}\ m^2$

$\therefore \quad i=$ current density $J\times$ Surface area

$=(639\cdot5\text{ amp}/m^2)\times(1\cdot57\times10^{-4}\ m^2)$

$=0\cdot1004$ amp. $=$ **100·4 milliampere.**

3. *The thermionic emission density from an emitter is $0\cdot29\times10^4$ amp/m^2 at temperature of 2500K. The work function for the emitter is 4·52 eV. Calculate the emission constant A_0 for the emitter. Given $e/k=11600$.*

According to Richardson's equation

$$J=A_0T^2\text{ exp. }(-\phi_0/kT)$$

Here $J=0\cdot29\times10^4$ amp/m^2, $\phi=4\cdot52\ eV=4\cdot52\times1\cdot6\times10^{-19}$ joule,

$T=2500K$ and $k=e/11600=1\cdot6\times10^{-19}/11600$ joule/K

Substituting these values, we get

$0\cdot29\times10^4$ amp/$m^2=A_0\ (2500K)^2$ exp $(-4\cdot52\times11600/2500)$

$$A_0=\frac{0\cdot29}{625}\text{ exp. }(20\cdot97)\text{ amp}/m^2K^2$$

Taking log of both sides

$\log A_0=\log 0\cdot29-\log 625+20\cdot97\log e$

$=\bar{1}\cdot4624-2\cdot7959+20\cdot97\times(0\cdot4343)$

$=-0\cdot5376-2\cdot7959+9\cdot1073$

$=5\cdot7738$

$\therefore \quad A_0=$ Antilog $5\cdot7738$

$=\mathbf{59\cdot40\times10^4\ amp}/m^2K^2.$

4. *There are two thermionic emitters identical in dimensions and operate at the same temperature of 2400 degree Kelvin. Emission constant A is the same for both. One emitter is tungsten while the other has a work function 0·6 times that of the first. Find the ratio of their emission currents. Assume suitable values of emission constant b for tungsten emitter.*

For tungsten,

$$I_{S_1}=AT^2\,e^{-b/T},$$

For the other emitter,

$$I_{S_2}=AT^2e^{-0\cdot6b/T}$$

since b is proportional to work function.

Hence, $$\frac{I_{S_2}}{I_{S_1}}=e^{-(0\cdot6-b)/T}=e^{0\cdot4b/T}.$$

For tungsten $b=52400$ degree Kelvin.

$$\text{Hence} \quad \log_e \frac{I_{S_2}}{I_{S_1}} = \frac{0{\cdot}4 \times 52400}{2400} = 8{\cdot}733$$

$$\therefore \quad \frac{I_{S_2}}{I_{S_1}} = 6194.$$

5. *A tungsten emitter operates at 2400 degree Kelvin. If the emission current density is to be reduced by 25%, what should the change in work fnuction be expressed in electron volts.*

Let I_{S_1} be the emission current density corresponding to constant b_1 in Richardson's equation, and I_{S_2}, the emission current density corresponding to constant b_2, then

$$I_{S_1} = AT^2 e^{-b_1/T}$$

and

$$I_{S_2} = AT^2 e^{-b_2/T}$$

or

$$\frac{I_{S_1}}{I_{S_2}} = \frac{e^{-b_1/T}}{e^{-b_2/T}} = e^{-b_1/T} \times e^{+b_2 T} = e^{(b_2 - b_1)/T}. \quad \ldots(1)$$

According to the problem

$$I_{S_2} = 0{\cdot}75\, I_{S_1},$$

$$\therefore \quad \frac{I_{S_2}}{I_{S_1}} = 0{\cdot}75 \text{ or } \frac{I_{S_1}}{I_{S_2}} = \frac{1}{0{\cdot}75} \quad \ldots(2)$$

From equations (1) and (2),

$$e^{(b_2 - b_1)/T} = \frac{1}{0{\cdot}75} = 1{\cdot}33.$$

or

$$\frac{(b_2 - b_1)}{T} = \log_e (1{\cdot}33). \quad \ldots(3)$$

We know that the constant b is related to the work function by the relation

$$b = \frac{E_a - E_f}{k}.$$

If ϕ is the voltage equivalent of work function, then

$$b = \frac{E_a - E_f}{k} = \frac{e\phi}{k} \text{ degree Kelvin.}$$

where e = charge on an electron in coulombs,
k = Boltzmann constant (Joules/degree Kelvin),
ϕ = voltage equivalent of work function of metal.

Substituting these values, we have

$$b = 11{,}605\phi \text{ degrees Kelvin,}$$

$$\therefore \quad b_1 = 11605\phi_1$$

and

$$b_2 = 11605\phi_2.$$

From equations (3) and (4),

$$\frac{11605\,(\phi_2 - \phi_1)}{T} = \log_e 1{\cdot}33$$

or

$$(\phi_2 - \phi_1) = \frac{T}{11605} \log 1{\cdot}33 = \frac{2400}{11605} \times 2{\cdot}3 \log_{10} 1{\cdot}33$$

$$= 0{\cdot}059 \text{ electron volt.}$$

6. *At what temperature will a pure tungsten filament gives an emission one fiftieth as much as that from a thoriated tungsten filament of same dimensions if their temperatures are the same? The value of constant A is 60 and 3 amperes cm.² per degree K² respectively, and value of b is 52,000 and 31,000 deg. K. respectively.*

For tungsten filament $I_{S_1} = AT^2 e^{-b_1/T} = 60T^2 e^{-52000/T}$

For thoriated tungsten $I_{S_2} = AT^2 e^{-b_2/T} = 3T^2 e^{-31000/T}$.

$$\therefore \quad \frac{I_{S_2}}{I_{S_1}} = \frac{3T^2 e^{-31000/T}}{60T^2 e^{-52000/T}}$$

$$= \frac{1}{20} e^{52000/T} \cdot e^{-31000/T}$$

$$= \frac{1}{20} e^{21000/T}.$$

According to the given problem $\frac{I_{S_2}}{I_{S_1}} = 50$.

$$\therefore \quad 50 = = \frac{1}{20} e^{21000/T},$$

$$1000 = e^{21000/T}$$

or $$\frac{21000}{T} = \log_e 1000 = 2{\cdot}303 \log_{10} 1000$$

or $$\frac{21000}{T} = 6{\cdot}909$$

or $$T = \frac{21000}{6{\cdot}909}$$

$$= \mathbf{3043°K}.$$

7. *Under space-charge limitea condition, the plate current of diode valve is 100 milliamperes for a plate voltage of 80 volts. Find the plate current for plate voltage of 150 and 200 volts under the space-charge limited condition.*

Under space charge limited condition,

$$I_P = KV_a^{3/2},$$

$$100 = K\ (80)^{3/2}. \qquad \ldots(1)$$

When $V_a = 150$ volts, then

$$I_{p_1} = K\ (150)^{3/2}. \qquad \ldots(2)$$

Dividing equation (2) by equation (1)

$$\frac{I_{P_1}}{100} = \left(\frac{150}{80}\right)^{3/2},$$

$$I_{p_1} = 100 \left(\frac{15}{8}\right)^{3/2}$$

$$= 100 \times 2{\cdot}58$$

$$= \mathbf{258\ milliamperes.}$$

For $V_a = 200$ volts.

$$I_{P_2} = K\ (200)^{3/2}. \qquad \ldots(3)$$

For equations (3) and (1), we have

$$\frac{I_{p_2}}{100} = \left(\frac{200}{80}\right)^{3/2},$$

$$I_{12} = 100 \left(\frac{200}{80}\right)^{3/2} = 100 \times 3{\cdot}95$$

$$= \mathbf{395\ milliamperes}.$$

8. *In a vacuum diode, the plate is a circular cylinder of length 12 cm. and diameter 2·0 cm. At the axis of the cylindrical plane is the tungsten filament at a temperature of 2400 degree Kelvin and it is 12 cm. long and has a diameter of 0·04 cm. Find.*

(i) saturation current,

(ii) plate voltage required to produce a space charge limited current equal to one third of the saturation current.

Given that β^2 *for ratio of plate to cathode diameter of 50 is 1·09*

$$J_S = AT^2\ e^{-b/T}$$

$$= 60{\cdot}2 \times 10^4 \times (2400)^2 \times e^{-62400/2400}$$

$$= \mathbf{1096\ amperes/sq.\ metre.}$$

Cathode surface $= 2\pi r_k \cdot l$

$$= 2\pi \times 0{\cdot}02 \times 10^{-2} \times 12 \times 10^{-2} \text{ square metre}$$

$$= 0{\cdot}48\pi 10^{-4} \text{ square metres.}$$

Therefore emission current

$$= 1096 \times 0{\cdot}48\pi \times 10^{-4} \text{ amperes}$$

$$= \mathbf{0{\cdot}1652\ ampere.}$$

The space-charge limited current required is

$$= \tfrac{1}{3} \times \mathbf{1652\ amperes.}$$

$$J_S = 0{\cdot}055 \text{ ampere.}$$

But $$J_S = 14{\cdot}68.10^{-6}.\ \frac{1}{r_a}.\ \frac{V_a^{3/2}}{\beta^2},$$

$l = 12$ cm., $\beta^2 = 1{\cdot}09$, since $\frac{r_a}{r_k} = \frac{1{\cdot}0}{0{\cdot}02} = 50.$

Putting the values,

$$0{\cdot}055 = 14{\cdot}68.10^{-6}\ \frac{12}{10} = \frac{V_a^{3/2}}{1{\cdot}09}.$$

$$\therefore \quad V_a^{3/2} = \frac{1{\cdot}09 \times 0{\cdot}055}{14{\cdot}68.10^{-6}.12} = 0{\cdot}309.10^3$$

$$= 309,$$

$$\mathbf{V_a = 45{\cdot}7\ volts.}$$

9. *In a high vacuum diode, the oxide coated thermionic emitter operates at 1000 degree Kelvin. The values of constant A and B in Richardson equation are :* $A = 0{\cdot}01 \times 10^4$ *amps./sq. metre/deg.* K^2 *and* $b = 1200$ *deg. K. Calculate the emission current density. If the emitter is in the form of a hollow cylinder of length 5 cm. and external diameter 0·2 cm., find the total emission current. If the anode voltage is 360 volts and the distance between anode and cathode is 0·4 cm., find (i) the percentage increase in emission current consider-*

ing the Schottky effect due to the electric field of anode, and (ii) the value of emission current considering the Schottky effect.

According to Richardson equation

$$J_s = AT^2 e^{-b/T}$$

$$= 0{\cdot}01 \times 10^4 \times (1000)^2 \text{ exp. } (-1200/1000)$$

$$= 630 \text{ amp./sq. m.}$$

Emitting surface $= (5 \times 10^{-2})\ \pi\ (0{\cdot}2 \times 10^{-2})$

$$= \pi \times 10^{-4} \text{ sq. m.}$$

$\therefore$ Total emission current in absence of electric field

$$I_s = 630 \times \pi \times 10^{-4} = 197{\cdot}8\ mA.$$

$$\text{Electric field strength at anode} = \frac{360 \text{ volt}}{0{\cdot}4 \times 10^{-2}}$$

$$= 9 \times 10^4 \text{ volts/m.}$$

Let I be the emission current considering the Schottky effect caused by the electric field of the anode voltage, then

$$I/I_s = K = e^{0{\cdot}44\sqrt{\xi}/T}$$

or
$$\log_{10} K = \frac{0{\cdot}44\sqrt{\xi}}{T} \log_{10} e$$

$$= \frac{0{\cdot}44 \times \sqrt{\cdot 9 \times 10^4)}}{1000} \cdot \frac{1}{2{\cdot}303} = 0{\cdot}0573$$

$\therefore$ $K = 1{\cdot}41$.

(*i*) The percentage increase in $I_s = \left(\frac{1{\cdot}41 - 1}{1}\right) \times 100$

$$= 41\%.$$

(*ii*) Emission current

$$I = 197{\cdot}8 \times 1{\cdot}41 = 225{\cdot}7 \text{ mA.}$$

EXERCISES AND PROBLEMS

1. What do you mean by electron emission ? Give the different categories of the electron emission.
2. What do you understand by bound and free electrons in the crystal lattice ?
3. What is work function ?
   Two thermionic emitters identical in dimensions operate at the same temperature of 2400 degree Kelvin. Emission constant A is the same for both. One is tungsten while the other has a work function constants 0·6 times that of the first. Find the ratio of their emission currents. Assume value of b for tungsten as 52400 deg. K.
4. Give thermodynamical derivation of the Richardson equation. What are the drawbacks of the derivation ?
5. Give the mathematical derivation of the Richardson equation by Fermi-Dirac Statistics.

6. What are the different types of emitters ? Give the advantages of indirectly heated emitters over directly heated emitters.
7. What do you mean by space-charge and temperature limited currents ?
8. For a plane parallel electrode, under space-charge limited condition derive the equation relating the anode current with anode voltage.
9. Describe Child-Langmuir's equation for space charge limited current. Discuss its applications to vacuum tubes.
10. Describe an experiment to determine the work function of the emitter (pure metal) of a vacuum tube.
11. At 2400°K a tungsten filaments is observed to increase its emission periodically by amounts upto 5%. What change in value of work function is required to explain this change ?
12. A flash of light from the ultraviolet line of mercury at 2537Å, lasting 0·1 second strikes a sodium surface of 1·5 cm^2. in photo tube. The surface has a work function of 1·9V, aud the light has a power density of 0·5 W/cm^2. If half the photons cause the emission of an electron, find the current during the flash.
13. Two diodes each have an anode 4 millimeters, in diameter and 2 centimeters long, but one of them has a filament 0·1 millimetre in diameter, while the other has an indirectly heated cathode 1 5 millimeters in diameter. Compare the space currents flowing in each when anode potential is 25 volts. (Neglect edge effects).
14. Find the ratio of emission currents for following emitters :
    (*a*) Tungsten emitter ; emitting surface=2 sq. cm. ; $A=60{\cdot}2\times10^4$ amperes/sq. metre/degree K^2 ; $b=52400$ deg. K and temperature $T=26000$ deg. K.
    (*b*) Thoriated tungsten emitting surface=3 sq. cm. ; $A=3\times10^4$ amperes/sq. metre/deg. K^2, $b=30000$ degree K. Temperature $T=1800$ degree K.
15. A vacuum diode has 0·8 inch diameter tungsten filament of 3 inches length. The temperature of the filament is adjusted to 2400 degree K. Assuming typical values of constants in Richardson's equation, calculate the emission current. Space charge limited current is 100 mA. with a plate voltage of 150 volts. Find the plate current with plate voltage of 300 and 600 volts and state whether these currents are temperature limited or space-charge limited. Find the values of current for the above three plate voltages if the temperature of filament is raised to 2450 degree K.
16. What are the different ways by which electrons can be ejected from metals.
17. Mark the correct answers :
    (*a*) The Richardson's equation is

$$(i)\quad (I_S)_f=I_S e^{0{\cdot}44\sqrt{\xi}/T},\quad (ii)\ I_f=A_f\xi^2 e^{-b_f/\xi}$$

(*iii*) $J_S = AT^2 e^{-b/T}$, (*iv*) $hf = \phi Q_e + \frac{1}{2}mv^2$.

(*b*) Field emission current density equation and Richardson's equation for thermionic current density are :
(*i*) Similar, (*ii*) Not similar.

(*c*) The Schottky effect is applicable to
(*i*) Thermionic emission, (*ii*) Secondary emission,
(*iii*) Photoelectric emission, (*iv*) Field emission.

(*d*) Operating temperature of oxide coated emitter in degree Kelvin is
(*i*) 1000, (*ii*) 1800, (*iii*), 2000, (*iv*) 2400.

[Ans. (*a*) (iii), (*b*) (i), (*c*) (iv) (*d*) (i)]

8 RECTIFIERS AND POWER SUPPLY

Alternating voltage is preferred for transmission and distribution means because of small losses and convenience. But for certain purposes, direct voltages are more useful. For such applications, we have to convert alternating voltage into direct voltage which is generally done by means of simple and efficient device named as electronic rectifier. Some other types of rectifiers may also be used.

By definition, rectifier is a device which converts alternating voltage or current into direct voltage or current due to a certain type of non-linearity of its characteristic.

We shall discuss some of the following two classes of rectifiers :

Case I. Vacuum tube and semi-conductor diode Rectifiers :

(A) Half wave vacuum tube *and semi-conductor diode* rectifiers,

(B) Full wave vacuum tube *and semi-conductor diode* rectifiers.

Case II. Gas Diode Rectifiers :

(A) Hot cathode gas diode rectifier,

(B) Cold cathode gas diode rectifier.

8·1. VACUUM TUBE AND SEMI-CONDUCTOR DIODE RECTIFIERS :

A half wave rectifier conducts during one half cycle of the applied voltage while in full wave rectifier, the rectification is in operation for the full wave of the applied a.c. voltage.

8·1-1. HALF WAVE RECTIFIER (Vacuum diode) :

A basic circuit of half wave rectifier is shown in figure 1 (*a*).

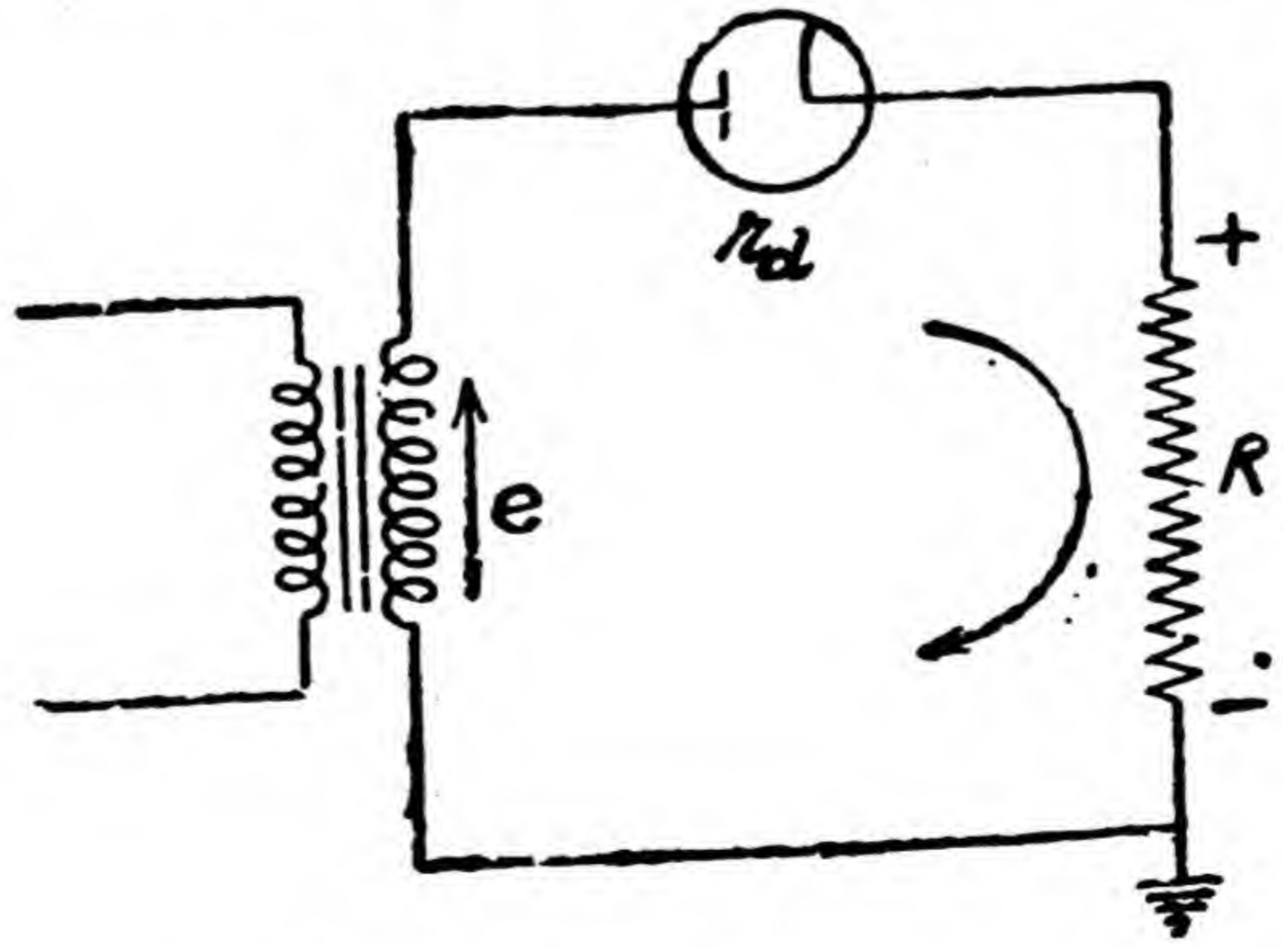


Fig. 1 (a) Basic half wave circuit.

If a curve relating the plate current and the applied voltage, known as static curve, is drawn then it is found to be non-linear.

The diode resistance r_a is measured as the reciprocal of the slope of this curve and, therefore, diode resistance is not constant.

When we insert some load resistance R in the circuit and then if we plot a curve in applied voltage and plate current it is found to be more linear than the previous case. The curve is known as a dynamic curve. Since slope of dynamic curve is less than static curve and the relation in applied voltage and current, in this case, depends upon the linear load resistance, the dynamic curve is more straight than static curve.

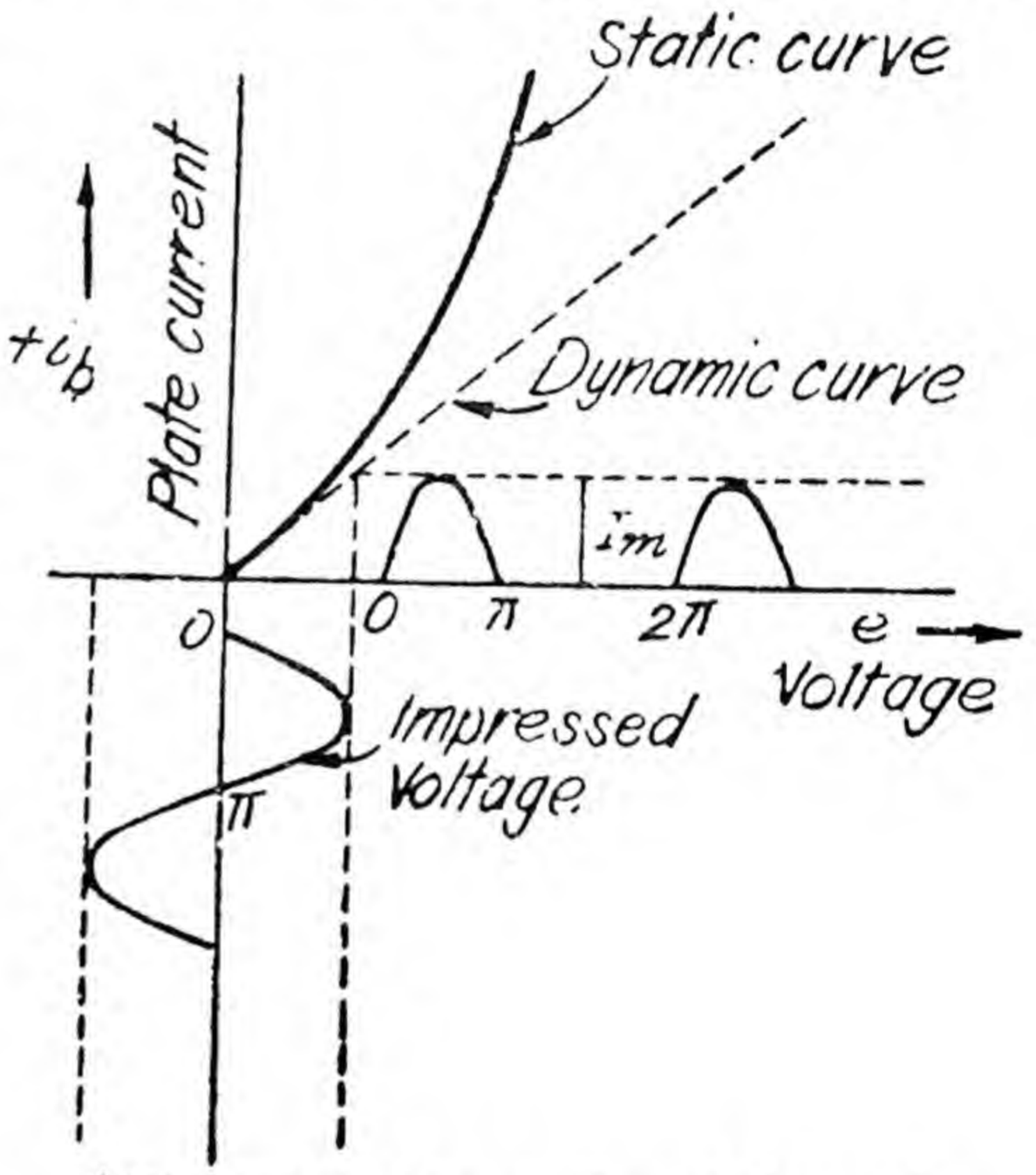


Fig. 1 (b) The static and dynamic characteristic of a rectifier.

The basic principle of rectifiction is as follows :

During the positive half cycle of a.c. *i.e.*, when plate is positive with a respect to cathode, the tube conducts. Electrons flow from cathode to plate and consequently the conventional current i_b flows in the load. The voltage across the load is $i_b.R$ where R is load resistance. During the negative half cycle of a.c. *i.e.* when plate is negative with respect to cathode, the electrons are not attracted towards the plate and hence the current i_b is zero. The voltage developed across the load is also zero. In this way the voltage developed across the load is either positive or zero with respect to lower end. Thus the out-put voltage is a undirectional or pulsating d.c. voltage.

Let i_b = instantaneous value of diode current

I_m = maximum amplitude of the output current

r_a = internal resistance of diode

R = load resistance

$I_{d.c.}$ = average value of the current recorded by d.c. meter placed in series with load R.

$I_{r.m.s.}$ = R.M.S. value of the current recorded by a.c. meter placed in series with load R.

and $e = E_m \sin \omega t$ = input voltage. ...(1)

Assuming the static curve to be linear, we can write

$$i_b = \frac{E_m}{r_a + R} \sin \omega t = I_m \sin \omega t, \quad 0 \leqslant \omega t \leqslant \pi.$$

$$i_b = 0, \qquad \pi \leqslant \omega t \leqslant 2\pi . \qquad \ldots(2)$$

where $$I_m = \frac{E_m}{r_a + R}.$$

D. C. Value of Current : The fuuction of a rectifier is to produce direct current. The average value of a.c. is expressed by that steady current which transfers, across any circuit the same charge as is transferred by that a.c. In symmetrical a.c. (two half cycles are exactly similar), the average value over a complete cycle is zero. The average value is obtained by adding or integrating the instantaneous values of the current over half cycles. In unsymmetrical a.c., the average value is taken over the whole cycle. The average value of d.c. component is given by

$$I_{d.c.} = \frac{1}{2\pi}\int_0^{2\pi} i_b\, d(\omega t) = \frac{1}{2\pi}\left[\int_0^{\pi} I_m \sin \omega t\, d(\omega t) + \int_{\pi}^{2\pi} 0.d(\omega t)\right]$$

$$= \frac{I_m}{2}\Big(-\cos \omega t\Big)_0^{\pi} = \frac{I_m}{\pi}, \text{ where } I_m = \frac{E_m}{r_a + R},$$

$$\therefore \quad I_{d.c.} = \frac{E_m}{\pi(r_d + R)}. \qquad \ldots(3)$$

R.M.S. Value of Current :

$$I_{r.m.s.} = \sqrt{\left\{\frac{1}{2\pi}\int_0^{2\pi} i_b^2\, d(\omega t)\right\}}$$

$$= \sqrt{\left\{\frac{1}{2\pi}\int_0^{2\pi} I_m^2 \sin^2 \omega t\, d(\omega t)\right\}}$$

$$= \sqrt{\left\{\frac{1}{2\pi}\int_0^{\pi} \frac{E_m^2 \sin^2 \omega t}{(r_a + R)^2}\, d(\omega t) + \frac{1}{2\pi}\int_{\pi}^{2\pi} 0.d(\omega t)\right\}}$$

$$= \sqrt{\left\{\frac{E_m^2}{2\pi(r_a + R)^2}\int_0^{\pi} \sin^2 \omega t\, d(\omega t)\right\}}$$

$$= \sqrt{\left\{\frac{E_m^2}{2\pi(r_a + R)^2}\int_0^{\pi} \frac{(1 - \cos 2\omega t)}{2}\, d(\omega t)\right\}}$$

$$= \sqrt{\left[\frac{E_m^2}{2\pi(r_a + R)^2}\left\{\frac{\omega t}{2} - \frac{\sin 2\omega t}{4}\right\}_0^{\pi}\right]}$$

$$= \sqrt{\left\{\frac{E_m^2}{2\pi\,(r_a + R)^2}\left(\frac{\pi}{2}\right)\right\}}$$

$$= \sqrt{\left\{\frac{E_m^2}{4\,(r_a + R)^2}\right\}}$$

$$= \frac{E_m}{2\,(r_a + R)}. \qquad \ldots(4)$$

Thus $$I_{r.m.s.} > I_{d.c.}$$

The maximum voltage across the load is given by

$$E_{max.} = I_m.R = \frac{E_m}{r_a + R}.\ R = \frac{E_m}{(1 + r_a/R)}. \qquad \ldots(5)$$

The d.c. voltage across the load is given by

$$E_{d.c.} = I_{d.c.} \quad R = \frac{E_m}{\pi\,(1 + r_d/R)} \qquad \ldots(6)$$

Frequency Components of Rectified Output : Fourier analysis of voltage pulses of a half wave rectifier gives the following frequency components :

$$e = E_m \left[\frac{1}{\pi} + \frac{1}{2}\sin \omega t - \frac{2}{3\pi}\cos 2\omega t - \frac{2}{15\pi}\cos 4\omega t - \ldots\right] \qquad \ldots(7)$$

The first term on right hand side of above relations (7) gives $_{d.c.}$ while the remaining terms are a.c. harmonics.

Efficiency of Rectifier : Efficiency of rectification is stated as

$$\eta = \frac{\text{Direct power output } (P_{d.c.})}{\text{Alternating input power } (P_{a.c.})} \times 100\%,$$

$$P_{d.c.} = I_{d.c.}^2 . R = \frac{E_m^2}{(r_d + R)^2} \cdot \frac{R}{\pi^2} . \qquad \ldots(8)$$

The input a.c. power is supplied by the transformer.
The power supplied to the circuit from a.c. source

$$P_{a.c.} = \frac{1}{2\pi}\int_0^{2\pi} e.i_b.d\,(\omega t)$$

$$= \frac{1}{2\pi}\int_0^{2\pi} i_b\,(r_d + R).i_b\,d(\omega t), \text{ since } e = i_b\,(r_d + R)$$

$$= \frac{1}{2\pi}\,(r_d + R)\int_0^{2\pi} i_b^2\,d(\omega t)$$

Since $I^2_{r.m.s.} = \frac{1}{2\pi}\int_0^{2\pi} i_b^2\,d(\omega t)$,

$$P_{a.c.} = (r_d + R)\,I^2_{r.m.s.}$$

$$= (r_d + R).\left[\frac{E_m}{2(r_d + R)}\right]^2 = \frac{E_m^2}{4(r_d + R)} \qquad \ldots(9)$$

$$\therefore \quad \text{Efficiency } \eta = \frac{E_m^2/(r_d + R)^2 \times R/\pi^2}{E_m^2/4\,(r_d + R)}\,.100\%$$

$$= \frac{4R \times 100}{\pi^2\,(R + r_d)} = \frac{40{\cdot}6}{\left(1 + \frac{r_d}{R}\right)}\%. \qquad \ldots(10)$$

Ripple Factor : Ripple factor γ is defined as

$$\gamma = \frac{\text{r m.s. value of a.c. component of wave}}{\text{average or d.c. value of wave}} \qquad \ldots(11)$$

R.m.s. value and d.c. value may be of voltage or current, *i.e.*

$$\gamma = \frac{I'_{r.m.s.}}{I_{d.c.}} = \frac{E'_{r.m.s.}}{E_{d.c.}}$$

where $I'_{r.m.s.}$ and $E'_{r.m.s.}$ denote the r.m.s. value of a.c. components only.

The instantaneous a.c. component of the current is given by

$$i' = i - I_{d.c.} \qquad \ldots(12)$$

From above, we can calculate the r.m.s. value of the current

$$I_{r.m.s.}=\sqrt{\left\{\frac{1}{2\pi}\int_0^{2\pi}(i')^2\,d(\omega t)\right\}}$$

$$=\sqrt{\left\{\frac{1}{2\pi}\int_0^{2\pi}(i^2-2iI_{d.c.}+I^2_{d.c.})\,d(\omega t)\right\}}.$$

where $i'=i-I_{d.c.}$

The integrals $\frac{1}{2\pi}\int_0^{2\pi} I^2_{d.c}\;d(\omega t)=I^2_{d.c.}$

$$\frac{1}{2\pi}\int_0^{2\pi} i^2\,d(\omega t)=I^2_{r.m.s.}$$

$$\frac{1}{2\pi}\int_0^{2\pi} 2i\,I_{d.c.}\,d(\omega t)=2I^2_{d.c.}$$

Therefore $I'_{r.m.s.}=\sqrt{(I^2_{r.m.s.}-2I^2_{d.c.}+I^2_{d.c.})}$ ...(13)

$$=\sqrt{(I^2_{r.m.s.}-I^2_{d.c.})}.$$

Now we can write for the ripple factor as

$$\gamma=\sqrt{\left\{\frac{I^2_{r.m.s.}-I^2_{d.c.}}{I^2_{d.c.}}\right\}}$$

$$\gamma=\sqrt{\left\{\left(\frac{I_{r.m.s.}}{I_{d.c.}}\right)^2-1\right\}} \qquad \ldots(14)$$

But as calculated before, $I_{r.m.s.}=\frac{I_m}{2}$

$$I_{d.c.}=\frac{I_m}{\pi}$$

$$\therefore\quad \gamma=\sqrt{\left\{\left(\frac{\pi}{2}\right)^2-1\right\}}=\sqrt{\{(1{\cdot}57)^2-1\}}$$

$$=1{\cdot}21$$

Since $\gamma > 1$, it shows that :

R.m.s. value of a.c. component of wave is greater than d.c. value of wave or in other words the ripple potential r.m.s. value exceeds the d.c. potential of the output which shows that without filter, half wave rectifier is poor device for the conversion of a.c. voltage into d.c. voltage.

Peak Inverse Voltage :

The peak inverse voltage (*P.I.V.*) is defined as the maximum voltage appearing across the rectifier diode during the period in which the diode is not conducting. In case of half wave rectifier, the diode is not conducting for the negative half cycle of the applied voltage and hence the voltage across the tube is the same as the applied voltage. In this case the peak inverse voltage is the maximum value of *e*, *i.e.* E_m.

$$\therefore\quad P.I.V.=E_m.$$

Sometimes for comparing various rectifier circuits a ratio of *P.I.V.* to the average output voltage is specified. For half wave rectifier, this ratio is

$$= \frac{P.\,I.\,V.}{\text{d.c. or average output}}$$

$$= \frac{E_m}{E_m/\pi}$$

$$= \pi$$

Voltage regulation of half wave rectifier :

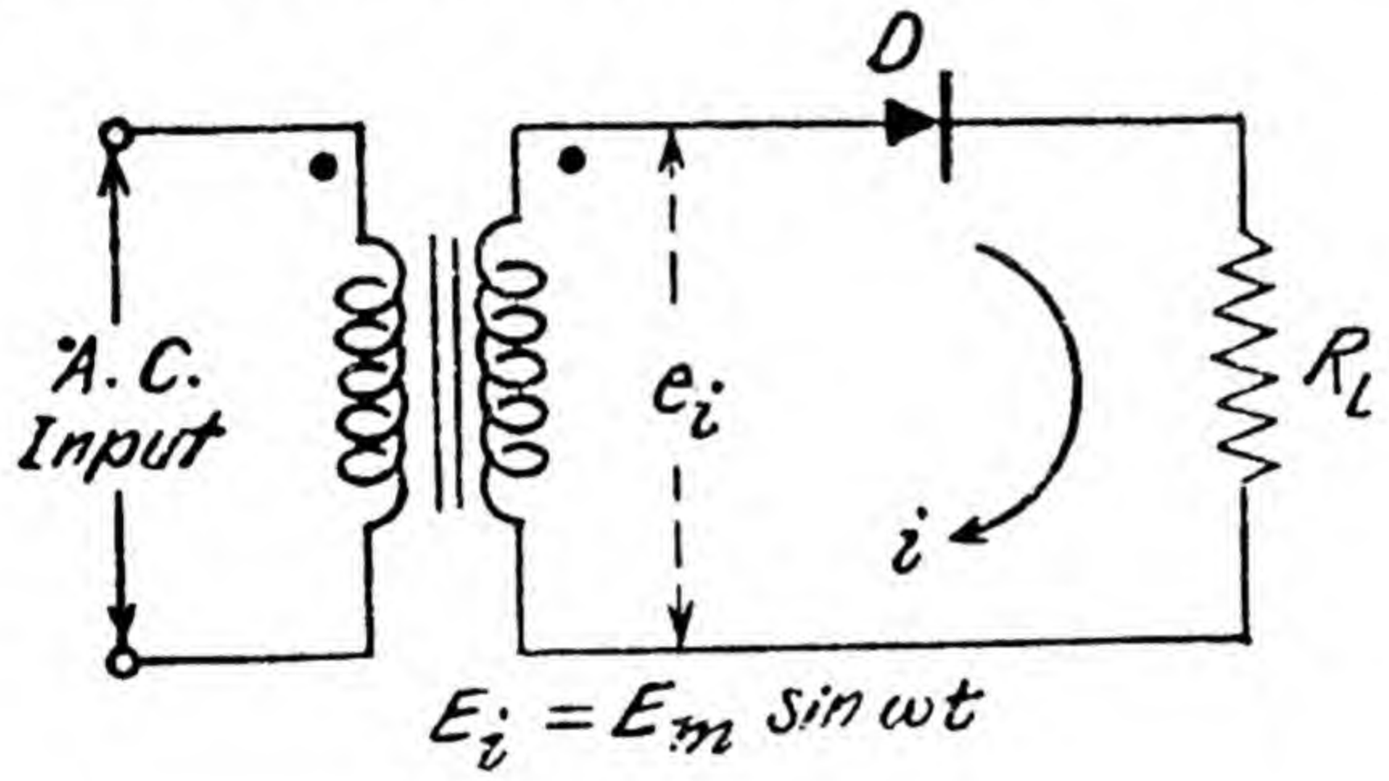
The voltage regulation is a measure of the ability of a rectifier to maintain a specified output voltage with the variation of load. and is defined as

$$\text{voltage regulation} = \frac{E_{no\ load} - E_{specified\ load}}{E_{specified\ load}},$$

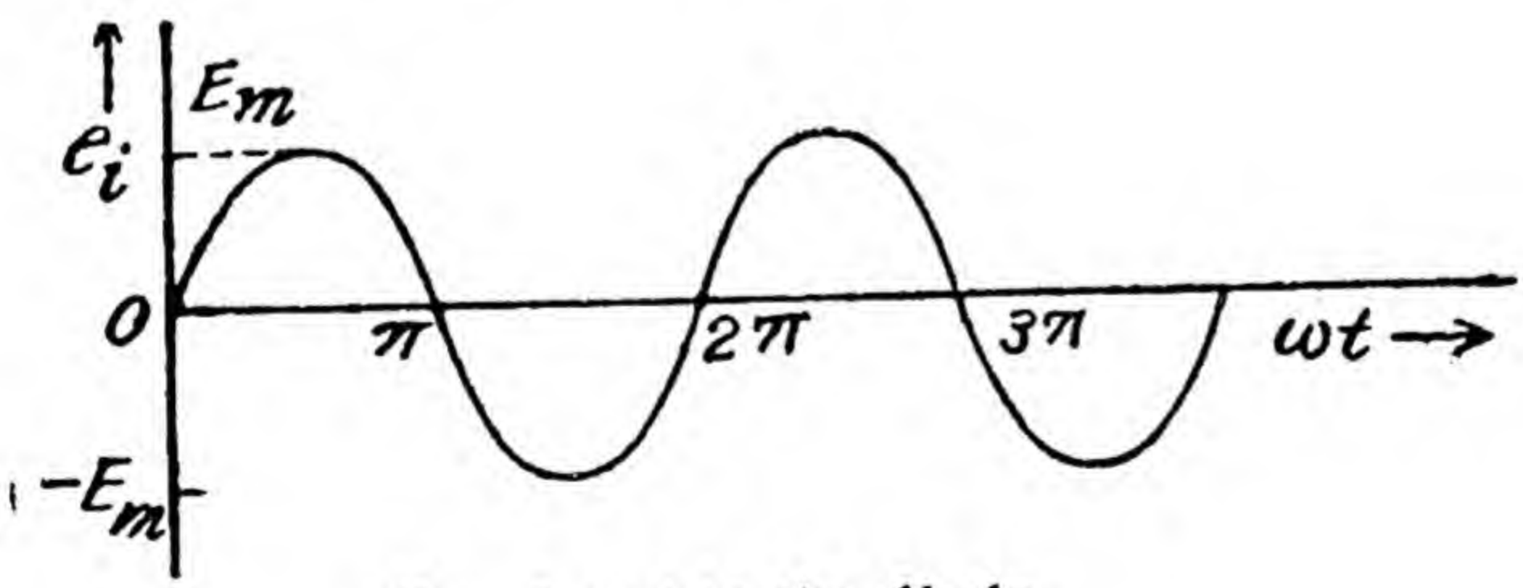
where E's refer to d.c. output voltage of the rectifier.

8·1-2. HALF WAVE RECTIFIER (Semiconductor Diode) :

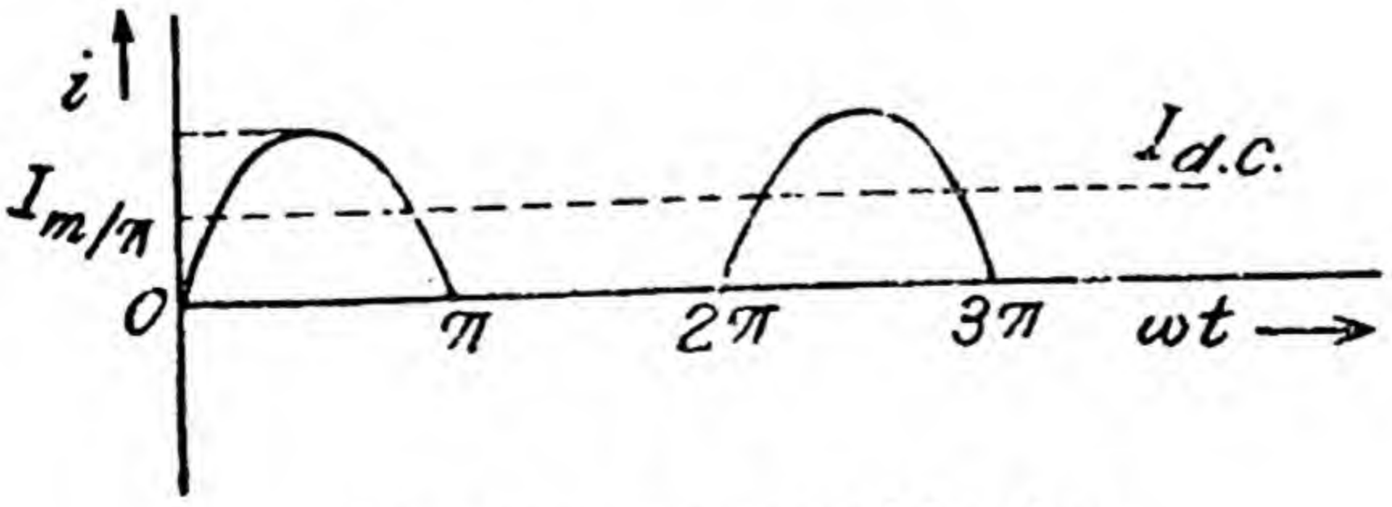
Fig. (2*a*) shows the basic circuit of a half wave rectifier. The a.c.



(a) Basic circuit



(b) Input to the diode



(c) Half wave rectification

Fig. (2) Half wave rectifier.

voltage which is to be rectified is applied to the input of a power transformer. The voltage developed across the secondary is available for rectification. So the transformer is used to provide the desired voltage into the rectifier. For high voltage d.c. power supplies, the transformer is obviously a step up transformer but for so many solid state equipment applications voltages of 4, 5, 6...50 volts etc. are used and hence stepdown transformer is used,

Let e_i be the input voltage to the rectifier, which is given by

$$e_i = E_m \sin \omega t \qquad ...(1)$$

The current i in the diode or through the load resistance R_L is given by

$$i = I_m \sin \omega t \text{ for } 0 \leqslant \omega t \leqslant \pi \qquad ...(2)$$

$$\text{and } i = 0 \text{ for } \pi \leqslant \omega t \leqslant 2\pi.$$

The peak current I_m is givin by

$$I_m = \frac{E_m}{R_f + R_L} \qquad ...(3)$$

where R_f = dynamic resistance of diode.

The current i flowing through load resistance R_L is unidirectional and its average value $I_{d.c.}$ is given by

$$I_{dc} = \frac{1}{2\pi}\int_0^{2\pi} i \, d\omega t \qquad ...(4)$$

As no current is flowing during the intervel π to 2π, hence

$$I_{dc} = \int_0^{\pi} I_m \sin \omega t \, d(\omega t) = \frac{I_m}{\pi} \qquad ...(5)$$

The value of I_{dc} can also be expressed as

$$I_{dc} = \frac{E_m}{\pi(R_f + R_L)} \qquad \because \; I_m = E_m/(R_f + R_L) \quad ...(6)$$

DC output voltage : The *dc* output voltage appearing across R_L will be

$$E_{dc} = I_{dc}\, R_L = \frac{I_m\, R_L}{\pi} = \frac{E_m}{\pi\{1 + (R_f/R_L)\}} \qquad ...(7)$$

RMS current and voltage : The *rms* value of current is given by

$$I_{rms} = \left[\frac{1}{2\pi}\int_0^{2\pi} i^2 \, d(\omega t)\right]^{1/2}$$

$$= \left[\frac{1}{2\pi}\int_0^{\pi} I_m^2 \sin^2 \omega t \, d(\omega t)\right]^{1/2} = \frac{I_m}{2}$$

$$= \frac{E_m}{2(R_f + R_L)} \qquad ...(8)$$

and $$E_{rms} = I_{rms}.\, R_L = \frac{E_m}{2\{1 + (R_f/R_L)\}} \qquad ...(9)$$

Rectifier efficiency—The rectifier efficiency is defined as the ratio of d.c. output power P_{dc} to the a.c. input power P_{in}. Hence

$$\eta = \frac{dc \text{ power delivered to the load}}{ac \text{ input power from transformer secondary}} = \frac{P_{dc}}{P_{ac}}$$

The useful power output is *d.c.* power P_{do} developed across the load, hence

$$P_{do} = I_{do}^2 R_L = \frac{I_m^2 R_L}{\pi^2} \qquad \ldots(10)$$

Power taken from a.c. source is used in two parts *i.e.*, (i) power dissipated at the junction of the diode and (ii) power dissipated in load resistance.

(i) power dissipated at junction of diode $= I_{rms}^2 R_f$

(ii) power dissipated at load resistance $= I^2_{rms} R_L$

$$\therefore \quad P_{ac} = I^2_{rms} (R_f + R_L) = \left(\frac{I_m}{2}\right)^2 (R_f + R_L)$$

Now $$\eta = \frac{I^2_m R_L/\pi^2}{I^2_m (R_f + R_L)/4} = \frac{4}{\pi^2} \cdot \frac{R_L}{(R_f + R_L)}$$

or $$\eta = \frac{0{\cdot}406}{1 + (R_f/R_L)} \qquad \ldots(11)$$

$$\text{Percentage rectifier efficiency} = \frac{40{\cdot}6}{1 + (R_f/R_L)} \qquad \ldots(12)$$

Theoretically the maximum value of rectifier efficiency of a half wave rectifier is 40·6 percent when $R_f/R_L = 0$.

Ripple factor. Ripple factor γ is defined as

$$\gamma = \frac{\text{ripple voltage}}{\text{d.c. value of wave}} = \frac{rms \text{ value of ac component}}{\text{dc value of wave}}$$

$$= \frac{E'_{rms}}{E_{dc}}$$

The ripple factor may also be defined in terms of ripple current and *dc* current as follows :

$$\gamma = \frac{I'_{rms}}{I_{dc}}$$

The current in half wave rectifier is given by

$$i \approx \frac{I_m}{\pi} + \frac{I_m}{2} \sin \omega t - \frac{2I_m}{3\pi} \cos 2\omega t - \frac{2I_m}{15\pi} \cos 4\omega t + \ldots$$

or $$i = I_{do} + i_r$$

where $I_{dc} = \frac{I_m}{\pi}$ and $i_r = \frac{I_m}{2} \sin \omega t - \frac{2I_m}{3\pi} \cos 2\omega t - \frac{2I_m}{15\pi} \cos 4\omega t$

Let I_{rms} be the value of current i, the total half wave rectified output current Similarly, let I'_{rms} be the *rms* value of ripple crurent i_r consisting of alternating components. Now

$$I_{rms}^2 = I_{do}^2 + I'^2_{rms}$$

or $$I'_{rms} = (I_{rms}^2 - I_{dc}^2)^{1/2}$$

Now $$\gamma = \frac{(I_{rms}^2 - I_{dc}^2)^{1/2}}{I_{do}} = \sqrt{\left\{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1\right\}}$$

$$= \sqrt{\left\{\left(\frac{I_m/2}{I_m/\pi}\right)^2 - 1\right\}} = \sqrt{\left(\frac{\pi^2}{4} - 1\right)} = 1{\cdot}21 \qquad \ldots(13)$$

This expression shows that the perentage of ac in the output is 121% of the dc voltage. Thus half wave rectifier is not very successful in converting ac into dc.

Regulation : By regulation we mean the variation of *dc* output voltage with change in *dc* load current. The percentage of regulation defined as

$$\text{Percentage regulation} = \frac{E_{no\ load} - E_{full\ load}}{E_{full\ load}} \times 100 \qquad \ldots(14)$$

For an ideal power supply, the percentage of regulation should be zero *i.e.*, the output voltage should be independent of load current.

For half wave rectifier, we have

$$E_{dc} = \frac{E_m R_L}{\pi(R_f + K_L)} = \frac{E_m}{\pi}\left[1 - \frac{R_f}{R_f + R_L}\right]$$

$$= \frac{E_m}{\pi} - I_{dc} R_f \qquad \ldots(15)$$

This equation suggests that a half wave rectifier functions as if it were a constant voltage source $E(=E_m/\pi)$ in series with internal resistance R_f. Thus E_{dc} is equal to E_m/π with no load ($I_{dc}=0$) and d.c. output voltage decreases linearly with the increase of d.c. output current I_{dc}.

8·1-3. FULL WAVE RECTIFIER (Vacuum diode) :

Fig. 3 (a) shows the circuit diagram of a full wave rectifier using vacuum diodes. There are two diodes and a transformer with centre tapped secondary. The two cathodes are joined together to form the positive of the d.c. output voltage across the load *R*. The negative

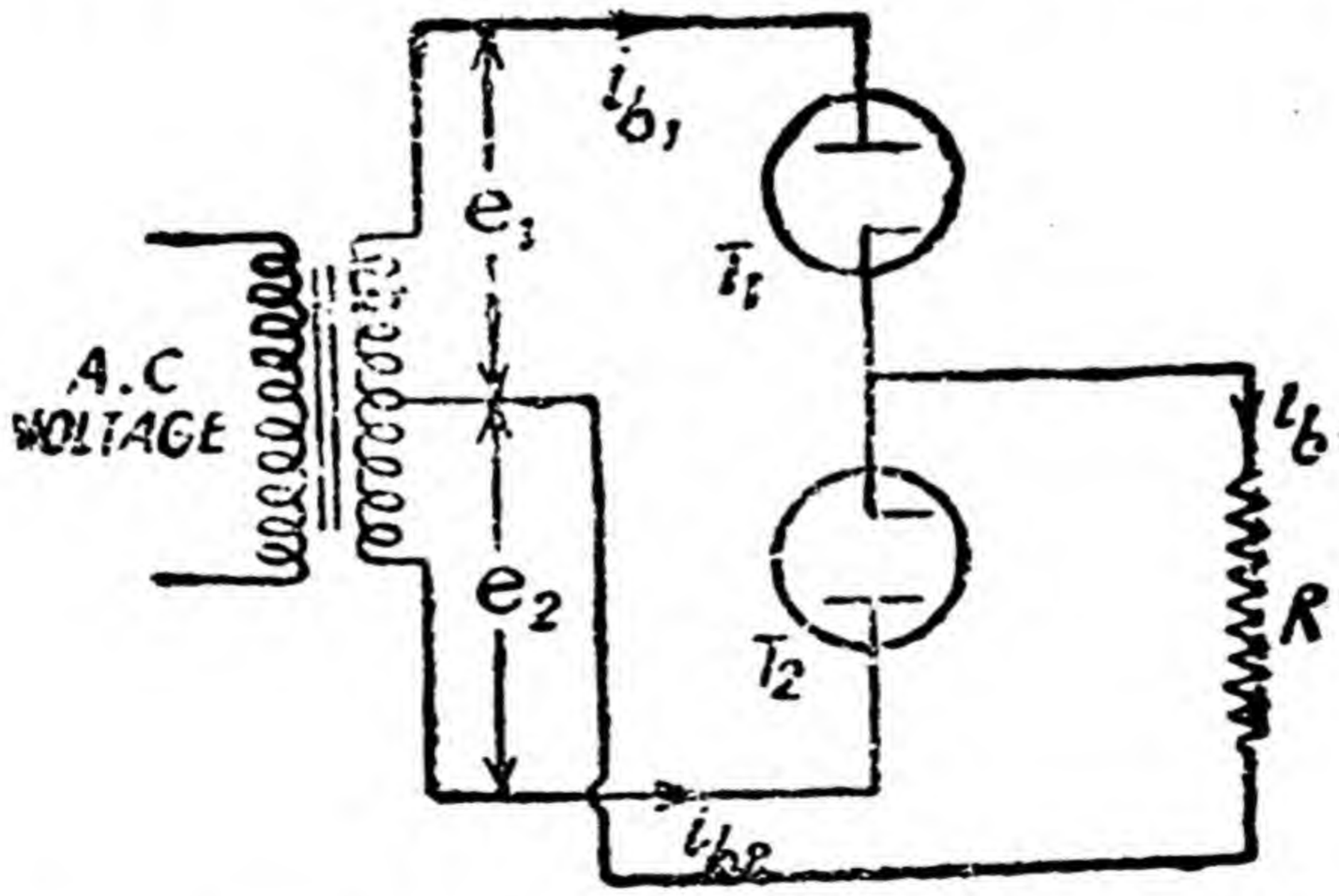


Fig. 3 (a) Circuit for full wave rectifier.

is connected to the centre tap of the secondary windings. Since the a.c. inputs of T_1 and T_2 are 180° apart, the tubes T_1 and T_2 conduct on alternate half cycles. The magnitude of e_1 is same as that of e_2. The total current through the load resistance will be equal to the sum of the individual currents of the tubes, *i.e.*,

$$i_b = i_{b_1} + i_{b_2}$$

$$\left.\begin{aligned} i_{b_1} &= \frac{E_m \sin \omega t}{r_d + R} \\ &= I_m \sin \omega t \\ i_{b_2} &= 0 \end{aligned}\right\} \text{for } 0 < \omega t <$$

and

$$\left.\begin{aligned} i_{b_1} &= 0, \\ i_{b_2} &= -\frac{E_m \sin \omega t}{r_d + R} \\ &= -I_m \sin \omega t \end{aligned}\right\} \text{for } \pi < \omega t < 2\pi.$$

The waveforms of input voltage and currents are shown in fig. 3*b*.

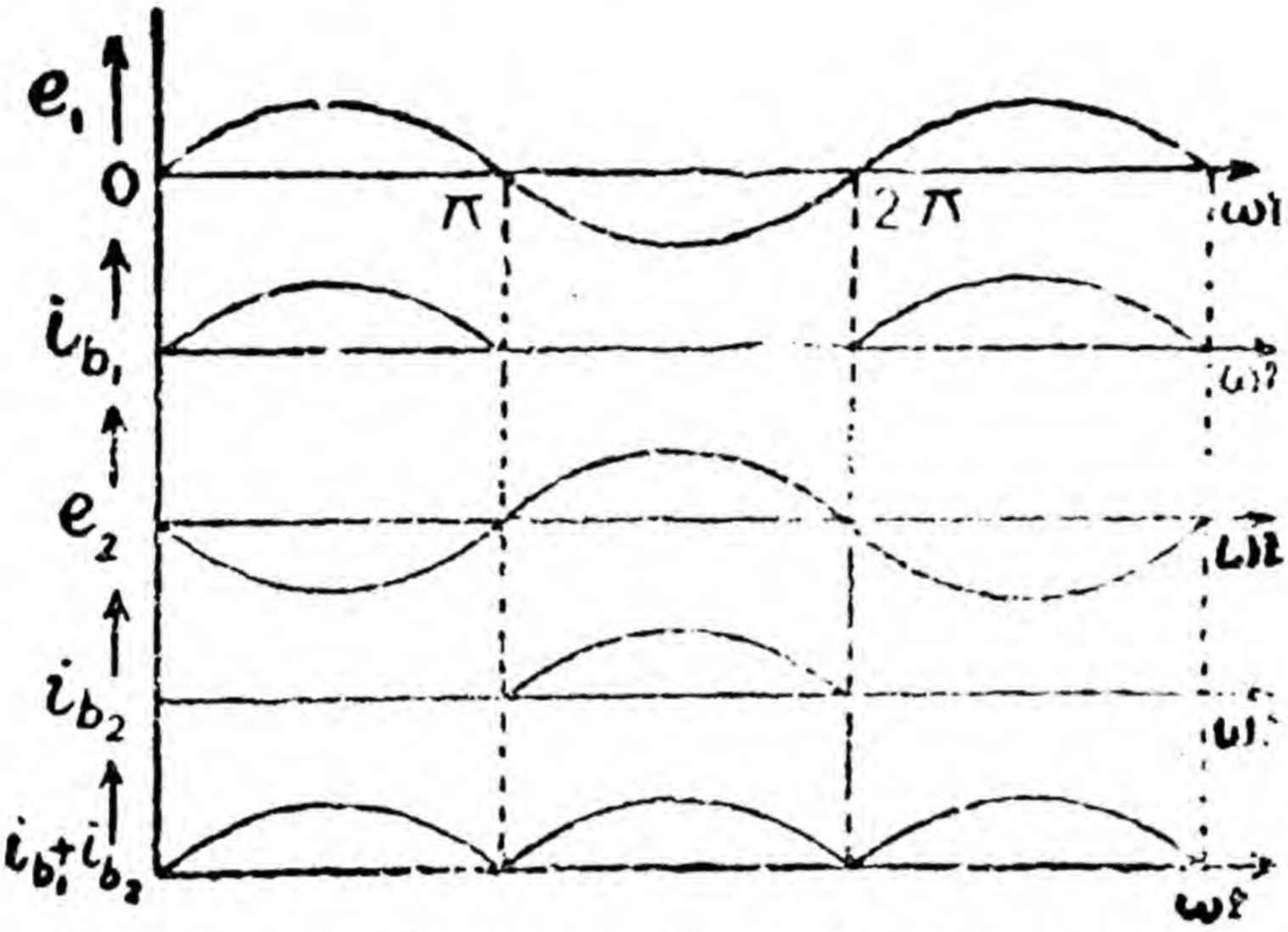


Fig. 3*b*. Waveforms in full wave rectification.

The general expression for current i_b can be written as

$$i_b = I_m \sin \omega t,$$

wnere I_m is maximum current $= \frac{E_m}{r_d + R}$ and E_m represents the peak value of e_1 or e_2.

D. C. or average value of current : As the form of current in both valves is same, we write

$$I_{d.c.} = \frac{1}{2\pi}\int_0^{\pi} i_{b_1}\, d(\omega t) + \frac{1}{2\pi}\int_{\pi}^{2\pi} i_{b_2}\, d(\omega t)$$

$$= \frac{1}{2\pi}\int_0^{\pi} \frac{E_m \sin \omega t}{(r_d + R)}\, d(\omega t) - \frac{1}{2\pi}\int_{\pi}^{2\pi} \frac{E_m \sin \omega t}{(r_d + R)}\, d(\omega t)$$

$$= \frac{E_m}{2\pi (r_d + R)}\Big[-(\cos \omega t)\Big]_0^{\pi} - \frac{E_m}{2\pi (r_d + R)}\Big[-\cos \omega t\Big]_{\pi}^{2\pi}$$

$$= \frac{E_m}{\pi (r_d + R)} - \frac{E_m}{2\pi (r_d + R)}[-1 - 1]$$

$$= \frac{E_m}{\pi (r_d + R)} + \frac{E_m}{\pi (r_d + R)} = \frac{2E_m}{\pi (r_d + R)}$$

$$= \frac{2}{\pi} I_m, \text{ since } I_m = \frac{E_m}{(r_d + R)}.$$

R.M.S. value of current :

$$I_{r.m.s.} = \left\{ \frac{1}{2\pi} \left[\int_0^{\pi} i_{b_1}^2 \, d(\omega t) + \int_{\pi}^{2\pi} i_{b_2}^2 \, d(\omega t) \right] \right\}^{1/2}$$

$$= \left\{ \frac{1}{2\pi} \left[\int_0^{\pi} \left(\frac{E_m \sin \omega t}{r_d + R} \right)^2 d(\omega t) + \int_{\pi}^{2\pi} \left(\frac{-E_m \sin \omega t}{r_d + R} \right)^2 d(\omega t) \right] \right\}^{1/2}$$

$$= \left\{ \frac{1}{2\pi} \frac{E_m^2}{(r_d + R)^2} \left[\int_0^{\pi} \sin^2 \omega t \, d(\omega t) + \int_{\pi}^{2\pi} \sin^2 \omega t \, d(\omega t) \right] \right\}^{1/2}$$

$$= \left\{ \frac{1}{2\pi} \frac{E_m^2}{(r_d + R)^2} \left[\int_0^{\pi} \left(\frac{1 - \cos 2\omega t}{2} \right) d(\omega t) + \int_{\pi}^{2\pi} \left(\frac{1 - \cos 2\omega t}{2} \right) d(\omega t) \right] \right\}^{1/2}$$

$$= \left\{ \frac{1}{2\pi} \frac{E_m^2}{(r_d + R)^2} \left[\frac{\pi}{2} + 0 + \left(\pi - \frac{\pi}{2} \right) + 0 \right] \right\}^{1/2}$$

$$= \left\{ \frac{1}{2} \frac{E_m^2}{(r_d + R)^2} \right\}^{1/2}$$

$$= \frac{I_m}{\sqrt{2}} \text{ where } I_m = \frac{E_m}{r_d + R}. \qquad \ldots(1)$$

Average voltage across the load ;

$$E_{d.c.} = I_{d.c.} R = \frac{2I_m}{\pi} . R$$

$$= \frac{2}{\pi} \cdot \frac{E_m}{r_d + R} . R = \frac{2}{\pi} \frac{E_m}{\left(1 + \frac{r_d}{R} \right)}. \qquad \ldots(2)$$

Rectifier Efficiency :

$$\eta = \frac{P_{d.c.}}{P_{a.c.}} \times 100\%.$$

$$P_{d.c.} = R I_{d.c.}^2 = R \left(\frac{E_m}{r_d + R} \right)^2 \cdot \frac{2^2}{\pi^2},$$

$$P_{a.c.} = I_{r.m.s.}^2 \, (r_d + R) = \left[\frac{E_m}{\sqrt{2} (r_d + R)} \right]^2 . (r_d + R)$$

$$= \frac{E_m^2}{2 (r_d + R)}.$$

$$\therefore \quad \eta = \frac{\frac{4}{\pi^2} \cdot \frac{E_m^2}{(r_d + R)^2} \times R}{E_m^2 / 2 (R + r_d)} \times 100\%$$

$$= \frac{8R}{\pi^2 (r_d + R)} . 100\%$$

$$= \frac{81 \cdot 2}{\left(1 + \frac{r_d}{R} \right)} \%. \qquad \ldots(3)$$

Ripple Factor : The purpose of the rectifier is to produce steady output voltage and current from the sinusoidal voltage source but this is not perfectly achieved. In other words the output lack smoothness of the waveform which is referred to as a ripple. The ripple voltage or current is measured by the *ripple factor*.

$$\text{Ripple factor } \gamma = \frac{\text{effective value of alternating component of voltage or current}}{\text{direct value of voltage or current}}$$

$$\text{Rectifier current} = I_{d\cdot c\cdot} + I_{a\cdot c\cdot},$$

$$I_{r\cdot m\cdot s\cdot} = \sqrt{(I^2_{d\cdot c\cdot} + I^2_{a\cdot c\cdot})},$$

$$I^2_{r\cdot m\cdot s\cdot} = I^2_{d\cdot c\cdot} + I^2_{a\cdot c\cdot}$$

$$I^2_{a\cdot c\cdot} = I^2_{r\cdot m\cdot s\cdot} - I^2_{d\cdot c\cdot}$$

$$I_{a\cdot c\cdot} = \sqrt{(I^2_{r\cdot m\cdot s\cdot} - I^2_{d\cdot c\cdot})},$$

$$\gamma = \sqrt{\left(\frac{I^2_{r\cdot m\cdot s\cdot} - I^2_{d\cdot c\cdot}}{I^2_{d\cdot c\cdot}}\right)}$$

$$= \sqrt{\left(\frac{I^2_{r\cdot m\cdot s\cdot}}{I^2_{d\cdot c\cdot}} - 1\right)}$$

$$\text{But } \frac{I_{r\cdot m\ s\cdot}}{I_{d\cdot c\cdot}} = \frac{I_m/\sqrt{2}}{2I_m/\pi} = \frac{\pi}{2\sqrt{2}} = 1\cdot 11,$$

$$\therefore \gamma = \{(1\cdot 11)^2 - 1\}$$

$$= 0\cdot 48.$$

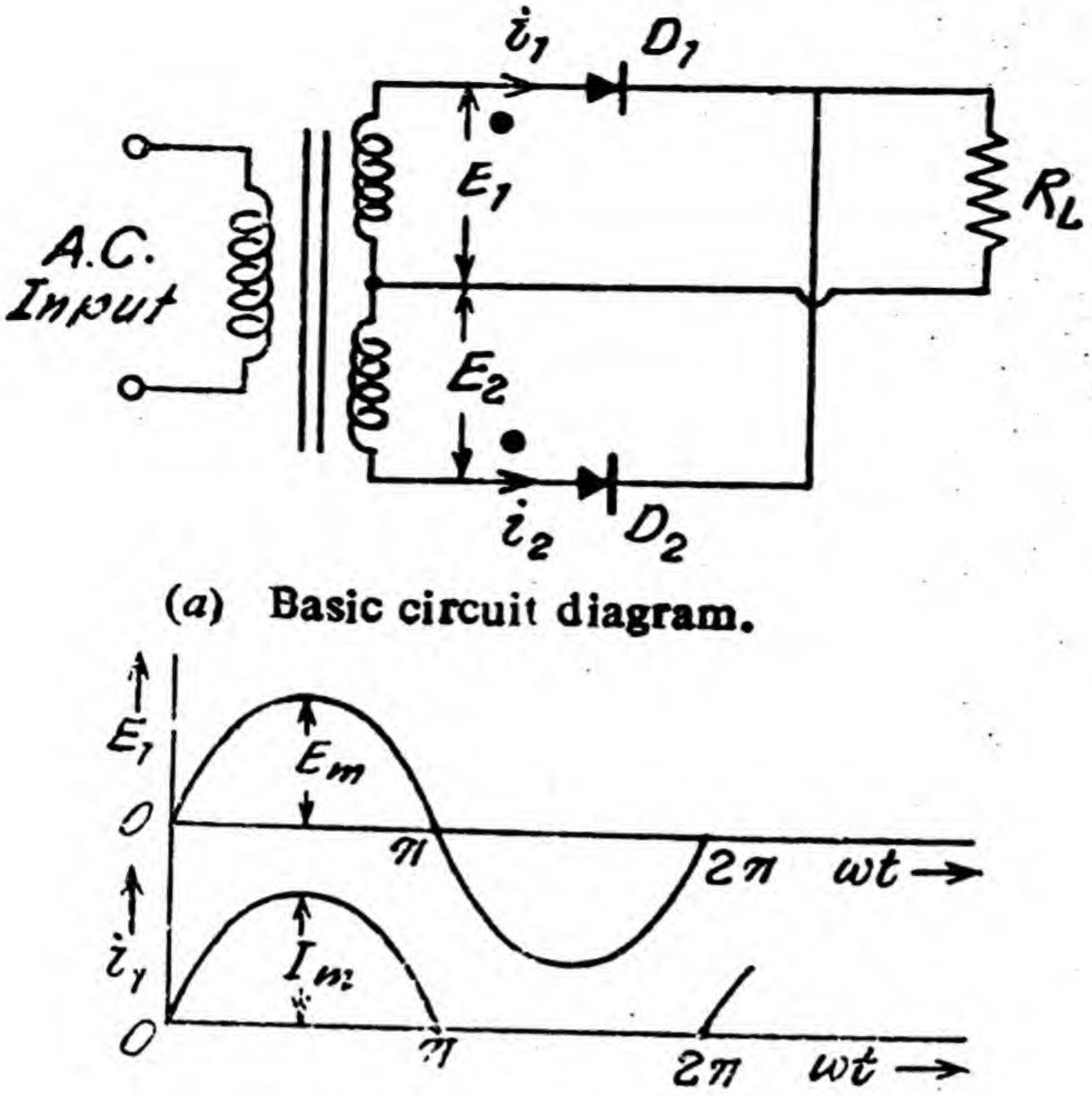
Peak Inverse Voltage :

The peak inverse voltage across T_1 can be calculated by assuming that T_1 is not conducting while T_2 is conducting to its maximum. Under this condition the centre tap is positive with respect to the cathode of valve T_1 and both e_1 and e_2 are passing through their maximum values. If we neglect the drop across the conducting diode, then the voltage developed across the load R will be E_m with upper end positive of the load with respect to centre tap. In a similar way, E_m will be developed with centre tap positive with respect to upper end of the load due to e_1. Thus P.I.V. across the non-conducting diode T_1 is equal to the sum of two voltages, *i.e.* $E_m + E_m = 2E_m$.

$\therefore$ P.I.V. for full wave wave rectifier $= 2E_m$.

8·1-4. FULL-WAVE RECTIFIER (Semiconductor diode)

Fig (4*a*) shows the basic circuit of a full wave rectifier. With the help of a centre tapped transformer, two equal and opposite voltages E_1 and E_2 are obtained. During one half cycle, the diode D_1 conducts and diode D_2 remains non-conducting. A current i_1 flows through the load resistor R_L. During the other half cycle, the diode D_2 conducts while diode D_1 remains non-conducting. Now a current i_2 flows through load resistor R_L. The waveforms of E_1 and E_2 and currents i_1 and i_2 are shown in fig. (4*b*).



(*a*) **Basic circuit diagram.**

(*b*) **Waveforms of applied a.c. voltages and output currents.**

Fig. 4. Full wave rectifier.

The total current flowing through R_L is given by

$$i = i_1 + i_2 \qquad \ldots(1)$$

where $i_1 = I_m \sin \omega t$ for $0 \leqslant \omega t \leqslant \pi$

$$i_1 = 0 \text{ for } \pi \leqslant \omega t \leqslant 2\pi$$

$$I_m = \frac{E_m}{R_f + R_L}$$

Similarly, $i_2 = 0$ for $0 \leqslant \omega t \leqslant \pi$

and $i_2 = I_m \sin \omega t$ for $\pi \leqslant \omega t \leqslant 2\pi$.

DC Output Current—Since the current is of the same form in the two halves of a.c. cycle

$$I_{dc}=\frac{1}{\pi}\int_0^{\pi} i\, d(\omega t)=\frac{1}{\pi}\int_0^{\pi} I_m \sin \omega t\, d(\omega t)$$

$$=\frac{2I_m}{\pi} \qquad \ldots (2)$$

D.C. Output Voltage—Following a similar discussion to that used for half wave rectifier, the dc output voltage across load is given by

$$E_{dc}=I_{dc}\, R_L=\frac{2I_m\, R_L}{\pi} \qquad \ldots(3)$$

Thus E_{dc} is twice that in a half wave rectifier.

RMS Current—$I_{rms}=\left[\frac{1}{\pi}\int_0^{\pi} i^2\, d(\omega t)\right]^{1/2}=\left[\frac{I_m^2}{\pi}\int_0^{\pi} \sin^2 \omega t\, d(\omega t)\right]^{1/}$

or
$$I_{rms}=\frac{I_m}{\pi} \qquad \ldots(4)$$

Rectifier Efficiency – We know that rectifier efficiency η is given by

$$\eta=\frac{P_{dc}}{P_{ac}}$$

Here $P_{ac}=I_{dc}^2\ R_L=\frac{4I_m^2}{\pi^2}\ R_L$

and $P_{ac}=I^2_{rms}\,(R_f+R_L)=\frac{I^2_m}{2}(R_f+R_L)$

$$\therefore \quad \eta=\frac{4I^2_m\, R_L/\pi^2}{I^2_m\ (R_f+R_L)/2}=\frac{8}{\pi^2}\cdot\frac{R_L}{(R_f+R_L)}$$

$$=\frac{0\cdot 812}{1+(R_f/R_L)}.$$

$$\text{Percentage efficiency}=\frac{81\cdot 2}{1+(R_f/R_L)}$$

Thus the full wave rectifier efficiency is twice that of half wave rectifier.

Ripple Factor—As $\frac{I_{rms}}{I_{dc}}=\frac{I_m/\sqrt{2}}{2I_m/\pi}=1\cdot 11$

$$\gamma=\sqrt{\left[\left(\frac{I_{rms}}{I_{dc}}\right)^2-1\right]}=\sqrt{[(1\cdot 11)^2-1]}=0\cdot 48$$

Regulation—$E_{dc}=\frac{2I_m\, R_L}{\pi}=\frac{2E_m\, R_L}{\pi\,(R_f+R_L)}$

$$=\frac{2E_m}{\pi}\left[1-\frac{R_f}{R_f+R_L}\right]=\frac{2E_m}{\pi}-I_{dc}\, R_f$$

8·2. COMPARISON OF HALF AND FULL WAVE RECTIFIERS :

(*i*) The ripple in full wave rectifier is small in comparison to half wave rectifier. Due to high ripple in half wave rectifier, an expensive smoothing filter system is required.

(*ii*) Efficiency of full wave rectifier is twice that of half wave rectifier.

(*iii*) The output of half wave rectifier is low. Transfer efficiency is also low.

(*iv*) Since in half wave rectifier current flows through the secondary in the same direction, it saturates the core of transformer which increases the hysteresis loss and produces harmonics in the secondary output.

TABLE

	Full wave	Half wave
1. Maximum efficiency	81·2%	40·6%
2. Average or d.c. current	$2I_m/\pi$	I_m/π
3. R.M S current	$I_m/\sqrt{2}$	$I_m/2$
4. Ripple factor	0· 8	1·21
5. P.I.V.	$2E_m$	E_m

8·3. GAS DIODE RECTIFIERS (Hot Cathode).

Gas diodes are also used in half wave and full wave rectifier circuits.

8·3-1. HALF WAVE GAS DIODE RECTIFIERS :

In fig. 5, the circuit of half wave rectifiers with a load resistance is shown. The tube conducts during one half cycle of input voltage.

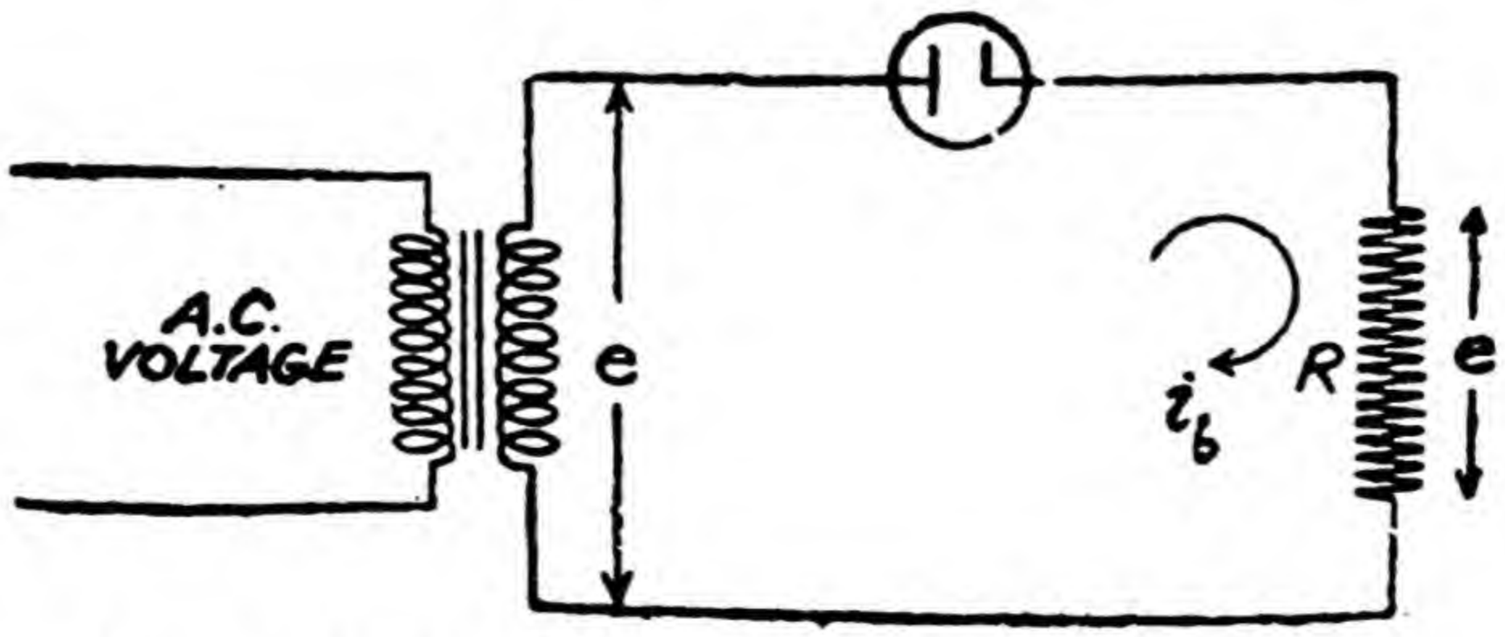


Fig. 5. Circuit of half wave rectifier using the gas diode.

Volt-ampere characteristic of a gas diode is given in figure 6. It is clear from the characteristic that for the conduction of the tube, the applied potential must exceed the ionisation potential E_0. Thus the potential drop appears across the tubes. Because of this potential drop, the tube conducts for less than half cycle as shown in Figure 7. The tube conducts at angle α_1 which ceases at α_2.

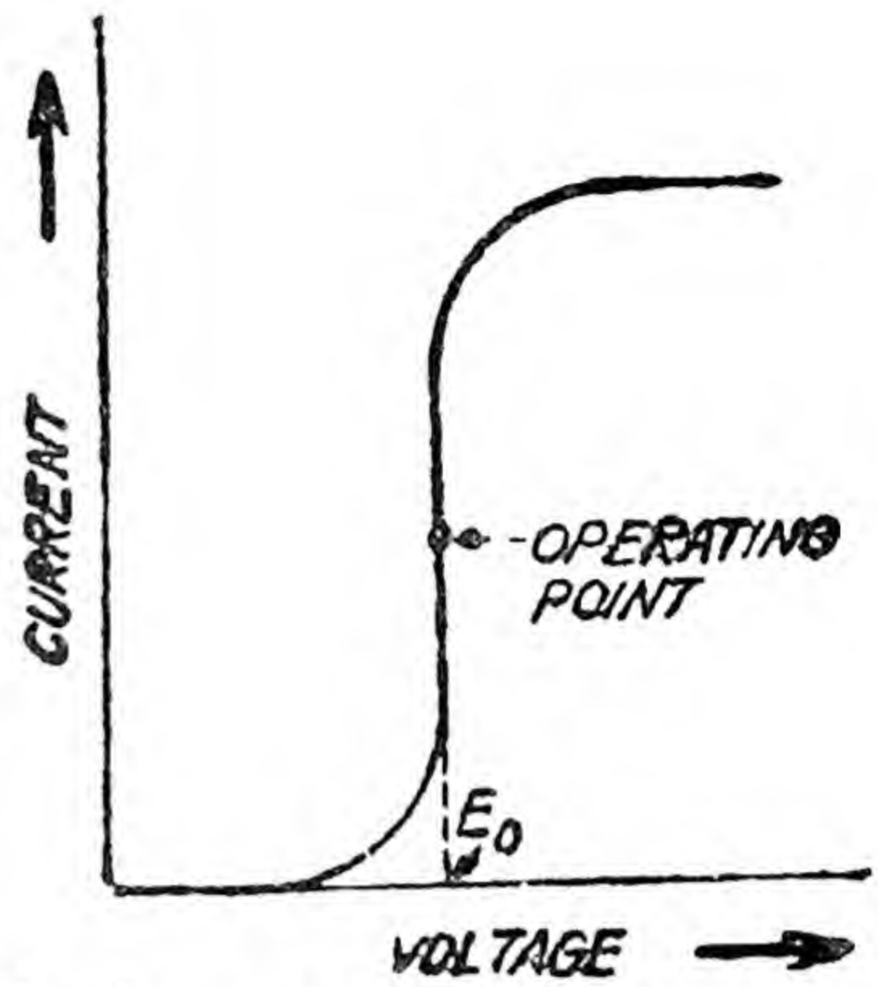


Fig. 6 Actual characteristic of a gas diode

Analysis : Applying Kirchhoff's law to the plate circuit, we have

$$e_0 = R.i_b = E_m \sin \omega t - E_0$$

$$\therefore \quad i_b = \frac{E_m \sin \omega t - E_0}{R}, \qquad \ldots(1)$$

where E_0 is voltage drop across tube during conduction.

Direct plate current will be obtained if we take average value of the instantaneous current given by

$$I_{d.c.} = \frac{1}{2\pi}\int_{\alpha_1}^{\alpha_2} \frac{E_m \sin \omega t - E_0}{R}.d(\omega t).$$

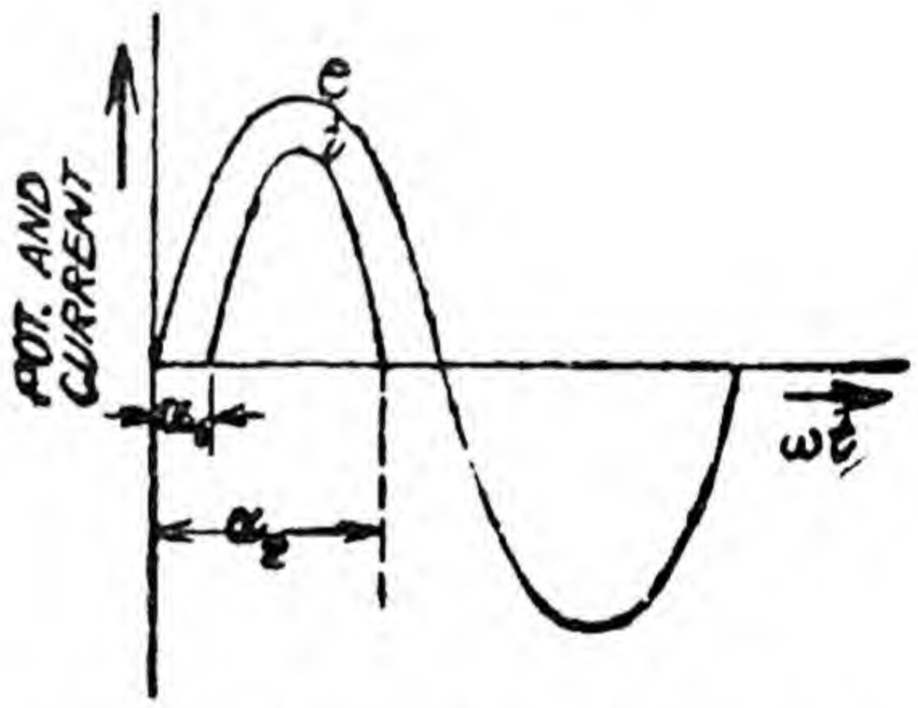


Fig. 7. Voltage and current phase of a gas diode half wave rectifier.

Since applied plate potential is much greater than E_0, we can take α_1 and $(\pi - \alpha_2)$ very nearly to zero. Thus

$$I_{d.c.} = \frac{1}{2\pi}\int_0^{\pi} \frac{E_m \sin \omega t - E_0}{R}\, d(\omega t)$$

$$= \frac{1}{2\pi}\cdot\frac{1}{R}\left[\int_0^{\pi} E_m \sin \omega t\, d(\omega t) - \int_0^{\pi} E_0\, d(\omega t)\right]$$

$$= \frac{E_m}{\pi R} - \frac{E_0}{2R} = \frac{E_m}{\pi R}\left(1 - \frac{\pi}{2}\frac{E_0}{E_m}\right). \qquad \ldots(2)$$

Therefore the output d.c. power

$$P_{d.c.} = I^2_{d.c.}\ R = \frac{E_m^2}{\pi^2 R}\left(1 - \frac{\pi}{2}\frac{E_0}{E_m}\right)^2. \qquad \ldots(3)$$

The input power to the plate circuit can be calculated by the expression

$$P_{a.c.} = \frac{1}{2\pi}\int_0^{\pi} e.i_b.d(\omega t)$$

$$= \frac{1}{2\pi}\int_0^{\pi} E_m \sin \omega t.\frac{E_m \sin \omega t - E_0}{R}\, d(\omega t)$$

$$= \frac{1}{2\pi}\int_0^{\pi} \frac{E_m^2 \sin^2 \omega t}{R}\, d(\omega t) - \frac{1}{2\pi}\int_0^{\pi} E_0 \frac{E_m \sin \omega t}{R}.d(\omega t)$$

$$= \frac{E_m^2}{4R} - \frac{E_0 E_m}{\pi R}$$

$$= \frac{E_m^2}{4R}\left(1 - \frac{4.E_0}{\pi E_m}\right). \qquad \ldots(4)$$

Therefore the efficiency of rectification is given by

$$\eta = \frac{P_{d.c.}}{P_{a.c.}} \times 100\% = \frac{\frac{E_m^2}{R\pi^2}\left(1 - \frac{\pi}{2}\frac{E_0}{E_m}\right)^2}{\frac{E_m^2}{4R}\left(1 - \frac{4}{\pi}\frac{E_0}{E_m}\right)} \times 100\%$$

$$= \frac{4}{\pi^2} \cdot \frac{\left(1 - \frac{\pi}{2}\frac{E_0}{E_m}\right)^2}{\left(1 - \frac{4}{\pi}\frac{E_0}{E_m}\right)} \times 100\%$$

which may be reduced to an expression.

$$\eta = 40{\cdot}6\ (1 - 1{\cdot}87\ E_0/E_m)\% \qquad \ldots(5)$$

Thus efficiency does not involve R or i_b in its expression which means it is independent of load resistance and load current.

8·3-2. FULL WAVE GAS DIODE RECTIFIER :

The circuit of a full wave rectifier using gas double diode is shown in fig. 8. The tube plates P_1 and P_2 conduct during both half cycles of input voltage because the inputs of two plates are 180° apart.

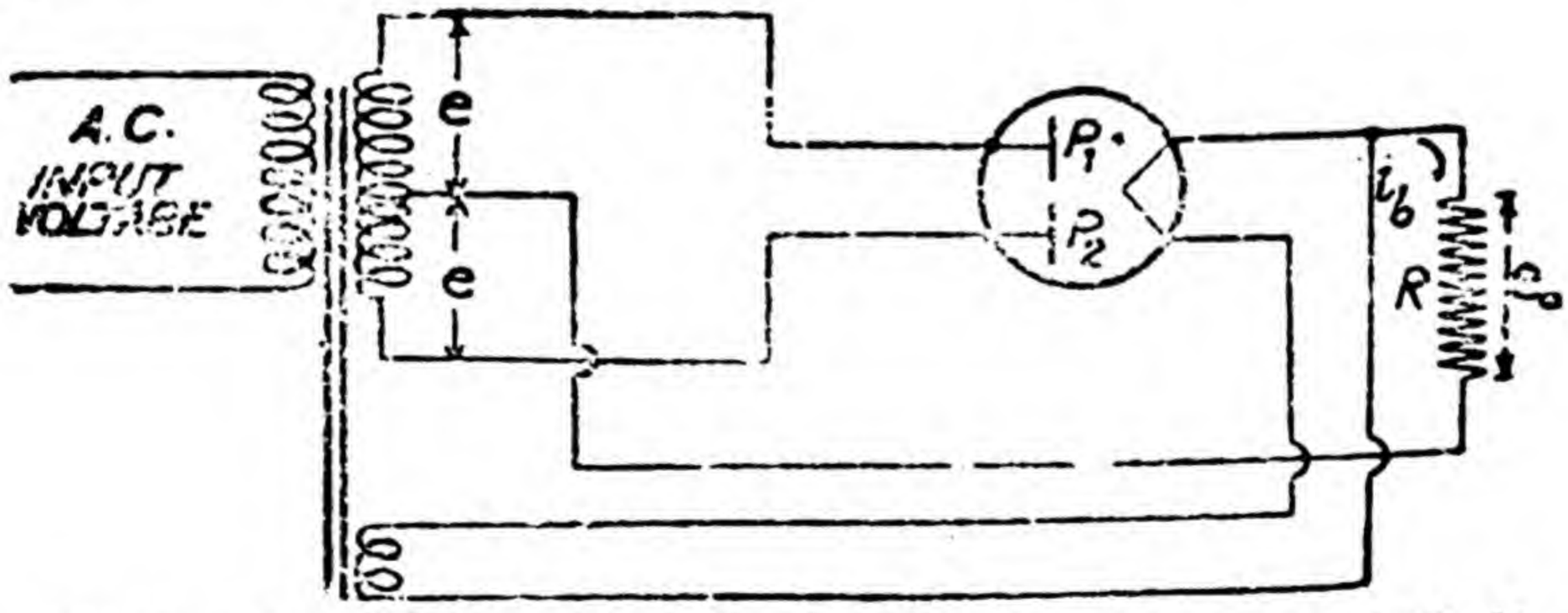


Fig. 8. Circuit of full wave rectifier using gas doubled diode.

In figure 9 the output voltage waveforms are also shown. It is apparent that each plate conducts for less than half cycle.

Analysis. We shall calculate the efficiency of a rectifier, given by

$$\eta = \frac{P_{d.c.}}{P_{a.c.}} \times 100\%.$$

$P_{d.c.}$ and $P_{a.c.}$ have been calculated in half wave rectifier analysis. In this case these values will be slightly different.

$$I_{a.c.} = 2.\frac{1}{2\pi}\int_0^{\pi} \frac{E_m \sin \omega t - E_0}{R}\, d(\omega t)$$

$$= 2.\frac{E_m}{\pi R}\left(1 - \frac{\pi}{2}\frac{E_0}{E_m}\right); \qquad (1)$$

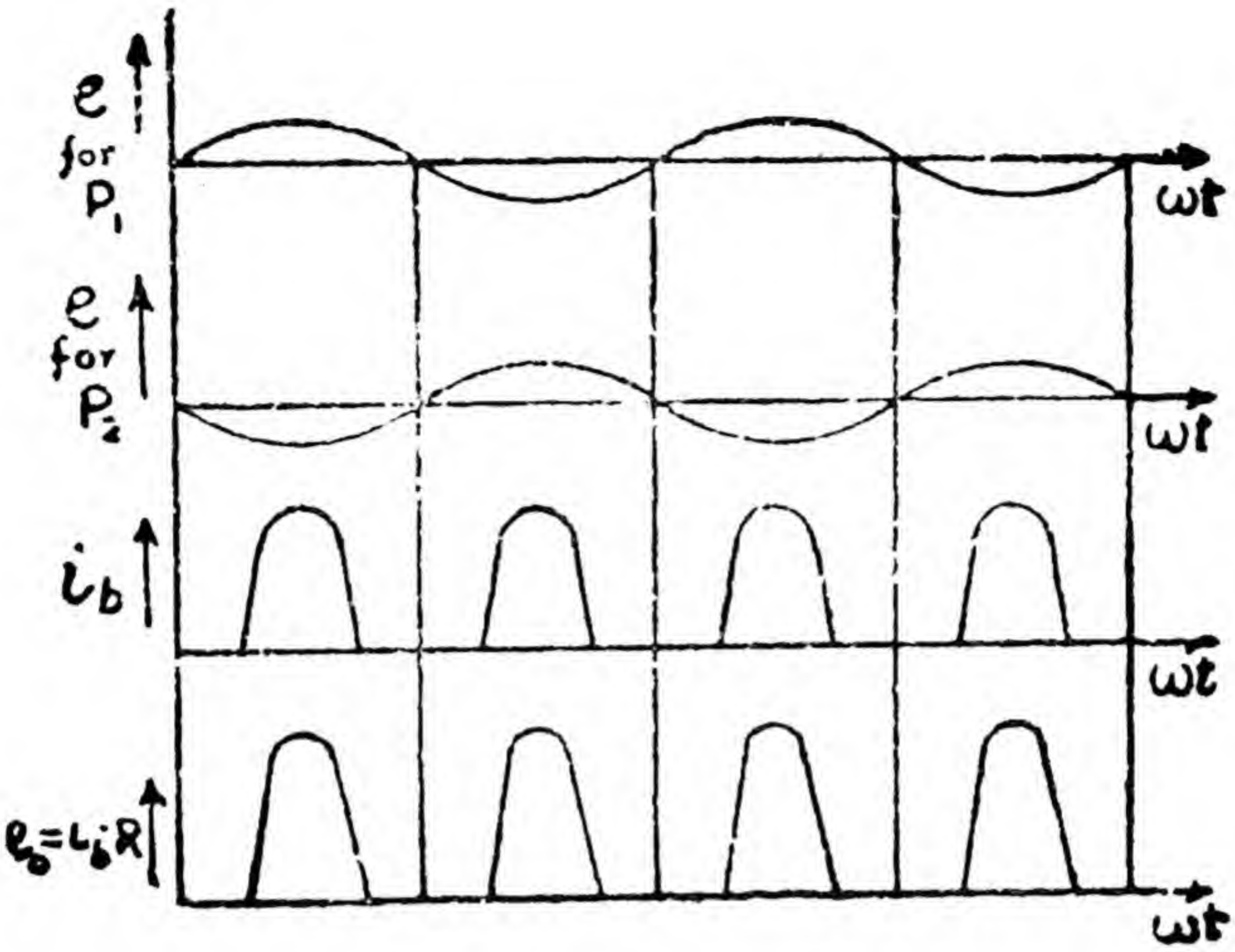


Fig. 9. Waveforms of output voltage and output current with input voltage.

$$\therefore \quad P_{d.c.} = R.I_{d.c.}^2$$

$$= \frac{4E_m^2}{\pi^2 R}\left(1-\frac{\pi}{2}\frac{E_0}{E_m}\right), \qquad \ldots(2)$$

$$P_{a.c.} = 2.\frac{1}{2\pi}\int_0^{\pi} e.i_b\, d(\omega t)$$

$$= 2.\frac{1}{2\pi}\int_0^{\pi} E_m \sin\omega t.\ \frac{E_m \sin\omega t - E_0}{R}.d(\omega t)$$

$$= \frac{E_m^2}{2R}\left(1-\frac{4}{\pi}\frac{E_0}{E_m}\right) \qquad \ldots(3)$$

$$\therefore \quad \eta = \frac{\frac{4}{\pi^2}.\frac{E_m^2}{R}\left(1-\frac{\pi}{2}\frac{E_0}{E_m}\right)^2}{\frac{E_m^2}{2R}\left(1-\frac{4}{\pi}\frac{E_0}{E_m}\right)} \times 100\%$$

$$= \frac{8}{\pi^2}\frac{\left(1-\frac{\pi}{2}\frac{E_0}{E_m}\right)^2}{\left(1-\frac{4}{\pi}\frac{E_0}{E_m}\right)} \times 100\%$$

$$= 81{\cdot}2\left(1-1{\cdot}87\frac{E_0}{E_m}\right)\% \qquad \ldots(4)$$

If we compare the efficiency of full wave vacuum tube rectifier with above, then we observe that both are same.

8·4. COMPARISON OF VACUUM AND GAS DIODE RECTIFIERS AND USES :

The vacuum tube rectifiers are used for low voltage, low power requirements, while gas diode rectifiers are used for high power

applications. In the latter case, generally hot cathode mercury vapour rectifiers are used.

In comparison with high vacuum tube rectifier, the hot cathode mercury vapour rectifier (*i*) requires low filament power, (*ii*) has small anode to cathode voltage drop, (*iii*) is cheaper in cost.

But for all these advantages it is essential that anode potential must be greater than or equal to 1000 volts and the current required be large.

The main disadvantage of hot cathode mercury vapour rectifier is that it has a tendency towards arc back and produces r.f. transients as the tube ionises. Due to all these defects and since momentary overload causes damage to cathode it is not used in home receivers.

8·5. FULL WAVE BRIDGE RECTIFIER CIRCUIT :

In previous articles half wave and full wave rectifier circuits have been discussed. A number of another rectifier circuit arrangements are also possible. An important rectifier circuit arrangement is a bridge rectifier in which diodes are arranged in the form of the four resistances in the wheatstone bridge as shown in figure (10). This circuit finds two important uses : (i) as power rectifier, and (ii) as a rectifying system in rectifier type a.c. meters. When this arrangement is used for power amplification, the rectifying elements are usually thermionic diode either of vacuum type or gas type. When it is used in a-c meters, the rectifying elements are metal rectifiers (either copper oxide type or selenium type). We know that in a *P-N* junction semiconductor diode, no heating power is required, and therefore the use of bridge rectifier with semiconductor diodes has increased considerably after the discovery of silicon diodes.

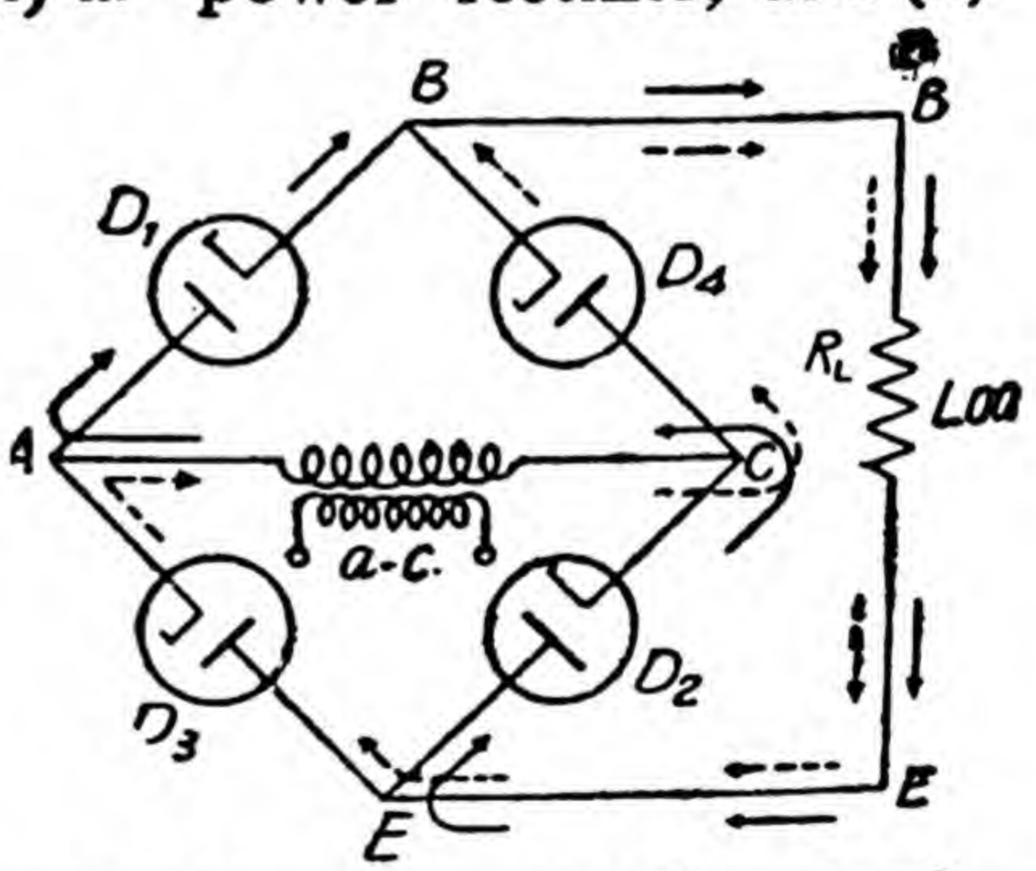


Fig. 10. Bridge rectifier circuit.

Operation : Let for the positive half cycle of the applied a.c. voltage, the point *A* is positive and *C* is negative. In this case the current flows through diode D_1, load resistance R_L and diode D_2 in the direction of thick arrow shown in the figure. The circuit is completed. The current through R_L flows in such a way that the point *B* is positive and *E* is negative. During this period, the diodes D_3 and D_4 are reversed biased and do not conduct the current. For the other half cycle, let the point *C* becomes positive and *A* negative. Now the current flows through diode D_4, load resistance R_L and the diode D_3. Again the circuit is completed. In

this case also the current flows through R_L so that point B is positive and E is negative as shown by dotted arrow.

Thus the current flows through the load in the same direction for both halves of the applied a.c. voltage. Thus the rectified current waveform is the same as that of a full wave rectifier. The diodes conduct simultaneously and in series.

Advantages : (*i*) No centre tap is required in the transformer secondary.

(*ii*) The peak inverse voltage across each tube is half as much as for the centre tap connection for the same $I_{d.c.}$ or $E_{d.c.}$

(*ii*) For a given power output, power transformer of a small size can be used as the current in both the primary and secondary of the plate supply transformer flows for the entire cycle.

Disadvantages : (*i*) As compared to a full wave rectifier, twice the number of rectifying diodes are required.

(*ii*) The two diodes are in series while in conduction and hence there is greater voltage drop, poor rectification efficiency and poor voltage regulation.

Use : The bridge rectifier is used in rectifier type voltmeter. The circuit arrangement is shown in figure (11). The rectifier elements are selenium or copper oxide rectifiers with a sensitive d. c. ammeter as a load. This circuit can be used for the measurement of a.c. as well as d.c. voltages and currents. The d.c. ammeter reads average value of currents. This may be calibrated to give r.m.s. values of applied sinusoidal voltages. When the applied voltage waveform differs from sinusoidal waveform, the values given out by this instrument is incorrect.

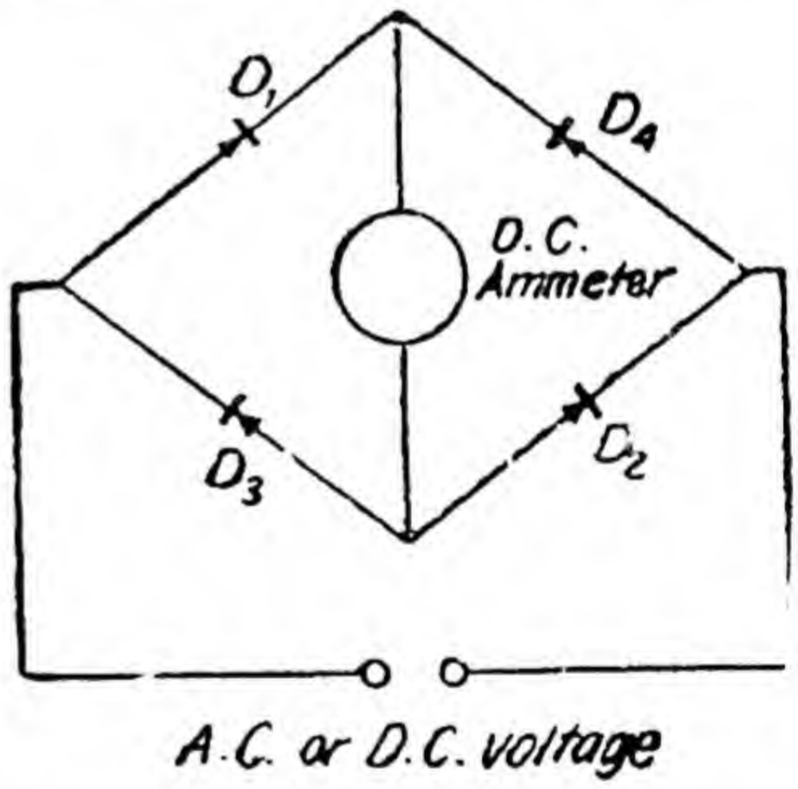


Fig. 11. Bridge rectifier type voltmeter.

8·6. METAL RECTIFIERS OR BARRIER LAYER RECTIFIERS :

The metal rectifiers utilise the unidirectional conduction property of the barrier layer between a good conductor and semiconductor. They differ from vacuum or gas filled diodes as they do not have thermionic emitters. As they do not require filament heating, they are most suitable for bridge rectifier circuits. Two types of such rectifiers are in common use :

(*i*) Copper oxide rectifier, and (*ii*) Selenium rectifier.

(1) Copper oxide rectifier : The copper oxide rectifier has a

disc of copper with one face coated with cuprous oxide (Cu_2O). The pure cuprous oxide is an insulator. When there is a slight excess of oxygen atoms, it loses its nonconducting property and becomes a semiconductor. This situation can be regarded as the missing of few copper atoms from the structure. If at any point in the structure, a copper atom is missing then there is an electron missing or there is a positive hole. Thus the conductivity of semiconducting Cu_2O is considered to be due to positive holes and it is said a *p*-type semiconductor. The oxide is heat treated in such a way as to give an excess of oxygen and a thin or barrier layer of Cu_2O at the surface of copper disc. The various layers of copper oxide rectifier are shown in fig (12). Thus a copper oxide rectifier can be regarded as consisting of a thin layer of pure Cu_2O sandwitched between copper and a layer of semiconducting Cu_2O with a top metallic plate. The top metallic plate acts merely as a contact.

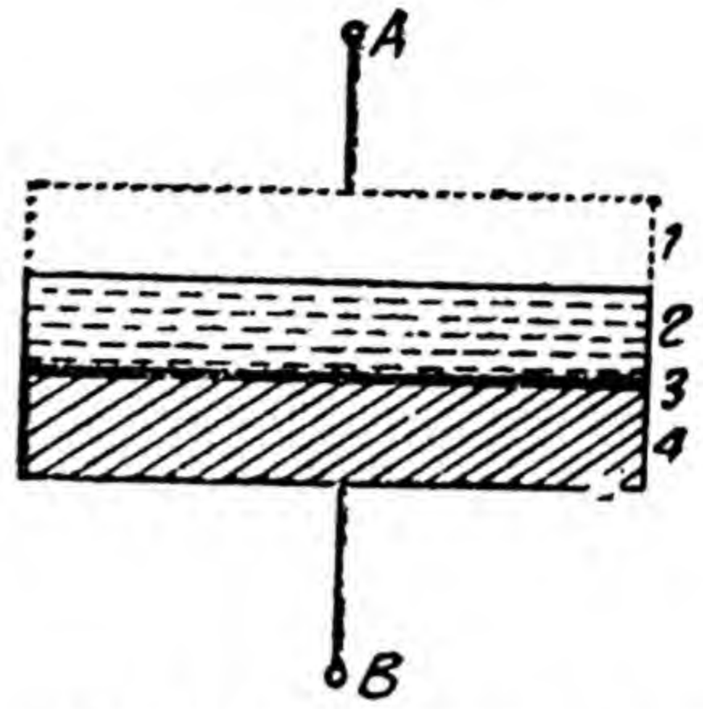
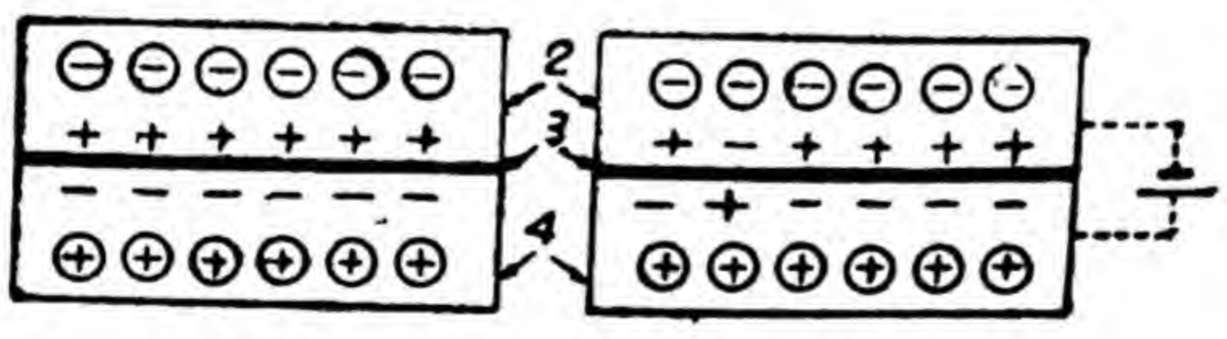


Fig. 12. Various layers of copper oxide rectifier. (1) Metallic plate, (2) p-type semiconductor, (3) Pure Cu_2O (Insulator), (4) Conductor (Copper plate).

Figure (13) shows the diffusion of electrons and holes; across the barrier layer. When the layers are formed, on one side, there will be mobile holes denoted by positive (+) sign in the semiconductor. At the same time there will be corresponding centre of negative charges denoted by negative ⊖ sign in fig. (13 *a*) as the



(*a*) Before diffusion (*b*) After diffusion

ig. 13.

material as a whloe is electrically neutral. On other side of barrier layer, there are mobile electrons denoted by negative (—) sign and fixed positively charged atoms denoted by ⊕ sign. Now mobile holes from the semiconductor and electrons from the metal tend to diffuse across the barrier layer. In this way there will be concentration of negative charge on semiconductor side and positive charge on metal side as shown in fig. (13 *b*). Thus a potential barrier is set up which opposes the diffusion. After sometime the equilibrium condition is reached when no further diffusion takes place. This develops an electric field across the barrier which is shown by a battery with dotted lines in fig. (13 *b*).

The principle of rectification is shown in figure (14). Considering figure (14 *a*), let during positive half cycle, *A* is posi-

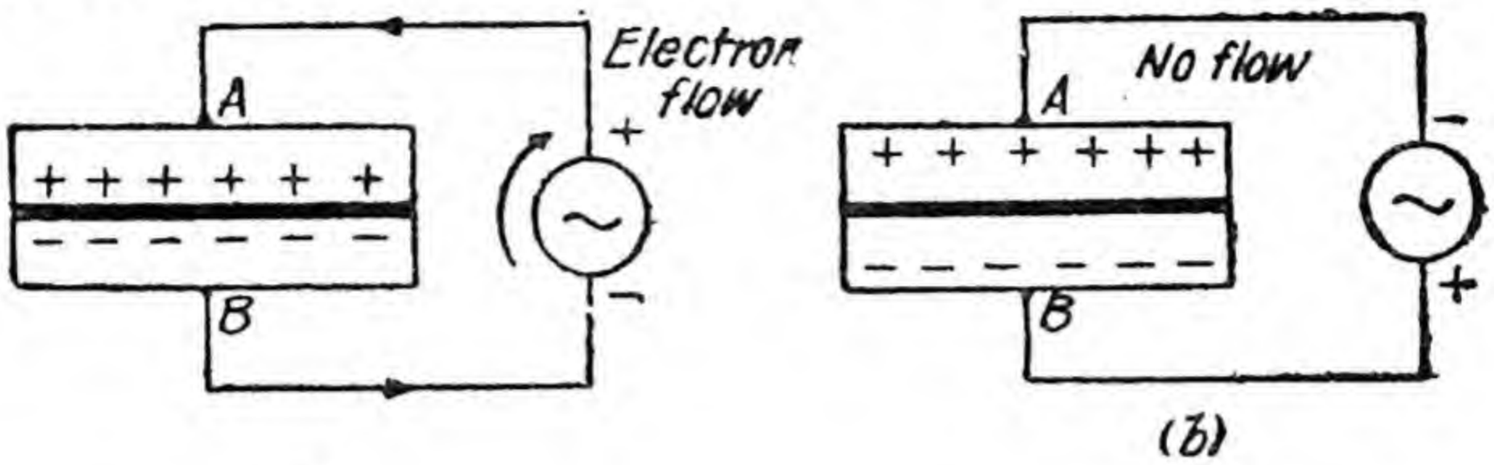
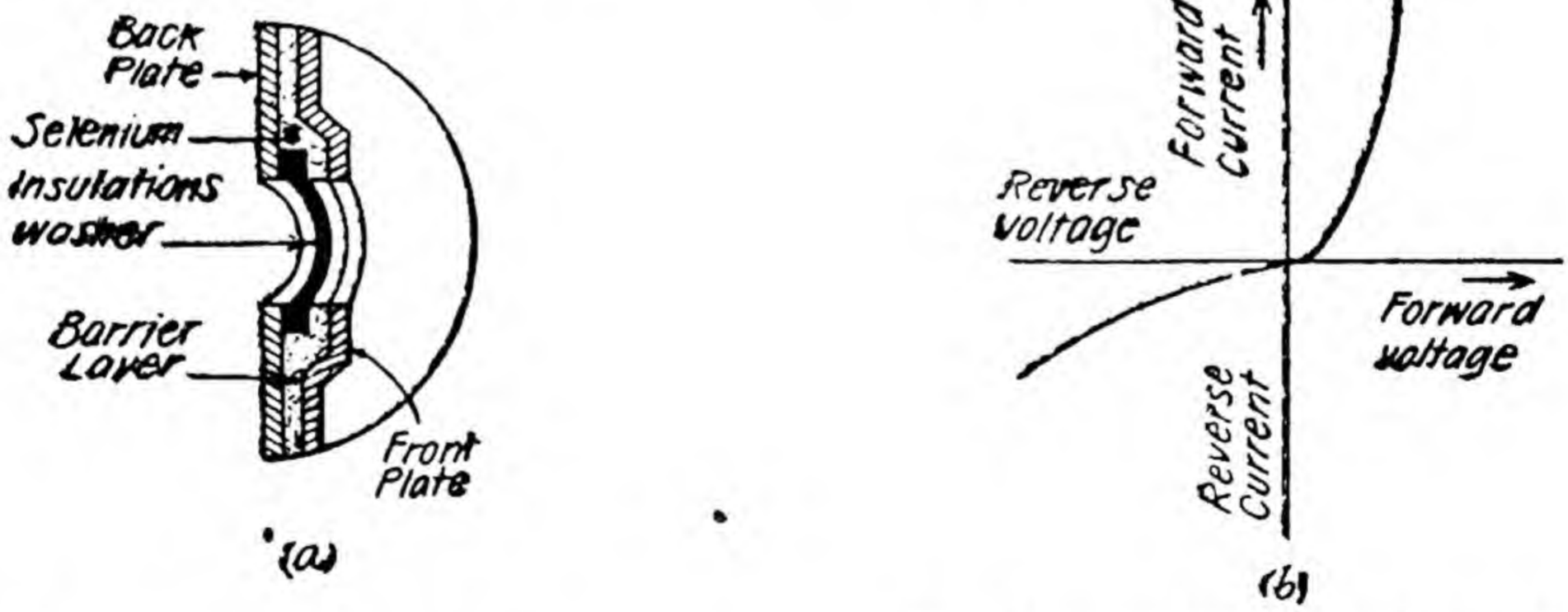


Fig. 14. Principle of rectification.

tive with respect to *B*. Now the holes of the semiconductor and electrons of metal are repelled towards barrier where they combine with each other. Immediately, an electron is given back to metal by a.c. source and another electron is absorbed from the semiconductor so that at every point in the circuit, a neutral charge is maintained. Thus a conventional current flows from semiconductor to metal inside the rectifier. As shown in fig. (14*b*). for the negative half cycle, the holes are attracted towards *A* while electrons are attracted towards *B*. Now no combination of holes and electrons takes place and hence no current flows in the circuit.

(ii) Selenium rectifier : In a selenium rectifier, a layer of selenium is formed on a steel plate which is the back plate. Selenium is normally a semiconductor but an insulating surface or barrier layer can be formed upon it. The surface is then sprayed with a metallic alloy to form the front surface. The cross section of a selenium rectifier is shown in figure (15 *a*). When selenium is



(*a*) Cross section of selenium rectifier.

(*b*) Current voltage characteristic of selenium rectifier.

Fig. 15.

positive with respect to the front plate, then for a low voltage drop, a large conduction current takes place. On the other hand, for a large voltage drop, a little conduction current takes place when selenium is negative with respect to front plate. This is analogus with a high vacuum diode, the semiconductor acts as anode while front plate as cathode. Fig. (15 *b*) shows the current voltage

characteristics of selenium rectifier. The direction of current now is known as forward direction while the opposite as reverse direction. If the applied voltage in the forward direction is below a certain value (threshold), very small current flows. The threshold value in selenium rectifier is at about 0·5 v.

8·7. MULTI-PHASE RECTIFIER CIRCUITS :

Multiphase rectifier circuits are used when large d.c. power is required. These circuits utilize standard three phase distribution system (according to modern practice the electric power is transmitted and distributed as three phase power). In case of a multiphase rectifier the output voltage is smoother than that of a single phase rectifier. The efficiency is also higher than a single phase rectifier. For large output powers, tubes having mercury pool cathodes are used. However tubes with hot cathodes can also be used. The commonly used multiphase rectifier circuits are :

(*i*) Three phase, Half wave system, (*ii*) Six phase, Half wave system, (*iii*) Three phase, Full wave system.

8·7-1. THREE PHASE, HALF WAVE SYSTEM :

Figure (16 *a*) illustrates the principle of a three phase, rectification circuit. The power is supplied through a three phase. delta wave connected transformer. The rectifying unit consists of either vacuum diodes or gaseous diodes or ignitons etc. having all cathodes connected together. The d.c. load is connected to the common

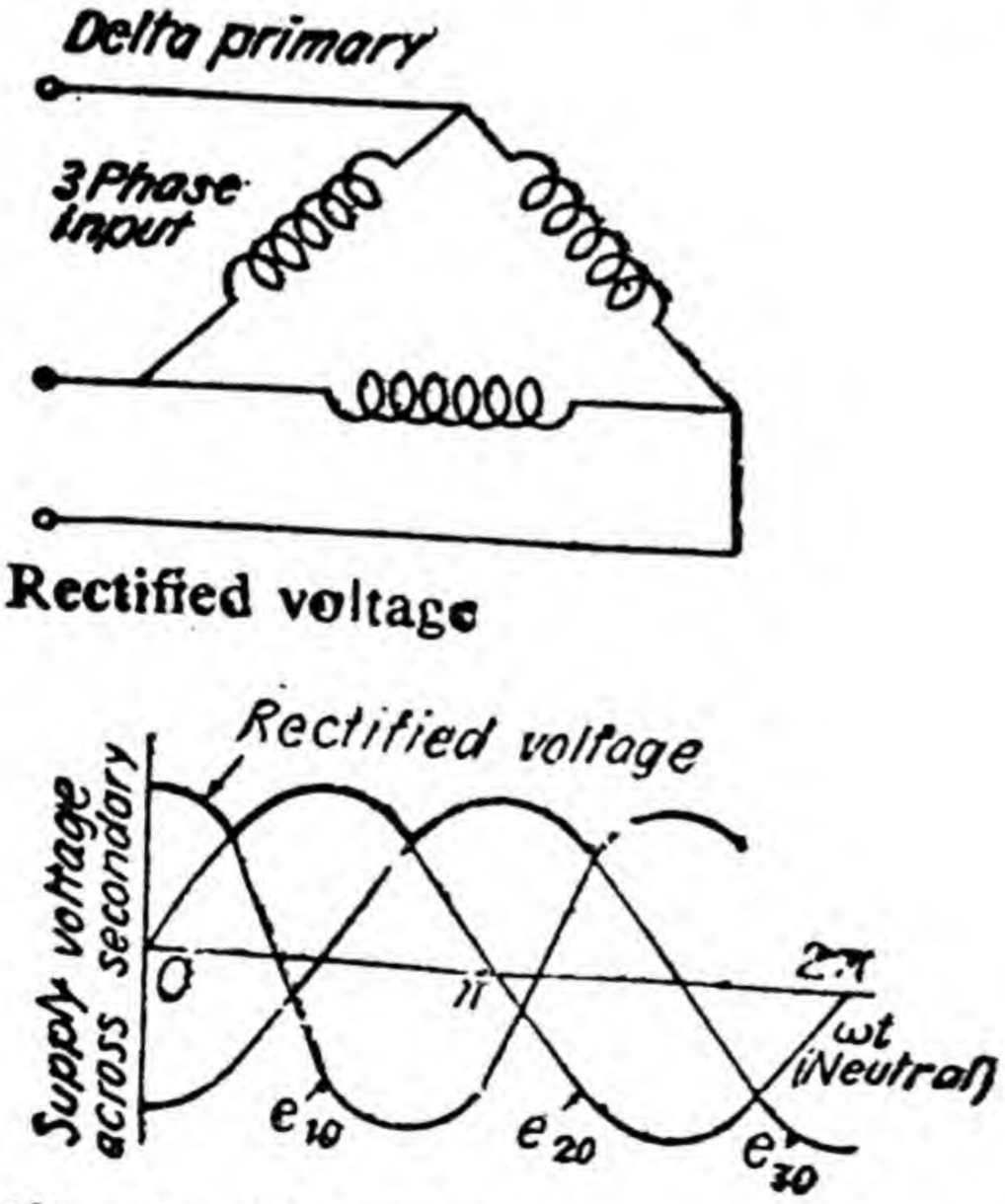


Fig. 16. Three phase, Half wave rectifier.

cathode terminal and to the neutral of the transformer secondary. The transformers serve two functions : (*i*) providing voltages on anodes, and (*ii*) providing the common connection, called the neutral connection. Each leg of the secondary wave provides one

phase. The voltages are given with respect to neutral connection. For sinusoidal input, the voltages applied to the anodes are sinusoidal, equal in magnitude but shifted from each other in phase by 120° as shown in fig (16 *b*). The anode voltages are e_{10}, e_{20} and e_{30} and each voltage is equal to the primary voltage multiplied by the transformer ratio. Each tube conducts for one third of a cycle and at any given instant, although more than one anode is positive with respect to the neutral but the anode having the highest positive supply voltage conducts. The arc transforms from one anode to the next anode when their voltages become equal. The anode currents for resistance load are shown in fig. (16 *c*).

As the waveform of load voltage or current repeats itself every third of a cycle, the average value of the load voltage is given by

$$E_{d.c.}=\frac{1}{(2\pi/3)}\int_{\pi/6}^{5\pi/6} E_m \sin \omega t.d(\omega t)$$

because e_{10} is positive relative to e_{20} and e_{30} when $\pi/6 < \omega t < 5\pi/6$.

$$=\frac{3E_m}{2\pi}\left[-\cos \omega t\right]_{\pi/6}^{5\pi/6}$$

$$=\frac{3E_m}{2\pi}\left[-\cos\frac{5\pi}{6}+\cos\frac{\pi}{6}\right]$$

$$=\frac{3E_m}{2\pi}\left[0{\cdot}866+0{\cdot}866\right]$$

$$=0{\cdot}827\ E_m \qquad \ldots(1)$$

The average current in the resistive load is

$$I_{d.c.}=\frac{0{\cdot}827\ E_m}{R_L}.$$

$\therefore$ Average tube current $=\frac{1}{3}.\ \frac{0{\cdot}827\ E_m}{R_L}$

Since the current $I_{d.c.}$ is supplied by the three tubes.

$$=0{\cdot}276\ E_m/R_L \qquad \ldots(2)$$

Here E_m is the amplitude of each phase voltage.

In order to compute the ripple factor, we first calculate the r.m.s. current

$$I_{r.m.s.}=\sqrt{\left[\frac{1}{(2\pi/3)}\int_{\pi/6}^{5\pi/6}\frac{E_m^2\sin^2\omega t}{R_L^2}\,d(\omega t)\right]}$$

$$=\frac{E_m}{R_L}\sqrt{\left[\frac{3}{2\pi}\int_{\pi/6}^{5\pi/6}\left(\frac{1-\cos 2\omega t}{2}\right)d(\omega t)\right]}$$

$$=\frac{E_m}{R_L}\left[\frac{3}{4\pi}.\left\{\omega t\right\}_{\pi/6}^{5\pi/6}-\left\{\frac{\sin 2\omega t}{2}\right\}_{\pi/6}^{5\pi/6}\right]^{1/2}$$

Solving it, we have

$$I_{r\cdot m\cdot s\cdot}=0\cdot 838\,\frac{E_m}{R_L} \qquad \ldots(3)$$

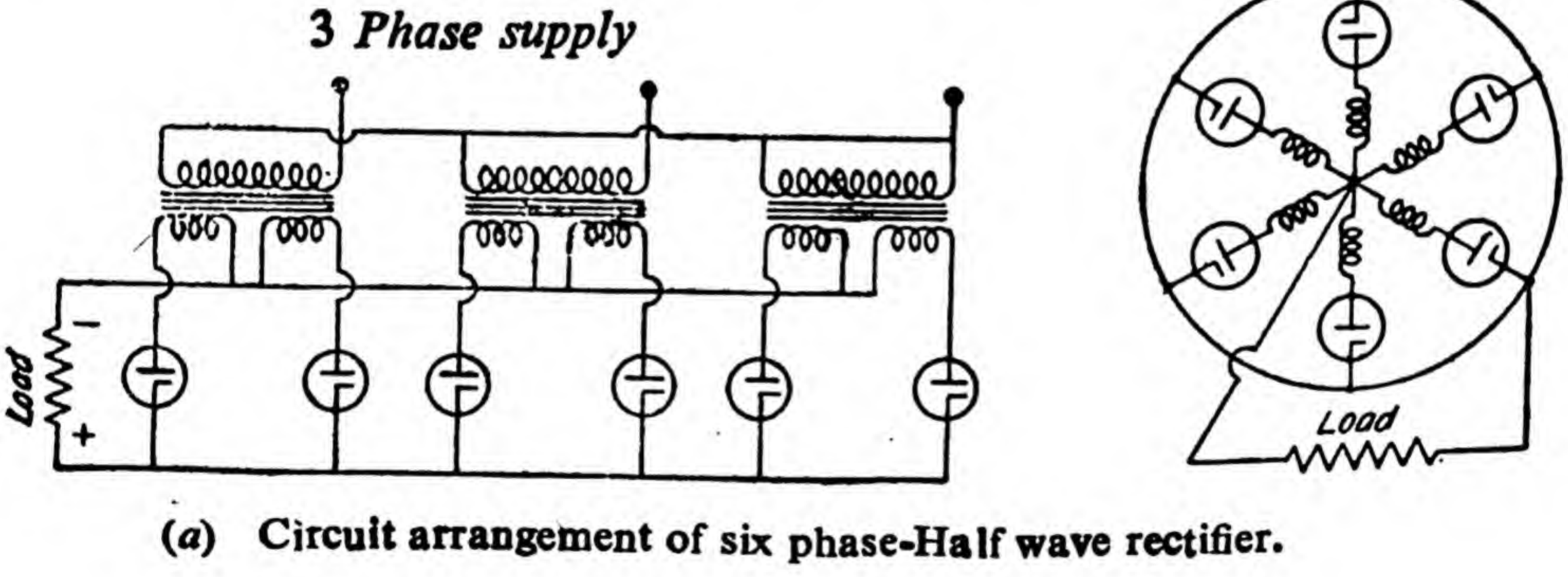
The ripple factor is given by

$$\gamma=\sqrt{\left\{\left(\frac{I_{r\cdot m\cdot s\cdot}}{I_{d\cdot c\cdot}}\right)^2-1\right\}}\times 100\%$$

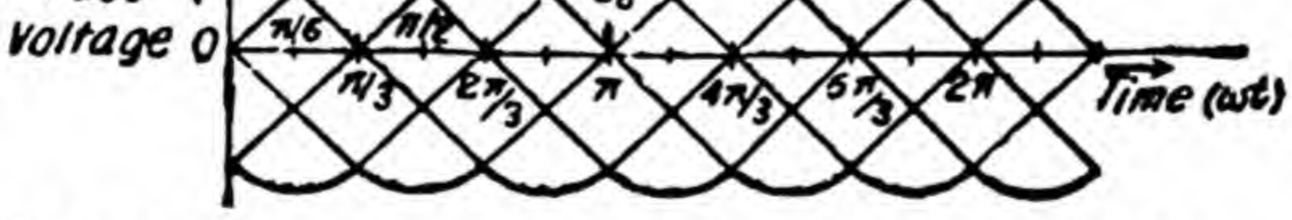
$$=\left[\left(\frac{\cdot 838}{0\cdot 827}\right)^2-1\right]^{1/2}\times 100\%=17\% \qquad \ldots(4)$$

8·7-2. SIX PHASE-HALF WAVE SYSTEM :

As compared to a single phase operation, three phase operation provides better conditions of rectification. Still better conditions can be obtained by using a number of phases. The three phase power can be converted into 6 or 12 phase power by means of transformers. The circuit and the operation of six phase half wave rectifier differ very little from the three phase half wave rectifier. The circuit arrangement is shown in figure (17 *a*). The circuit consists of six phases which are arranged in a symmetrical manner and each phase supplying one tube.



(*a*) **Circuit arrangement of six phase-Half wave rectifier.**



(*b*) **Voltage wave form of a six phase rectifier.**

Fig. 17. Delta, six phase, star rectifier circuit.

Since there are six phases each tube conducts for only one-sixth of the cycle as shown in fig. (17 *b*). The fundamental ripple frequency component is six times the supply frequency which results more smooth load voltage waveform as compared to three phase system. The analysis of this rectifier is as follow :

$$E_{d\cdot c\cdot}=\frac{E_m}{(2\pi/6)}\int_{\pi/6}^{\pi/2} E_m \sin\omega t\; d(\omega t)$$

$$= \frac{6E_m}{2\pi}\left[-\cos \omega t\right]_{\pi/6}^{\pi/2}$$

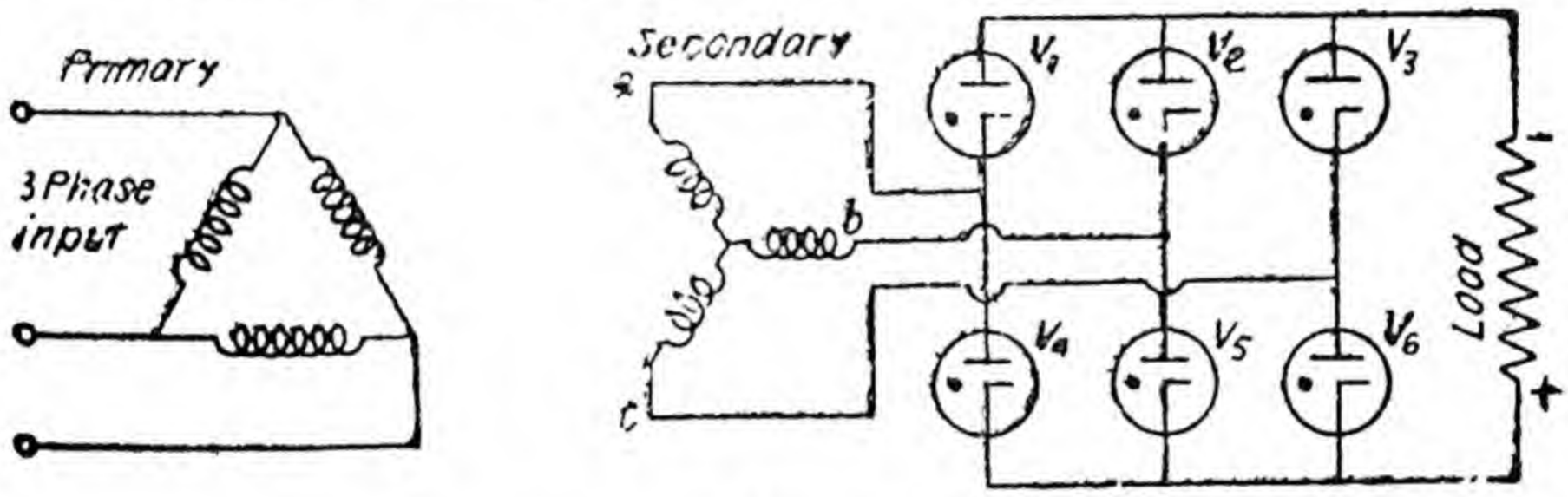
$$= \frac{3E_m}{\pi}\left[-\cos \pi/2 + \cos \pi/6\right] = \frac{3E_m}{\pi}\left[0 + 0{\cdot}866\right]$$

$$= 0{\cdot}82\ E_m \qquad \ldots(5)$$

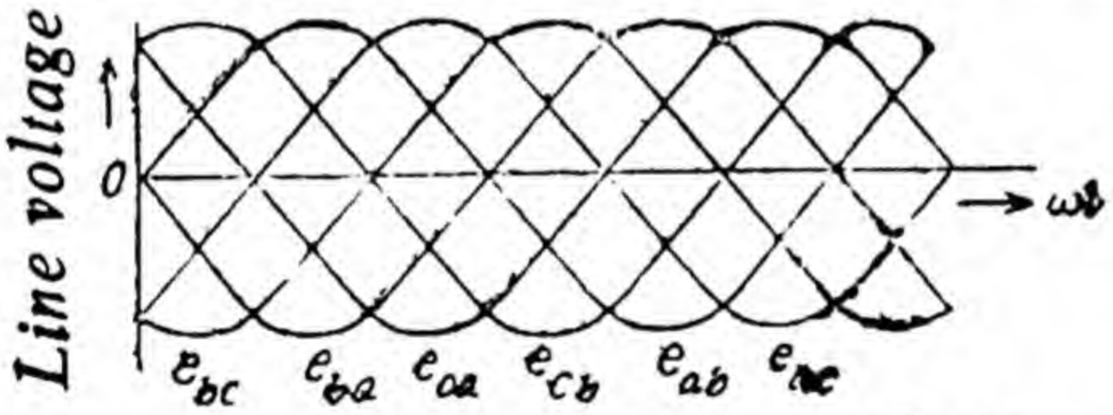
The other quantities as $I_{d.c.}$, $I_{r.m.s.}$ and γ can be calculated in a similar way as in case of 3 phase half wave rectifier system. The only difference is that here the valve conducts for 1/6 th of a cycle while in three phase system each valve conducts for 1/3 rd of the cycle.

8·7-3. THREE PHASE-FULL WAVE SYSTEM :

The circuit diagram of a three phase-full wave system is shown in figure (18*a*). In this rectifier two tubes always conduct as in case of a single phase bridge rectifier and hence this is sometimes



(*a*) Three phase-Full wave rectifier circuit.



(*b*) Output voltage waveform.

Fig. 18. Three phase—Full wave system.

known as three phase bridge rectifier. During positive halves of the phase voltages, the tubes V_4, V_5 and V_6 conduct and produce positive output voltage with respect to neutral. During negative halves of the phase voltages, the tubes V_1, V_2 and V_3 conduct and produce negative output voltage with respect to neutral. The two outputs are added in series and the average value of the output voltage is twice that of either taken alone. Fig. (18 *b*) shows the waveforms of voltages.

It is seen from fig. (18*b*) that when line to line voltage e_{ab} is at its maximum positive value, the current flows through V_4, the load and valve V_2. After half cycle, e_{ab} is at its maximum negative value and the current flows through V_5, the load and valve V_1. In this way different pairs of tubes conduct during the positive and negative half of the line to line voltage. The direction of current through the load in every case is from *A* to *B*. Similarly, different

other pairs of tubes conduct for line voltages e_{ob} and e_{ao}. If the tube drops are considered negligible, the load voltage is always equal to the most positive voltage.

8·8. CONTROLLED RECTIFIERS

Often in many industrial control devices, it is required to control the d.c. current in load. This can be achieved by the following two ways : (i) by providing taping on the input transformer of the rectifier, and (ii) by inserting a controlling resistor in the output circuit. The first method is not desirable when high load currents are involved moreover this is not an economical method The second method increases the losses and gives poor efficiency. A most efficient and inexpensive method of controlling the output current of a rectifier is by using thyratrons, ignitrons, excitrons or cold cathode gas triodes instead of diodes for rectification. The current in the load can be controlled by controlling the conduction angle at our will. This is not possible in case of diodes, however the conduction angle can be controlled without much expense of energy in thyratrons or ignitrons. There is no wear and tear of mechanical parts (hence long life) in thyratrons or ignitrons.

8·9. THYRATRON CONTROLLED RECTIFIER

In a thyratron controlled rectifier circuit, the load is fed through an a.c. source in series with a thyratron as shown in figure (19). To analyse the action of a thyratron in a controlled circuit, the *first step* is to study the firing characteristics of the thyratron *i.e.* a curve showing the critical grid voltage for different values of anode voltage. The left side of figure (20*a*) shows the firing characteristic of a negative controlled thyratron. For different values of anode voltage, this curve gives the minimum grid voltage for conduction to take place. The *second step* is to find the critical grid voltage from this firing characteristic for different points on the a.c. cycle of the plate voltage. The *third step* is to plot this critical grid voltage as a function of time as shown in fig. (21*a*) for negative and in fig. (21*b*) for a positive control thyratron tubes.

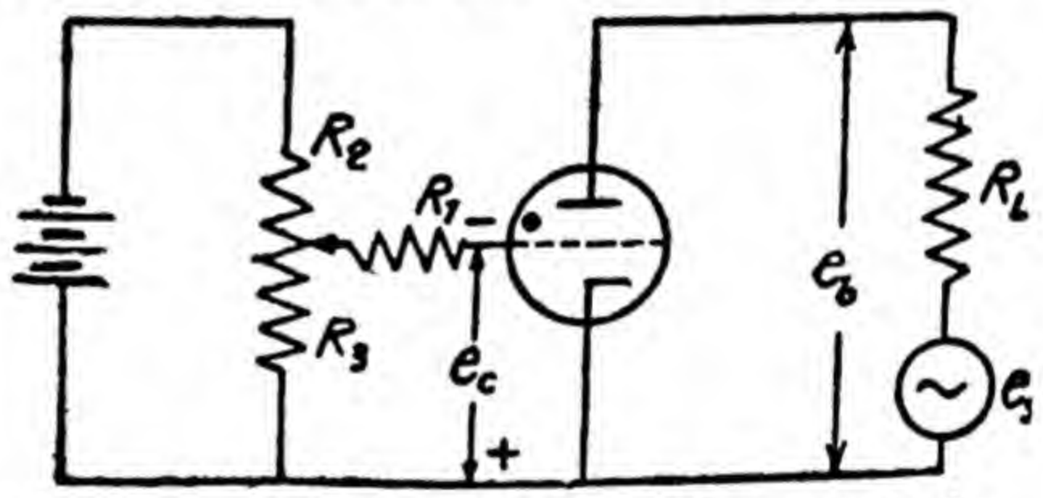


Fig. 19. Thyratron rectifier circuit.

At the instant t_1, the anode voltage is e_{b1} and the corresponding critical grid voltage as obtained from the characteristic is e_{c1} ; thus e_{c1} is taken just below t_1. Similarly, at instant t_2, the anode voltage is e_{b2} and the corresponding critical grid voltage as obtained from characteristic is e_{c2} ; therefore e_{c2} is taken just below t_2. In this way several points are noted. The curve joining these points is called as critical grid voltage curve. This curve signifies as a function of time the grid voltage that will just allow the tube to fire if the anode voltage follows the curve e_s. Figure (21) shows ths potential and current wave shapes of a thyratron for various

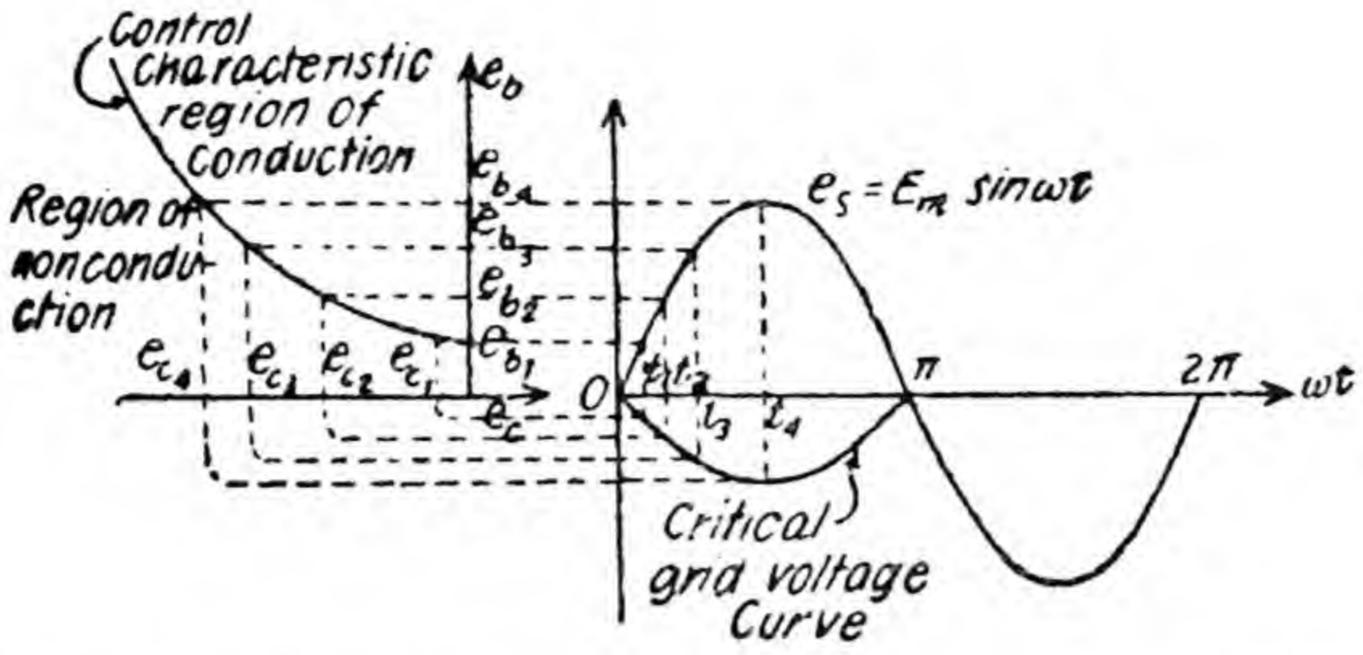


Fig. 20. (*a*) Critical grid voltage curve for a negative control thyratron.

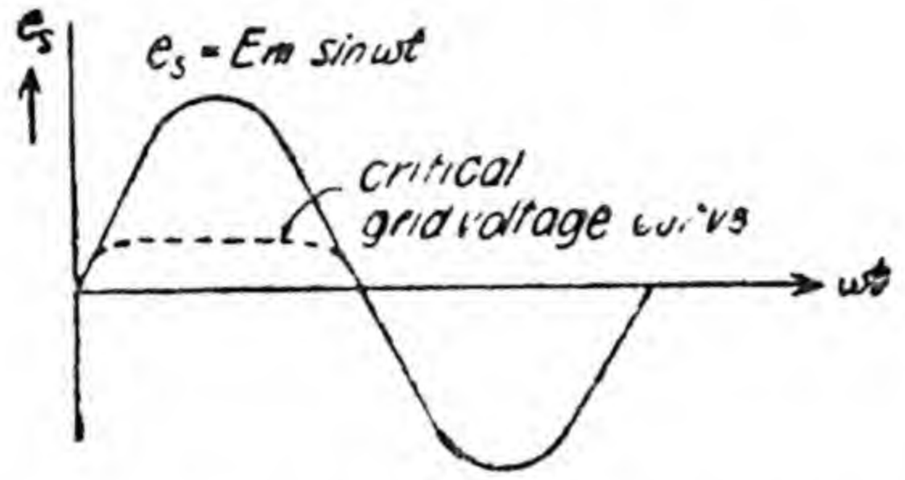
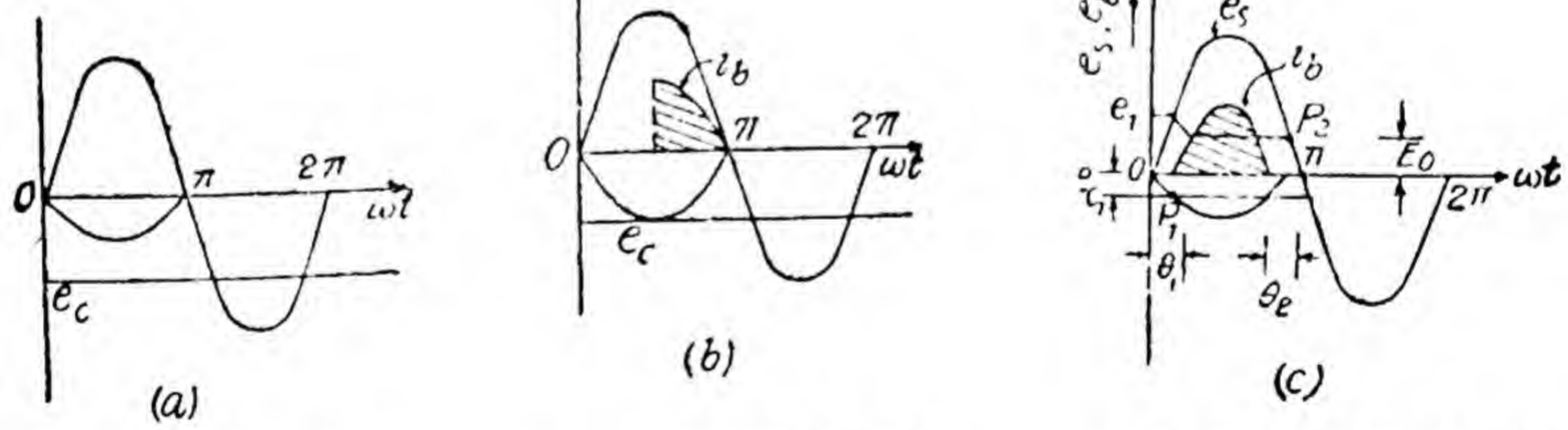


Fig. 20. (*b*) Critical grid voltage curve for a positive control thyratron.

values of grid bias. When e_c is more negative than the critical grid voltage, the tube does not fire at all as shown in fig (21 *a*). If the grid bias is reduced towards zero, then at a particular time, the grid bias is equal to the critical value and the tube starts conducting as shown in fig. (21*b*). This passes a current until the plate voltage falls to zero. The plate current is represented by a quarter sinusoids As the bias voltage is further reduced [fig. 21*c*],



e_c blow critical grid voltage.

e_c = critical grid voltage.

bias voltage is further reduced.

Fig. 21. Wave shapes illustrating the d.c control of thyratron

the tubes fires earlier in the cycle and plate current increases. The conduction is shown by a point P_1. Now the grid loses control and regains control only when the arc is extinguished. Corresponding to point P_1, the conduction angle is θ_1. When conduction starts, plate voltage e_b drops from e_1 to E_0, where E_0 is the constant voltage drop of the tube. The current rises from zero suddenly and varies according to the variation of applied voltage. At angle

$(\pi-\theta_2)$, the applied plate voltage is just equal to E_0, the conduction ceases. During the negative half cycle, no conduction takes place.

Mathematical analysis of thyratron rectifier : The analysis is based on the following assumptions :

(*i*) The tube fires at an angle θ_1 corresponding to grid bias e_{c1} (fig. 21 *c*).

(*ii*) Voltage across the tube after firing falls to a low value E_0 and remains practically constant till the tube conducts.

(*iii*) Tube extinguishes when the a.c. applied voltage falls to E_0. The extinction angle is then $(\pi-\theta_2)$.

Let the applied voltage in the plate circuit be $e_s=E_m \sin \omega t$. If the tube drop after the conduction has begun is E_0. then the current in the plate load resistance R_L during the conduction portion of the cycle is

$$i_b=\frac{E_m \sin \omega t-E_0}{R_L},\ \theta_1 < \omega t < \pi-\theta_2$$

where E_m is the peak value of the applied potential, θ_2 is the phase angle at which conduction starts and $\pi-\theta_2$, is the angle at which conduction ceases. After this the tube current remains zero until the phase angle θ_1 is again reached in the next cycle. The average rectified current which will be read on a *d.c.* ammeter is given by

$$I_{d.c.}=\frac{1}{2\pi}\int_{\theta_1}^{\pi-\theta_2} i_b\, d(\omega t)=\frac{1}{2\pi}\int_{\theta_1}^{\pi-\theta_2}\frac{(E_m \sin \omega t-E_0)}{R_L}\, d(\omega t)$$

$$=\frac{E_m}{2\pi R_L}\int_{\theta_1}^{\pi-\theta_2}\left(\sin \omega t-\frac{E_0}{E_m}\right) d(\omega t)$$

$$=\frac{E_m}{2\pi R_L}\left[-\cos \omega t-\frac{E_0}{E_m}\omega t\right]_{\theta_1}^{\pi-\theta_2}$$

$$=\frac{E_m}{2\pi R_L}\left[\cos \theta_1+\cos \theta_2-\frac{E_0}{E_m}(\pi-\theta_2-\theta_1)\right], \quad \ldots (1)$$

where θ_2 is defined by the relation that when $\pi-\theta_2=\omega t$,

$E_m \sin(\pi-\theta_2)=E_0$; $E_m \sin \theta_2=E_0$ or $\sin \theta_2=E_0/E_m$ ...(2)

If the ratio E_0/E_m is very small, then E_0/E_m and θ_2 may be taken zero, and equation (1) reduces to

$$I_{d.c.}=\frac{E_m}{2\pi R_L}(1+\cos \theta_1) \quad \ldots (3)$$

It is clear from this expression that $I_{d.c.}$ can be controlled by varying ignition angle θ_1 from 0 to π and that d.c. current is maximum when θ_1 is zero, being equal to $E_m/\pi R_L$. Now equation (3) can be written as

$$\frac{I_{d.c.}}{(I_{d.c.})_{max}}=\frac{1+\cos \theta_1}{2} \quad \ldots (4)$$

We see from equations (3) and (4) that $I_{d.c.}$ can be varied by varying angle θ_1 *i.e.*, by varying the position in the a.c. cycle at which the d.c. grid potential exceeds the critical grid voltage. For the large positive grid voltage, the conduction may take place over almost complete half cycle. For zero grid voltage, conduction takes place over a large portion of the half cycle and hence $I_{d.c.}$ has a large value. If the grid voltage is made more negative, conduction does not take place at all and $I_{d.c.}$ is then zero.

Average value of anode voltage is given by

$$E_{d.c.}=\frac{1}{2\pi}\int_0^{2\pi} e_b\, d(\omega t)$$

$$=\frac{1}{2\pi}\left[\int_0^{\theta_1} E_m \sin \omega t\, d(\omega t)+\int_{\theta_1}^{\pi-\theta_2} E_0\, d(\omega t)+\int_{\pi-\theta_2}^{2\pi} E_m \sin \omega t.d(\omega t)\right]$$

where $e_b=E_m \sin \omega t \quad 0 < \omega t < \theta_1$

$=E_0 \quad \theta_1 < \omega t < \pi-\theta_2$

$=E_m \sin \omega t \quad \pi-\theta_2 < \omega t < 2\pi.$

$$=\frac{1}{2\pi}[-E_m \cos \theta_1+E_m+E_0(\pi-\theta_2-\theta_1)+E_m \cos(\pi-\theta_2)-E_m \cos 2\pi]$$

$$=\frac{E_0}{2\pi}(\pi-\theta_2-\theta_1)-\frac{E_m}{2\pi}(\cos \theta_1+\cos \theta_2) \qquad ...(5)$$

If $E_m \gg E_0$, then equation (5) reduces to

$$E_{d.c.}\approx-\frac{E_m}{2\pi}(1+\cos \theta_1) \qquad ...(6)$$

The negative sign in the above expression shows that the cathode is more positive than the plate and the positive terminal of voltmeter must be connected to the cathode in order to read in the positive direction.

The r.m.s. value of plate current is given by

$$I_{r.m.s.}=\sqrt{\left[\frac{1}{2\pi}\int_0^{2\pi} i_b^2\, d(\omega t)\right]}$$

$$=\left[\frac{1}{2\pi}\int_{\theta_1}^{\pi-\theta_2}\left(\frac{E_m \sin \omega t-E_0}{R_L}\right)^2 d(\omega t)\right]^{1/2} \qquad ...(7)$$

The power supplied to the plate circuit is given by

$$P=\frac{1}{2\pi}\int_0^{2\pi} e_b\, i_b\, d(\omega t)$$

$$=\frac{1}{2\pi}\int_{\theta_1}^{\pi-\theta_2} E_m \sin \omega t.\left(\frac{E_m \sin \omega t-E_0}{R_L}\right) d(\omega t) \qquad ...(8)$$

8·10. PHASE SHIFT CONTROL :

One of the most satisfactory methods of varying the current through a thyratron is known as phase shift control. In this method *the conduction point of the cycle is controlled by controlling the phase angle between the plate and the grid potentials.* This is achieved by the application of an a.c. voltage to the plate as well as to the grid. Both the a.c. plate voltage and a.c. grid voltage are taken from some a.c. voltage source so that they have the same frequency. The phase angle between the two voltages is varied so that the grid voltage lags behind the plate voltage. Fig. (22) shows how the conduction angle of a thyratron is affected by the phase angle between the plate and the grid voltages. From fig. (22*a*) it is observed that when the two voltages are in phase, the tube conducts for 180° and the direct current through the thyratron is maximum, *i.e.*,

$$I_{d.c.}=\frac{E_m}{\pi\ R_L}.$$

When the grid voltage lags behind the plate voltage by an angle θ_1, the current is given by

$$I_{d.c.}=\frac{E_m}{2\pi\ R_L}\ (1+\cos\ \theta_1).$$

Figs. 22 (*b*) and 22 (*c*) show two cases of phase difference and their corresponding conduction angles.

When the grid voltage leads the plate voltage fig. 22 (*d*) the tube conducts for 180° and the grid does not have any control on the current flow.

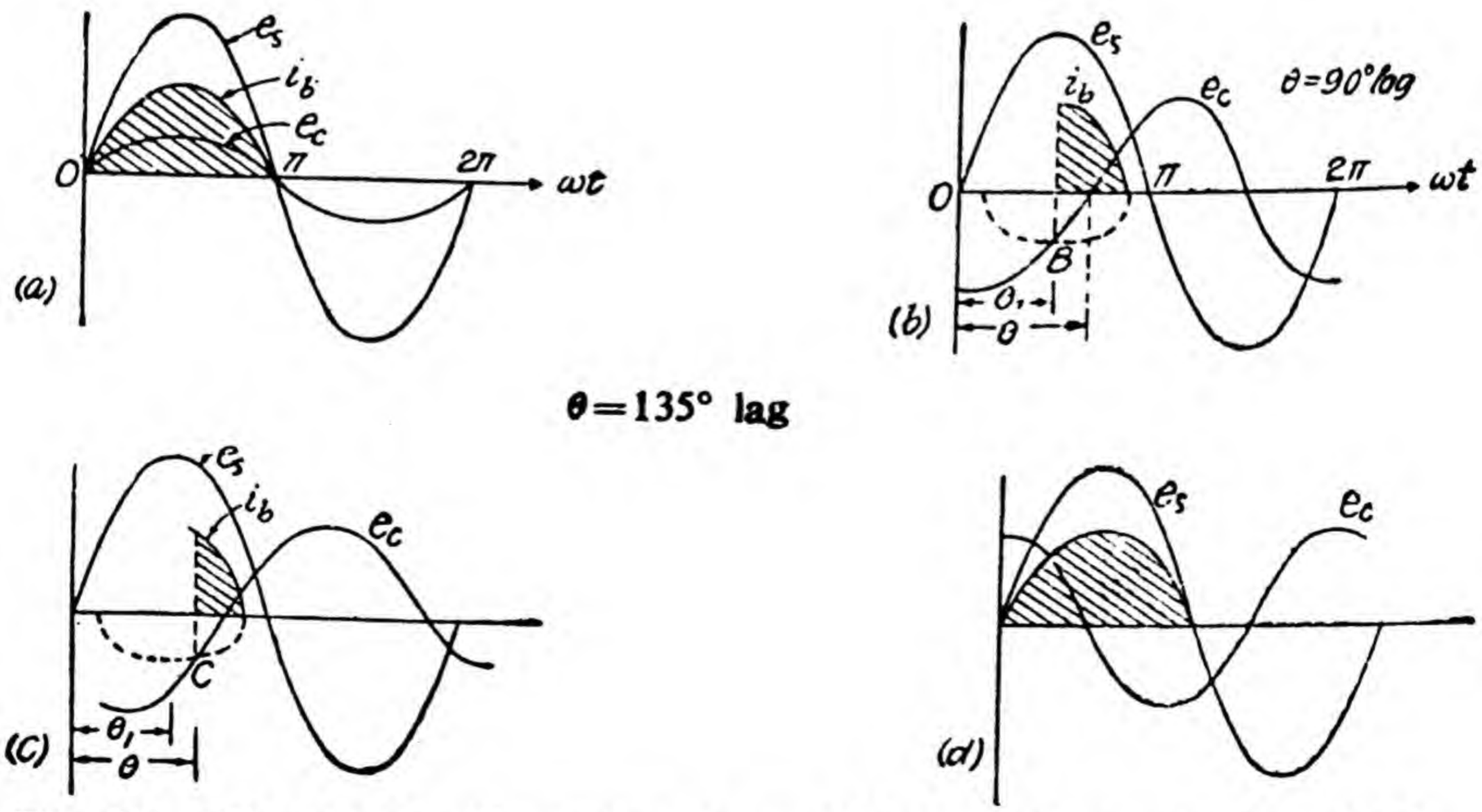


Fig. 22. Waveshapes illustrating phase shift control (*a*), (*b*), (*c*) and (*d* show various cases of phase difference and the corresponding angle

Figure (23) shows a circuit in which a variable phase can be obtained. The phase shift between grid and plate voltage is controlled by the impedances Z_1 and Z_2. Using a centre tapped transformer voltages E and E' are equal in magnitude and phase voltage E drives current through thyratron plate circuit. The phase shift of the grid supply voltage may be adjusted by varying the magnitude of either Z_1 or Z_2. A vector diagram may be drawn to show phase and magnitude relations of various voltages and currents.

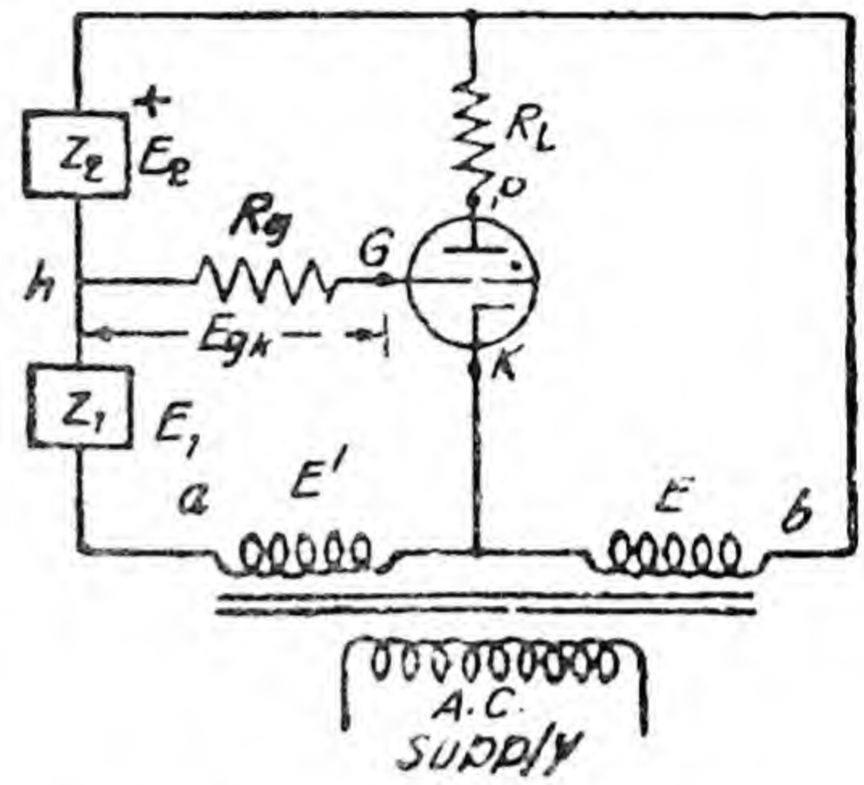
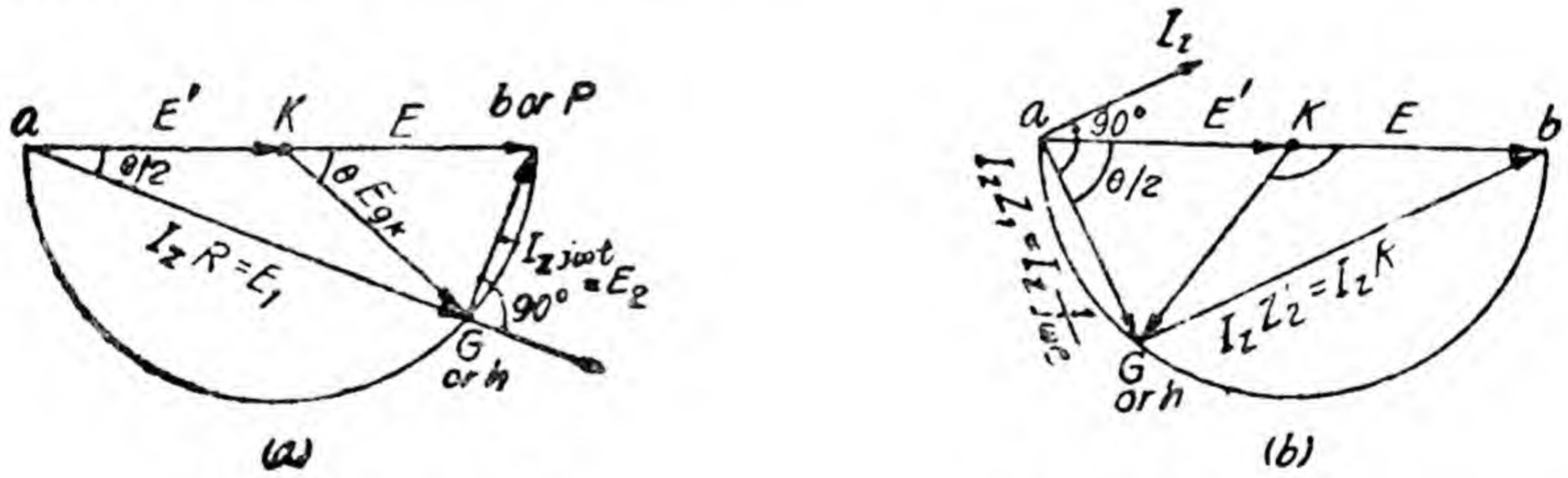


Fig. 23. Simple phase-shifting network

Figure (24 *a*) shows the phaser diagram for the combination Z_1 equal to R and Z_2 equal to $j\omega L$. Here the current I_2 through the impedances Z_1 and Z_2 lags behind the applied voltage $(E+E')$. On resolving the voltage sum $(E+E')$ into two components across the impedances Z_1 (R) and Z_2 $(j\omega L)$, a right angle triangle is oqtained as a phasor diagram. Now joining the centre K of the voltage vector $(E+E')$ to the vertex of the triangle, the grid supply voltage vector E_{gk} can be obtained. This lags behind the vector E by an angle θ. It is obvious from the figure that the vertex G of the phasor diagram always lies on the circle on vector $(E+E')$ as the diameter. The magnitude of the vector E_{gk} is, therefore, always constant and equal to E.



Phasor diagram when $Z_1=R$ and $Z_2=j\omega L$ — Phasor diagram when $Z_2=1/j\omega C$ and $Z_2=R$

Fig. 24. Phasor diagrams of phase shifting circuits.

The phase angle of E_{gk} with respect to E is given by

$$\theta=-2\tan^{-1}\frac{\omega L}{R}. \qquad \ldots(1)$$

Fig. (24 *b*) shows the phasor diagram for $Z_1=1/j\omega C$ and $Z_2=R$. In this case it is also found that E_{gk} has a constant magnitude equal to E and a phase angle which is lagging and is given by

$$\theta=-2\tan^{-1}\omega CR \qquad \ldots(2)$$

θ may be varied by varying either C or R.

8·11. ON OFF CONTROL :

Figure (25) shows the circuit diagram of a thyratron rectifier having on-off control. When the switch S is open the thyratron is biased negatively to such a value that the discharge can not be started at any time dur ng the cycle. In this case the negative grid bias E_{oo} is greater than the critical grid voltage. When the switch is closed, the grid voltage becomes almost equal to the plate voltage. The thyratron now fires and continue to conduct until the end of the half cycle. It fires at the beginning of every positive half cycle as long as the switch remains closed.

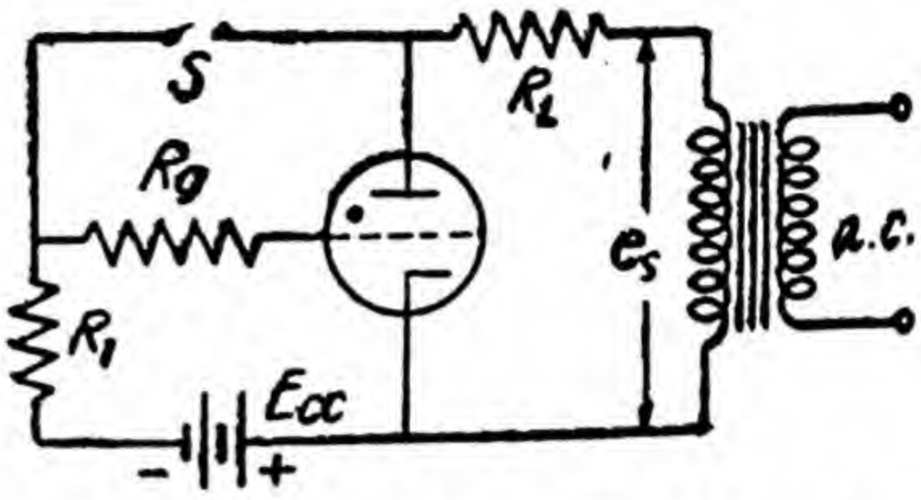


Fig. 25. On-off thyratron control.

8·12. IGNITRON RECTIFIERS AND IGNITRON CONTROL CIRCUITS :

The formation of an arc in a thyratron is delayed depending upon the grid voltage; on the other hand, the arc is initiated in ingnitron by passing the current through the ignitor. When large power is required, the ignitron rectifiers are used because there is no limitation of the supply of electrons from mercury cathode. Various methods have been devised for applying the ignitor current pulse depending upon the use to which ignitor is put. When this is used to rectify an alternating plate supply voltage, the ignitor pulses must be periodic and must be synchronised with the a.c. anode voltage in such a way that ignition of main arc takes place at the same phase angle in every plate voltage cycle. The output is controlled by controlling the time phase of firing or igniting. A few common circuits are briefly discussed below :

(1) Anode ignition with diode : The circuit arrangement using ignitron and its firing arrangement is shown in fig. (26). A gas filled thermionic diode T_2 is connected in the ignitor circuit This serves two functions : (i) it supplies the ignitor current and (ii) prevent reverse current through the ignitor during the negative half cycle of anode voltage. For the positive half cycle of the applied voltage, when the voltage becomes equal to the breakdown potential of gas filled tube T_2, the tube begins to conduct. Now a current flows through ignitor and the mercury pool. As the supply voltage increases, the current through T_2 continuously increases until it reaches a critical value at which the main arc in T_1 starts. When T_1 conducts, a heavy current flows through the load

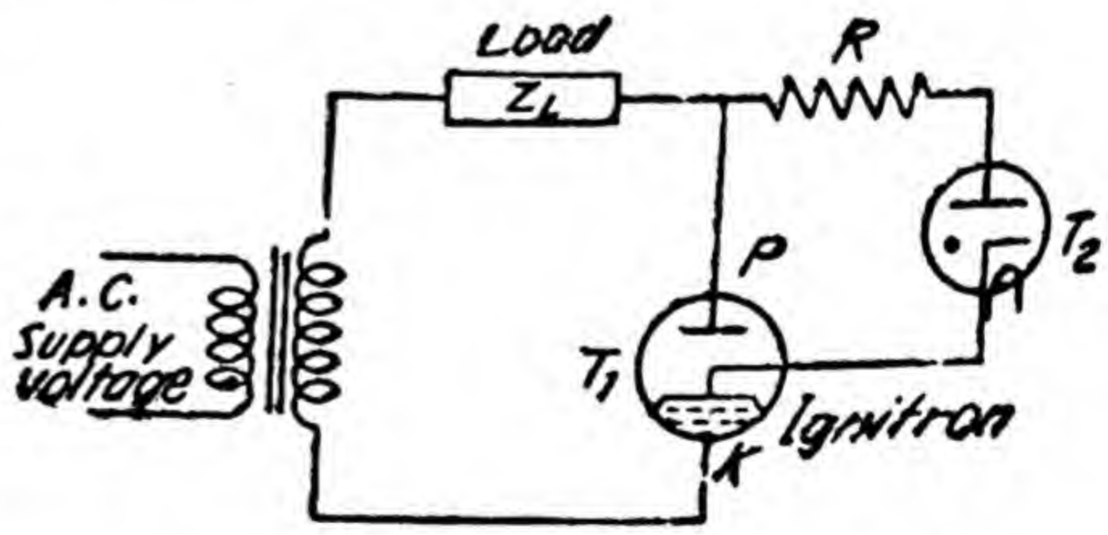


Fig. 26. Ignitron as a half wave rectifier.

Z_L and the voltage applied to T_2 and ignitor decreases to a low constant value. The ignitor current becomes negligibly small. During negative half cycle, the tube T_1 stops conducting and no current flows through the load. The process is again repeated for positive half cycle.

(2) Anode ignition with thyratron : A circuit arrangement of anode ignition with thyratron is shown in fig. (27). In this case diode T_2 is replaced by a thyratron. e_{sg} is the control grid voltage of thyratron T_2. This is an alternating voltage of the same frequency as anode voltage but delayed in phase.

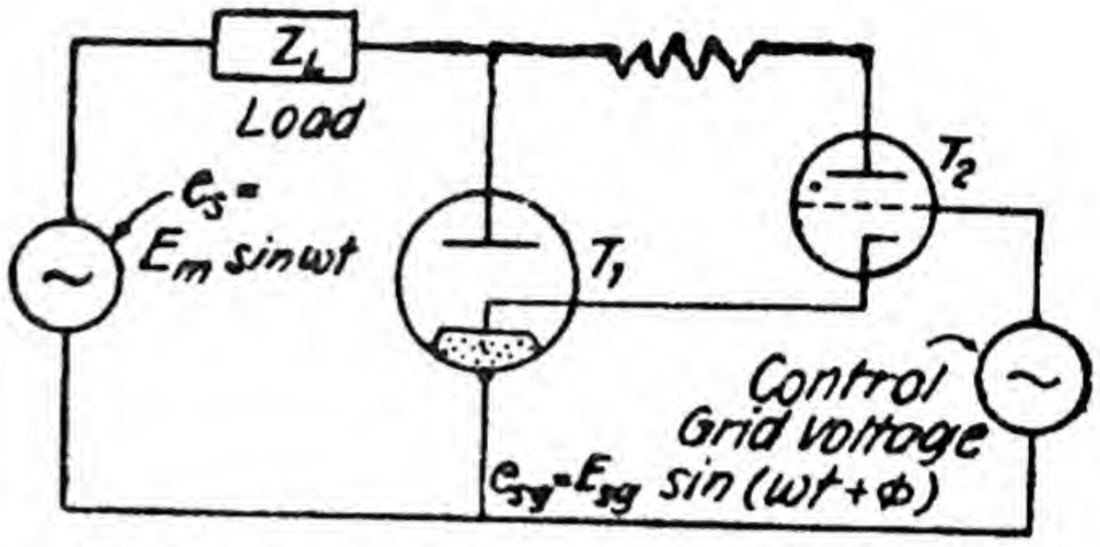
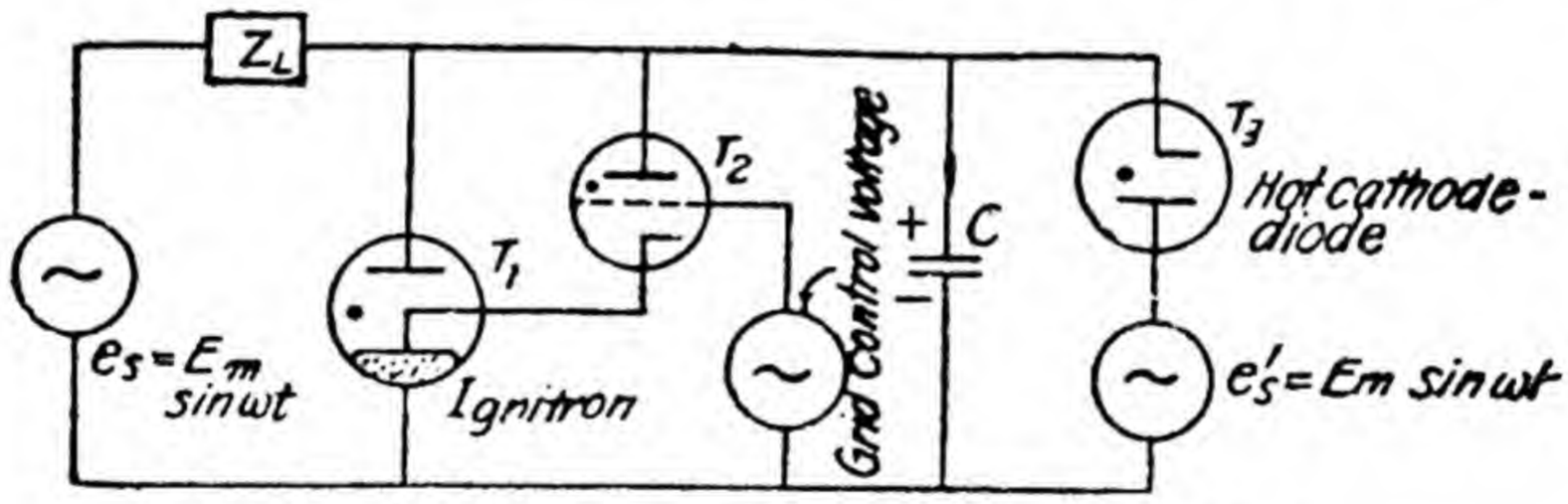


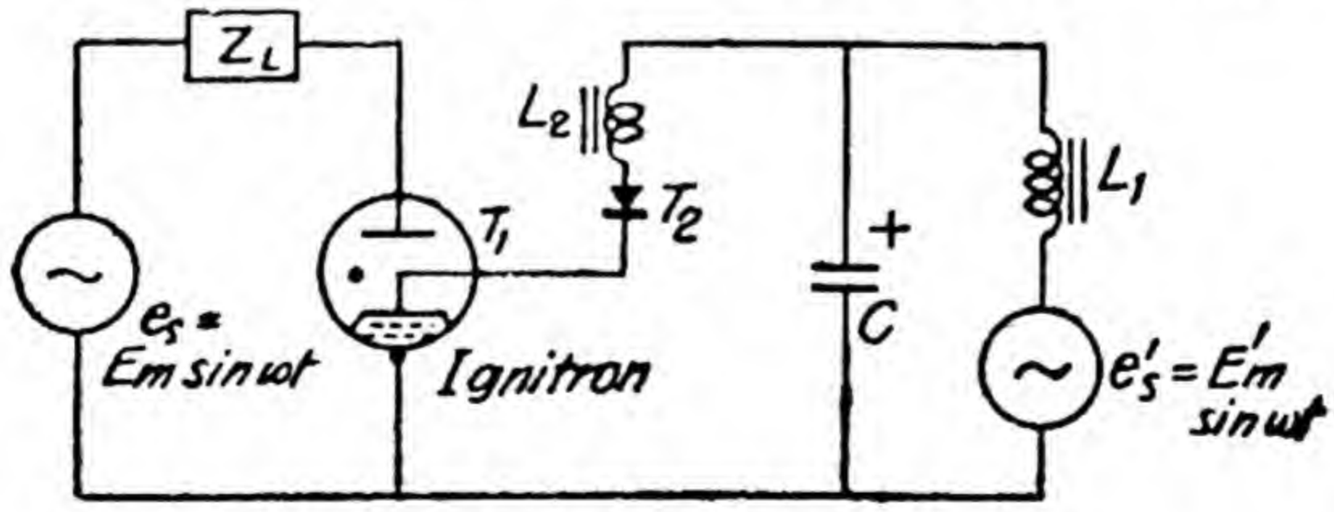
Fig. 27. Anode ignition with thyratron.

The ignition of thyratron is thus delayed depending upon the delay angle ϕ. There are two advantages of the above two methods : (i) Ignition cannot take place at the beginning of the cycle. (ii) Ignitor current flows through the load impedance Z_L restricting its suitability for relatively heavy load currents only. These defects are avoided in the following two methods.

(3) Thyratron and capacitor ignition : The circuit arrangement is shown in fig. (28 a). Whenthe positive half cycle starts, the capacitor C is gradually charged through the hot cathode gas diode T_3. As soon as the anode voltage of the tube T_2 reaches a definite value, the condenser discharges through the thyratron tube T_2. Now a sudden current through ignitor causes ignition in ignitron. The average value of rectifier current can be controlled by adjusting the grid control voltage.



(*a*) Thyratron and capacitor ignitron.



(*b*) Saturation reactor ignitron.

Fig. 28. Ignitor excitation circuits with no ignitor current through load.

(4) Saturating-reactor ignition : The circuit arrangement is shown in figure (28 *b*). For the positive half cycle of the applied voltage, the condenser C charges through L_1. Simultaneously, current gradually builds up in the circuit containing reactor L_2, tube T_2 and ignitor T_1. The core of L_2 is so designed that as soon as the current through the reactor reaches a critical value, the core gets saturated suddenly, *i.e.* its inductance decreases. Now the condenser C discharges suddenly through the circuit containing L_2, T_2 and ignitor. This causes an ignition. The purpose of rectifier T_2 is to prevent the flow of current in reverse direction.

8·13. POWER SUPPLY :

The circuit of a power supply using full wave rectifier is shown in fig. 29.

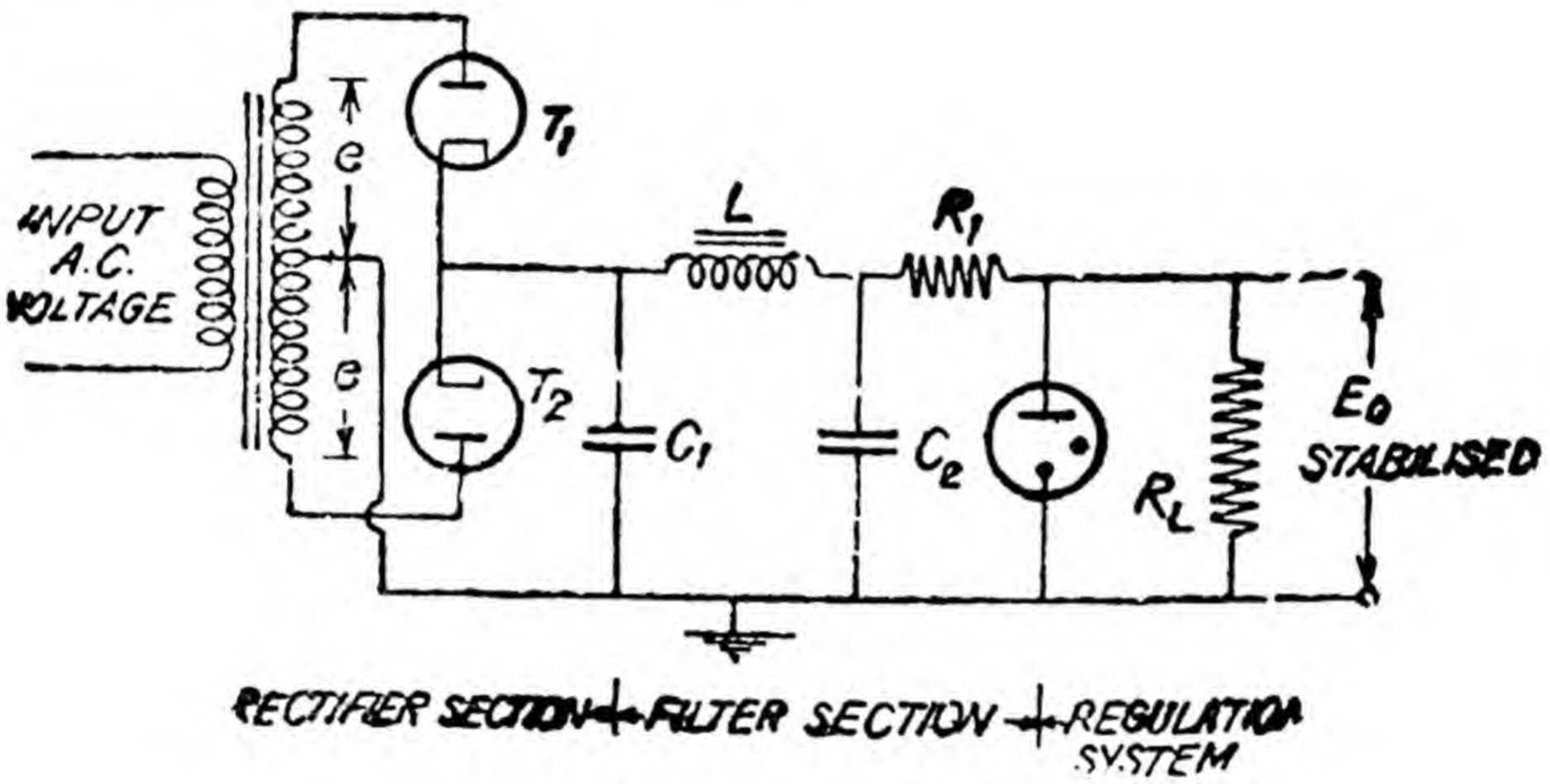


Fig. 29. Power supply unit.

The power supply system includes the following sections :

1. Rectifier section
2. Filter section
3. Regulation system
4. Bleeder resistor.

We shall now describe the different sections.

The rectifier section consists of two vaccum diodes (if full wave rectifier) or only one diode (if half wave rectifier) along with a transformer. The performance of half wave and full wave rectifiers has been discussed in the previous articles.

FILTER SECTION

We shall discuss this section in the light of following points :

(*a*) Ripple factor,

(*b*) Output voltage, and

(*c*) Voltage regulation.

The effect of output resistor on the above points will also be considered.

The output of a rectifier contains a.c. component of considerable magnitude. The effect of this a.c. component is to vary the output d.c. voltage. The filter system is used to reduce the magnitude of this ripple and to provide a regulated and constant voltage. The following filters are used :

(i) **The series inductor filter :** The rectifier with series inductor is shown in fig. 30.

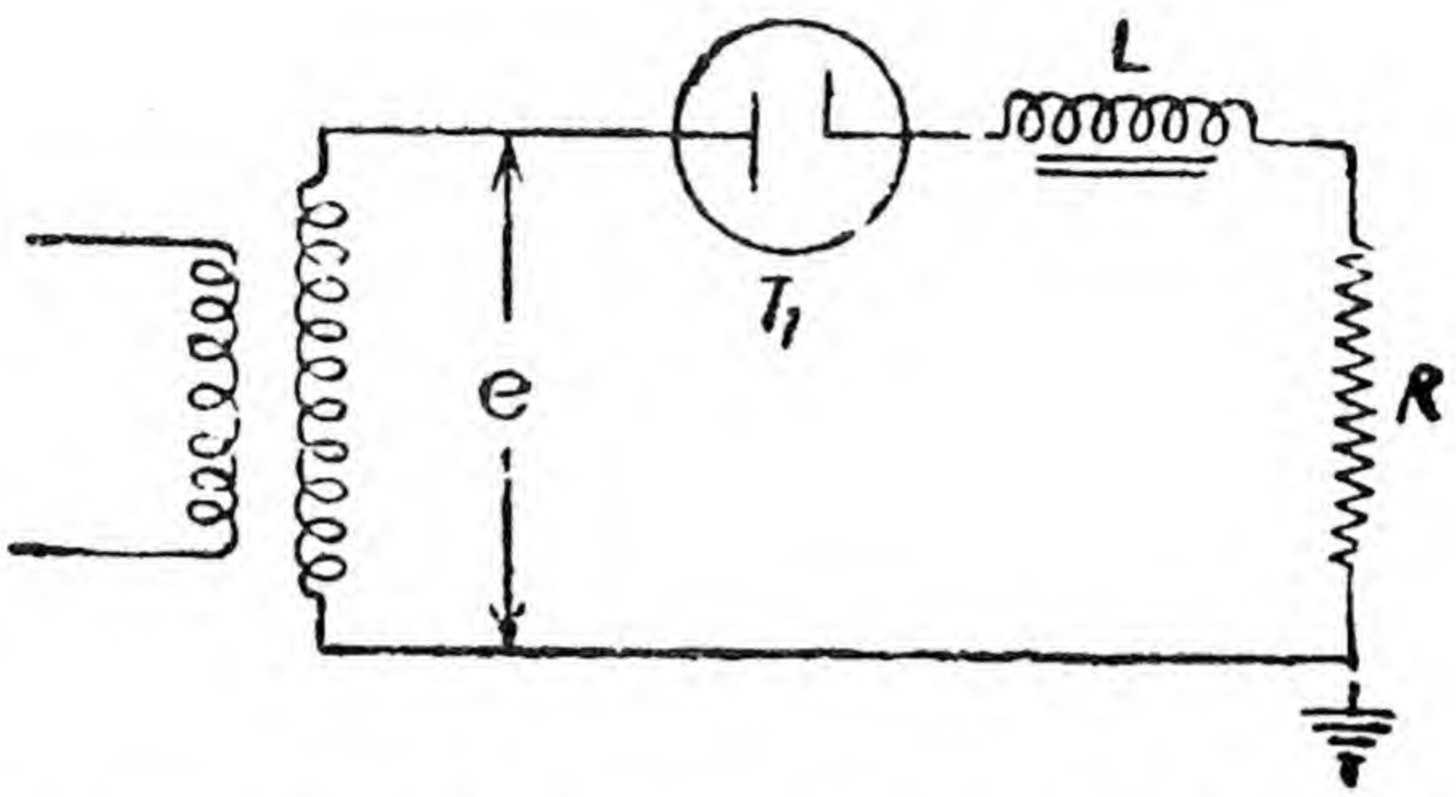


Fig. 30. Rectifier with series inductor filter.

The inductor in series with load serves as a filter. The operation of the inductor filter depends upon the property of the inductance to oppose any change of current that may flow through it. During the positive half cycle of the input voltage, the inductance of the choke retards the rise of current through the rectifier. After the impressed voltage reaches its positive peak, the energy stored in the magnetic field will continue the d.c. pulse even after the applied voltage becomes negative. Thus the conduction takes place for more than half a cycle. As the inductor L offers an extremely high impedance to alternating component of current in the diode output, the amplitude of a.c. component is reduced considerably. Since the d.c. resistance of large inductance L is very small compared to R, the d.c. component is not reduced at all.

If the input voltage be $e = E_m \sin \omega t$, then the differential equation for the conduction interval is

$$L\frac{di_b}{dt} + Ri_b = E_m \sin \omega t, \qquad \ldots(1)$$

where i_b is the diode current and R is the load.

The equation (1) can be written as

$$\frac{di_b}{dt} + \frac{Ri_b}{L} = \frac{E_m}{L} \sin \omega t. \qquad \ldots(2)$$

The solution of equation (2) consists of two parts : (*i*) complementary function, (*ii*) particular integral. The complementary

function can be obtained by putting the right hand side of equation (2) equal to zero and solving it for the current i_b, we have

$$i_b = Ae^{-Rt/L}, \qquad \ldots(3)$$

where A is constant.

The particular integral can be obtained by solving equation (2) as it is, we have

$$i_b = \frac{E_m}{(R^2+\omega^2L^2)^{1/2}} (\sin \omega t - \phi), \qquad \ldots(4)$$

where $$\phi = \tan^{-1}\frac{\omega L}{R}.$$

Therefore, the diode current is

$$i_b = Ae^{-Rt/L} + \frac{E_m}{(R^2+\omega^2L^2)^{1/2}} \sin (\omega t - \phi),\ 0 \leqslant \omega t \leqslant \theta_2 \qquad \ldots(5)$$

$i_b = 0$ $\qquad \theta_2 \leqslant \omega t \leqslant 2\pi$

where θ_2 is the ignition angle (conduction angle).

Assuming that the positive cycle begins at $t=0$, *i.e.* $i_b=0$ at $\omega t=0$ the value of the constant A can be calculated. Applying this condition in equation (5), we have

$$0 = A + \frac{E_m}{(R^2+\omega^2L^2)^{1/2}} \sin (-\phi)$$

$$\therefore \quad A = \frac{E_m}{(R^2+\omega^2L^2)^{1/2}} \sin \phi.$$

Substituting this value in equation (5), we get

$$i_b = \frac{E_m}{(R^2+\omega^2L^2)^{1/2}} \sin \phi\, e^{-Rt/L} + \frac{E_m}{(R^2+\omega^2L^2)^{1/2}} \sin (\omega t - \phi)$$

or $$i_b = \frac{E_m}{(R^2+\omega^2L^2)^{1/2}} [\sin (\omega t - \phi) + e^{-Rt/L} \sin \phi]. \qquad \ldots(6)$$

The first term has a sinusoidal nature and becomes zero for certain value of ωt. The second term representing the exponential component is always positive. For large L and small R, the exponential component decreases very slowly (as time constant L/R is quite large) and i_b may not be reduced to zero and it flows continuously. This occurs in case of full wave rectifier; moreover, in case of half wave rectifier, the current is cut off to zero.

The cut off angle θ_2 (greater than π) can be found by the condition that $\omega t = \theta_2$, the current $i_b = 0$. Applying this condition to equation (6), we have

$$0 = \frac{E_m}{(R^2+\omega^2L^2)^{1/2}} [\sin (\theta_2 - \phi) + e^{-Rt/L} \sin \phi]$$

or $\sin (\theta_2 - \phi) + e^{-Rt/L} \sin \phi = 0$ because $\frac{E_m}{(R^2+\omega^2L^2)^{1/2}} \neq 0$

or $\sin \theta_2 \cos \phi - \cos \theta_2 \sin \phi = e^{-Rt/L} \sin \phi$

$$\text{or} \quad \sin\theta_2 \cos\phi = \sin\phi\,(\cos\theta_2 - e^{Rt/L})$$

$$\text{or} \quad \sin\theta_2 \cos\phi = \sin\phi\,(\cos\theta_2 - e^{-R\theta_2/\omega L})$$

$$\text{or} \quad \cot\phi = \frac{\cos\theta_2 - e^{-R\theta_2/\omega L}}{\sin\theta_2}$$

$$\text{or} \quad \frac{R}{\omega L} = \frac{\cos\theta_2 - e^{-R\theta_2/\omega L}}{\sin\theta_2} \left(\because \quad \tan\phi = \frac{\omega L}{R}\right) \qquad ..(7)$$

If $R/\omega L$ is known, the value of θ_2 can be calculated with the help of equation (7). On getting the value of conduction angle θ_2 from equation (7) and the magnitude of current from equation (6), the general character of the solution of differential equation for the current can be depicted as a shown in fig. 31.

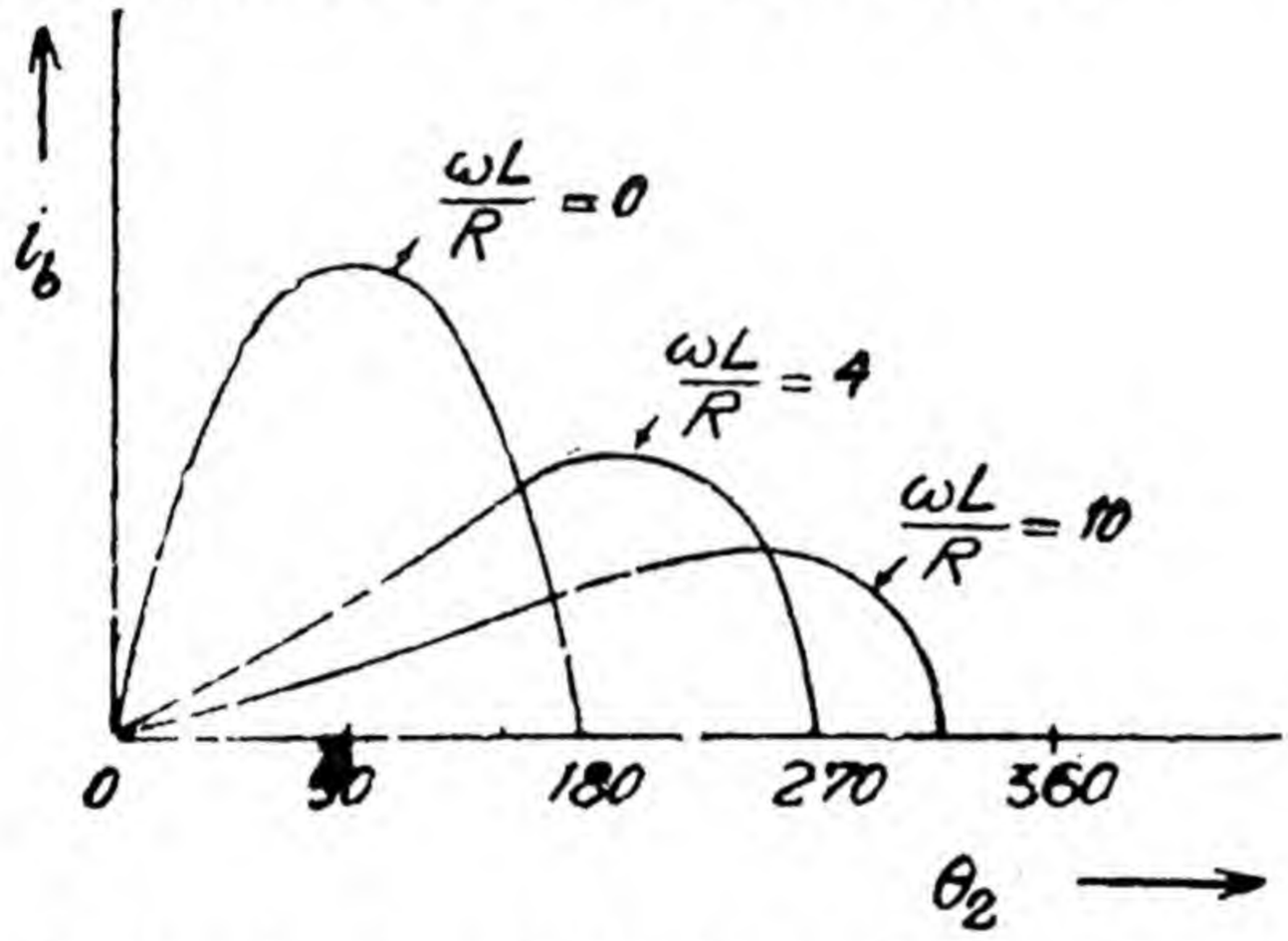


Fig. 31. Showing the effect of L on the waveform of the current on the conduction angle θ_2 in half wave rectifier with inductor filter.

The average value of direct current $I_{d.c.}$ can be calculated by averaging the current over a complete cycle *i.e.* over the interval 0 to θ_2 and θ_2 to 2π.

$$\therefore \quad I_{d.c.} = \frac{1}{2\pi}\int_0^{\theta_2} \frac{E_m}{(R^2+\omega^2L^2)^{1/2}} [\sin(\omega t - \phi) + e^{-R\theta_2/\omega L} \sin\phi]\, d(\omega t)$$

$$+ \frac{1}{2\pi}\int_{\theta_2}^{2\pi} d(\omega t)$$

$$= \frac{E_m}{2\pi R\left(1+\frac{\omega^2L^2}{R^2}\right)^{1/2}} \int_\theta^{\theta_2} [\sin(\omega t - \phi) + e^{-R\theta_2/\omega L} \sin\phi]\, d(\omega t)$$

If $\frac{\omega L}{R} \ll 1$, then $\frac{R}{\omega L}$ tends towards infinity so that $e^{-R\theta_2/\omega L} \to 0$.

$$\therefore \quad I_{d.c.} = \frac{E_m}{2\pi R\left(1+\frac{\omega^2L^2}{R^2}\right)^{1/2}} \int_0^{\theta_2} \sin(\omega t - \phi)\, d(\omega t)$$

$$= \frac{E_m}{2\pi R\left(1+\frac{\omega^2 L^2}{R^2}\right)^{1/2}} \int_0^{\theta_2} \sin \omega t \, d(\omega t)$$

because $\phi = \tan^{-1} \frac{\omega L}{R}$ and when $\frac{\omega L}{R} \ll 1$

$$\phi = 0.$$

Neglecting $\frac{\omega^2 L^2}{R^2}$ being very small, we get

$$I_{d.c.} = \frac{E_m}{2\pi R} \left[-\cos \omega t \right]_0^{\theta_2}$$

or

$$I_{d.c.} = \frac{E_m}{2\pi R} [1 - \cos \theta_2] \quad \ldots (8)$$

$$E_{d.c.} = I_{d.c.} \times R$$

$$= \frac{E_m}{2\pi} [1 - \cos \theta_2]. \quad \ldots (9)$$

$$\text{Again Efficiency} = \frac{E_{d.c.}}{E_m} \times 100$$

$$= \frac{(1 - \cos \theta_2)}{2\pi} \times 100. \quad \ldots (10)$$

Equations (8) and (9) show that as θ_2 approaches 360°, the d.c. output voltage and current are small in terms of E_m. For small value of L, θ_2 approaches 180°.

(ii) **Shunt Capacitor Filter :** In fig. 32, the rectifier with shunt capacitor filter is shown. The capacitor shunted across the load resistance acts as a simple filter.

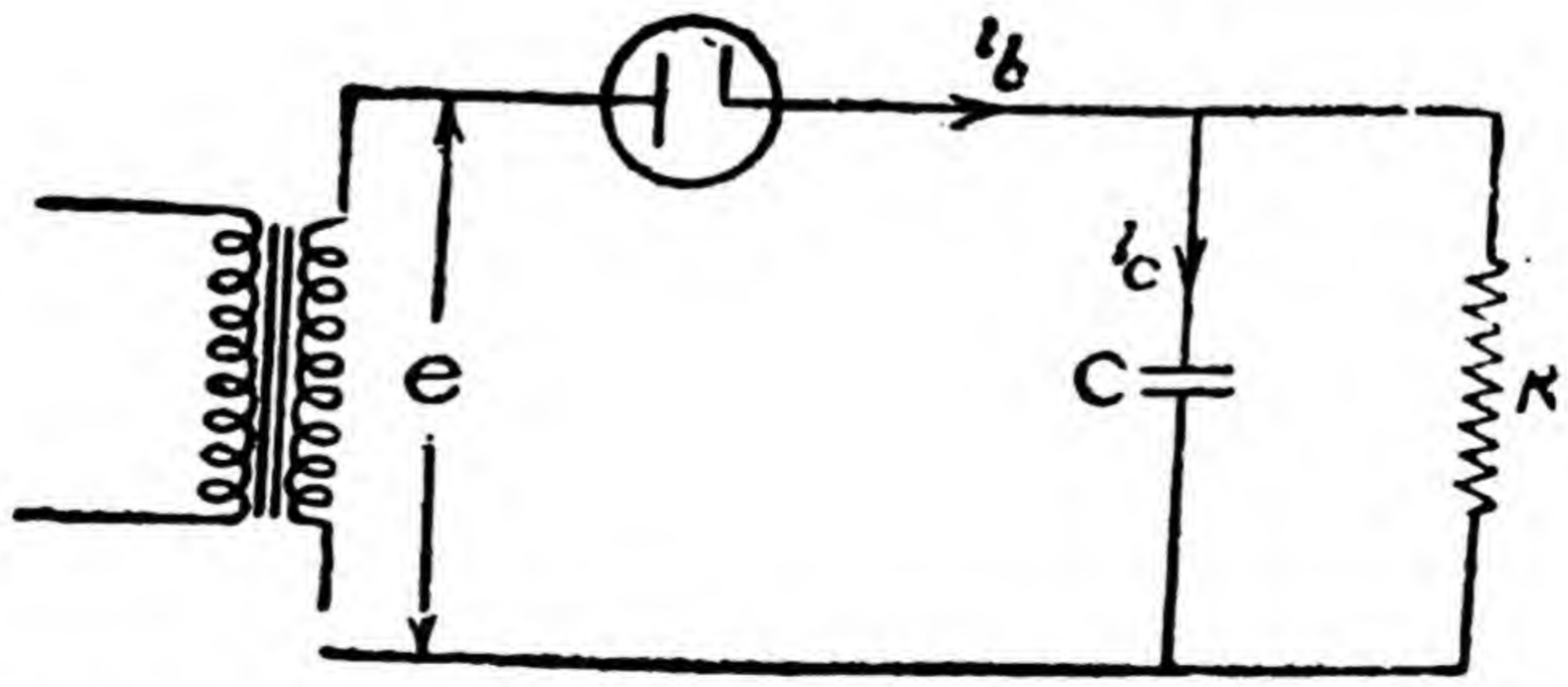


Fig. 32. Circuit of rectifier shunt capacitor filter.

As the direct current can not flow through capacitor, almost all the d.c. component flows through R and hence the direct output voltage is not reduced at all. The plate current flows when the diode conducts for the positive half cycle. This current charges the capacitor to nearly the peak value of the input voltage. During negative half cycle when the tube ceases to conduct, the capacitor

discharges through the output resistance and thus maintains the output voltage at a relatively high level through the input voltage cycle.

The waveform of currents (i_b=diode current, i_c=current through capacitor and i_R=current through resistor R) are shown in fig. 33.

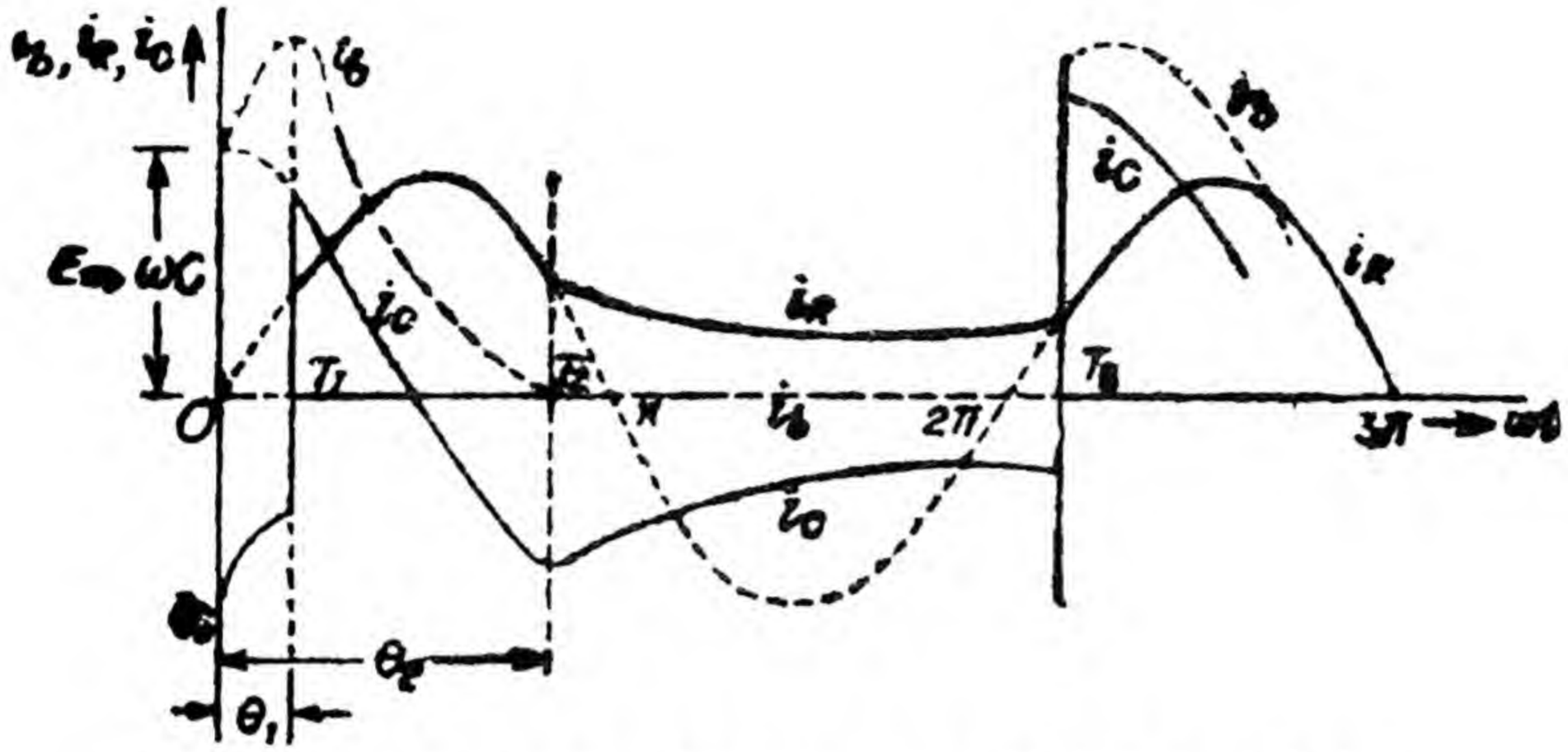


Fig. 33. Waveform of currents in half wave rectifier.

Ignition angle θ_1 corresponds to start of conduction period. Extinction angle θ_2 corresponds to the end of conduction period. It is observed from the figure that from the instant $\omega t_1=\theta_1$ to the instant $\omega t_2=\theta_2$ the diode conducts, while it does not conduct from the instant ωt_2 to $(2\pi+\omega t_1)$. The capacitor C now discharges its voltage through R. We shall obtain the expressions for the diode.

The voltage across the capacitor during the charging interval, ωt_1 to ωt_2, is equal to supply voltage, *i.e.*,

$$e_c=e=E_m \sin \omega t,\ \omega t_1 < \omega t < \omega t_2. \qquad \ldots(1)$$

The current i_R through the resistance during the same interval is given by,

$$i_R=\frac{E_m \sin \omega t}{R},\ \omega t_1 < \omega t < \omega t_2,$$

and the current i_c through the capacitor,

$$i_c=C\frac{de_c}{dt}\ \omega C\ E_m \cos \omega t,\ \omega t_1 < \omega t < \omega t_2.$$

The diode current i_b, during tne conduction interval is the sum of the two currents, *i.e.* $i_b=i_c+i_R$.

$$\therefore \quad i_b=E_m\left[\frac{\sin \omega t}{R}+\omega C \cos \omega t\right] \qquad \ldots(2)$$

At the conduction ends, the diode current is zero, *i.e.*

$$i_b=i_c+i_R$$
$$i_c=-i_R \text{ when } t_b=0,$$
$$\omega C \cos \omega t_2=-\frac{\sin\ \omega t_2}{R} \text{ (as } t=t_2 \text{ at conduction ends).}$$

$$\therefore \quad \omega t_2 = \tan^{-1}(-\omega CR). \qquad \ldots(3)$$

Equation (2) may be written in the following form

$$i_b = \frac{E_m}{R}\sqrt{(1+\omega^2 C^2 R^2)}\sin(\omega t+\phi)$$

where $\phi = \tan^{-1}(\omega CR) = \pi - \tan^{-1}(-\omega CR)$.

Thus,

$$i_b = \frac{E_m}{R}\sqrt{(1+\omega^2 C^2 R^2)}\sin(\omega t_2 - \omega t),\ \omega t_1 < \omega t < \omega t_2. \qquad \ldots(4)$$

After $\theta_2 = \omega t_2$, the capacitor discharges through R exponentially with a time constant CR until the next positive cycle of the applied voltage equals e_R or e_c. At this stage the capacitor begins to recharge and the diode starts conducting. This situation exists at T_3 at phase angle of $(2\pi+\theta_1)$.

The controlling differential equation during the non-conducting portion of the cycle is

$$i_b = C\frac{de_c}{dt} + \frac{e_c}{R} = 0 \quad (i_b = 0, \text{ during non-conducting portion})$$

or

$$\frac{de_c}{dt} + \frac{e_c}{RC} = 0. \qquad \ldots(5)$$

This is the differential equation for the interval ωt_2 to $2\pi+\omega t_1$. The solution is given by

$$e_c = Ae^{-t/CR}, \qquad \ldots(6)$$

where A is a constant, whose value can be obtained by applying the boundary condition $e = e_c = E_m \sin \omega t_2$ when $\omega t = \omega t_2$,

$$\therefore \quad E_m \sin \omega t_2 = Ae^{-t_2/CR}.$$

or

$$A = E_m \sin \omega t_2\, e^{t_2/CR} \qquad \ldots(7)$$

Substituting this value of A in equation (6), we get

$$e_c = E_m \sin \omega t_2 . e^{-(t-t_2)/CR}$$

or

$$e_c = E_m \sin \omega t_2 . e^{-\omega(t-t_2)/\omega CR},\ \omega t_2 < \omega t < 2\pi+\omega t_1 \qquad \ldots(8)$$

It is clear from the equation (8) that for large ωCR that is, for large R (light current load) or large C, the exponential term will approximate unity and hence the voltage during the interval will not drop greatly below at ωt_2. The ripple is thus small. For small ωRC, *i.e.* small R (heavy current load), the exponential term has greater effect and voltage drops considerably below its value at ωt_2. *Thus, the ripple with a capacitor filter increases with the increase in load current.*

$$E_{d.c.} = \frac{1}{2\pi}\int_{\omega t_1}^{\omega t_2} E_m \sin \omega t\, d(\omega t) + \frac{1}{2\pi}\int_{\omega t_2}^{2\pi+\omega t_1} E_m \sin \omega t\, e^{-\omega(t-t_2)/\omega CR}\, d(\omega t). \qquad \ldots(9)$$

At $\omega t_1 + 2\pi$,

$$e_c = E_m \sin(\omega t_1 + 2\pi) = E_m \sin \omega t_1$$

and from equation (8) we get

$$e_o = E_m \sin \omega t_2 . e^{-(\omega t_1 - \omega t + 2\pi)/\omega C R}$$

Equating the two we have,

$$E_m \sin \omega t_1 = E_m \sin \omega t_2 . e^{-(\omega t - \omega t_2 + 2\pi)/\omega C R}$$

or $$\sin \omega t_1 = \sin \omega t_2 \, e^{-(\omega t_1 - \omega t_2 + 2\pi)/\omega C R} \qquad \ldots(10)$$

Substuting equations (3) and (10) in equation (9) and integ-rating we have

$$E_{d.c.} = \frac{E_m}{2\pi} \sqrt{(1+\omega^2 C^2 R^2)} \, [1 - \cos (\omega t_2 - \omega t_1)]. \qquad \ldots(11)$$

If a curve of $E_{d.c.}/E_m$ is plotted against $\tan^{-1} \omega CR$, a complex waveform is obtained and it is difficult to calculate the ripple factor. If the output waveform is taken approximately as shown in fig. 34, the result can be simplified.

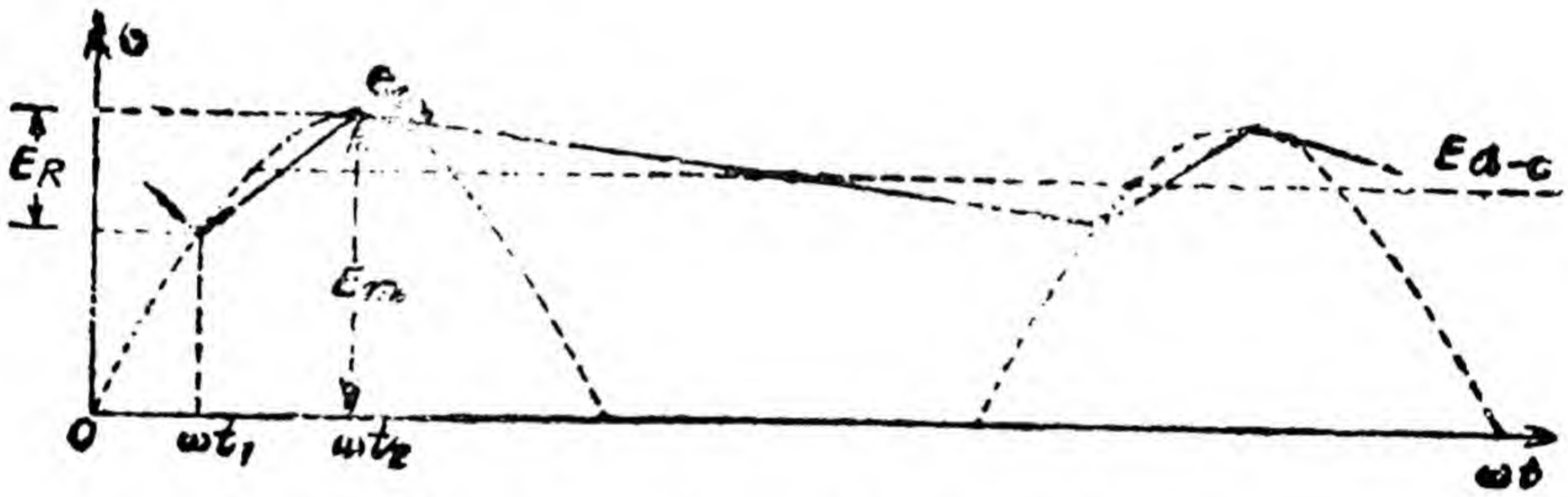


Fig. 34 Approximate output waveform.

From the figure,

$$E_{d.c.} = E_m - \frac{E_R}{2} \qquad \ldots(12)$$

During the discharge interval ωt_2 to $2\pi + \omega t_1$, the capacitor is ssumed to loose charge at a output rate with a current $I_{d.c.}$ The ate of loss of potential is

$$\frac{de_o}{d(\omega t)} = \frac{E_R}{2\pi + \omega t_1 - \omega t_2} = \frac{1}{\omega C} \frac{dq}{dt}.$$

We know that $\frac{dq}{dt} = I_{d.c.}$

$$\therefore \quad \frac{de_o}{d(\omega t)} = \frac{1}{\omega C} I_{d.c.} = \frac{E_R}{2\pi + \omega t_1 - \omega t_2}$$

or $$E_R = \frac{(2\pi + \omega t_1 - \omega t_2) \, I_{d.c.}}{\omega C} \qquad \ldots(13)$$

Taking the time axis along the $I_{d.c.}$ curve, it can be shown hat for such a triangular wave, the r.m.s. value of a.c. component of voltage is given by

$$E_{a.c.} = \frac{E_R}{2\sqrt{3}} \qquad \text{(See numerical example-1)} \qquad \ldots(14)$$

Substituting the value of E_R in equation (14) from equation (13), we get $$E_{a.c.}=\frac{(2\pi+\omega t_1-\omega t_2)\ I_{d.c.}}{2\sqrt{3}.\omega C}. \qquad ...(15)$$

Hence ripple factor

$$\gamma=\frac{E_{a.c.}}{E_{d.c.}}=\frac{(2\pi+\omega t_1-\omega t_2)\ I_{d.c.}}{2\sqrt{3}\ \omega C\ E_{d.c.}}$$

$$=\frac{(2\pi+\omega t_1-\omega t_2)\ I_{d.c.}}{2\sqrt{3}\ \omega C\ (I_{d.c.}\ R)}$$

$$=\frac{(2\pi+\omega t_1-\omega t_2)}{2\sqrt{3}\ \omega CR}, \qquad ...(16)$$

when $\omega t_1 \simeq \omega t_2$ is very small as compared to 2π, it can neglected and the ripple factor is given by

$$\gamma=\frac{2\pi}{2\sqrt{3}.2\pi f.CR}=\frac{1}{2\sqrt{3}f.CR}. \qquad ...(17)$$

The ripple factor in full wave vacuum tube rectifier can be obtained in a similar manner. The only difference is that two pulses of current are received by the capacitor from the source in each time.

$$\textit{Ripple factor}\ \gamma=\frac{(\pi+\omega t_1-\omega t_2)}{2\sqrt{3}\omega CR} \qquad ...(18)$$

The approximate output waveform of full wave vacuum tube rectifier is shown in fig. 35.

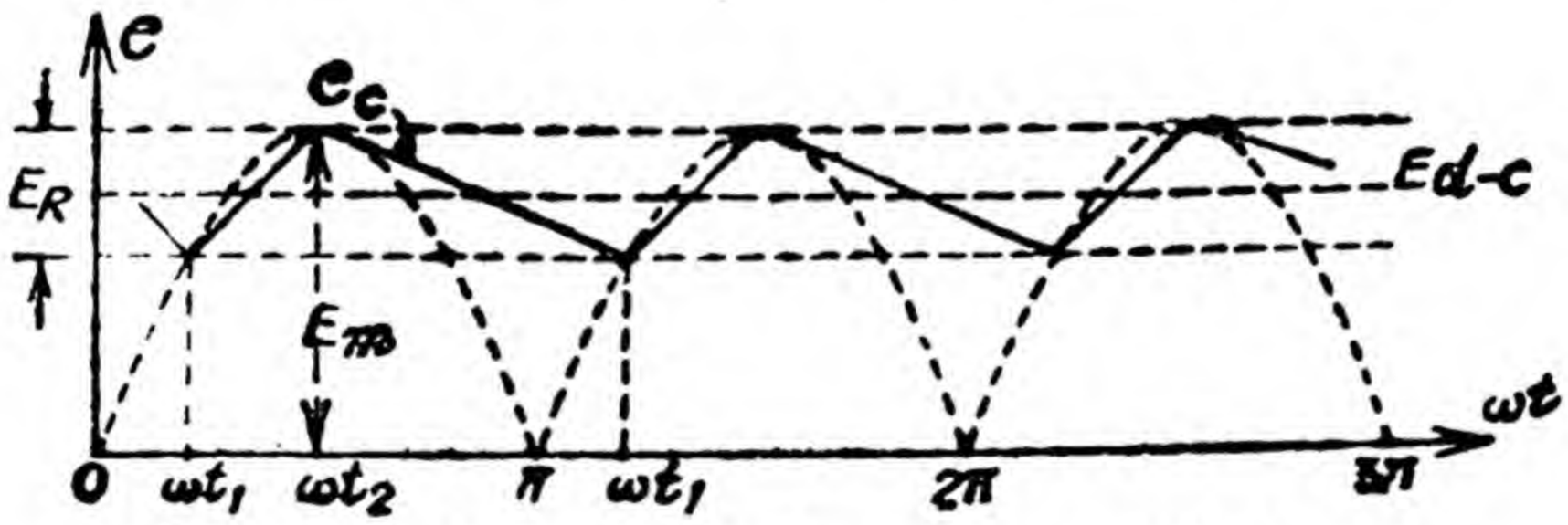


Fig. 35. Saw tooth voltage representation in the case of full wave rectifier.

(iii) Full wave rectifier with series inductor filter : Fig. 36 shows the circuit diagram of a full wave rectifier with an inductor filter at its output end. As no cutoff occurs in the current, the analysis of full wave rectifier with series inductor filter is different from that used in case of half wave rectifier with series inductor. The Fourier series shows that e_{AB} is of the form

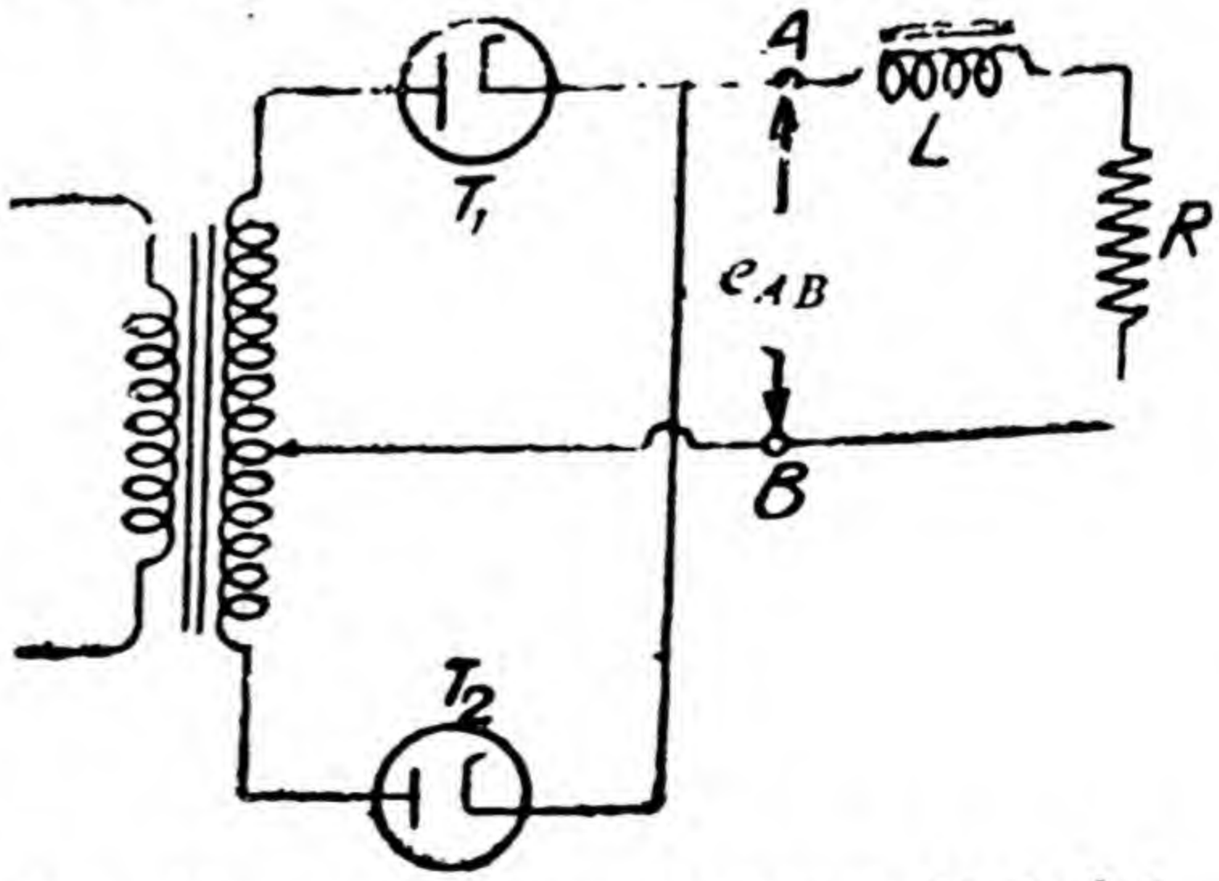


Fig. 16. Full wave rectifier with series inductor filter.

$$e_{AB}=\frac{2E_m}{\pi}\left[1-\frac{2}{3}\cos 2\omega t-\frac{2}{15}\cos 4\omega t\right] \qquad ...(1)$$

As far as the amplitude of the output voltage is concerned, the only two terms are important ; therefore, the analysis is approximated only by the first two terms *i.e.*

$$e_{AB}=\frac{2E_m}{\pi}-\frac{4E_m}{3\pi}\cos 2\omega t. \qquad ...(2)$$

If R_{ch} is the choke coil resistance, the d.c. current produced by the first term is

$$I_{d.c.}=\frac{2E_m}{\pi\,(R_{ch}+R)}, \qquad ..(3)$$

where R is the load resistance.

Similarly the current due to second term (fundamental a.c. component) is given by

$$\frac{4E_m}{3\pi}\,\frac{\cos\,(2\omega t-\phi)}{\sqrt{[(R_{ch}+R)^2+4\omega^2L^2]}}, \qquad ...(4)$$

where $$\tan\phi=\frac{2\omega L}{(R_{ch}+R)}.$$

The total current in the load is

$$I=\frac{2E_m}{\pi\,(R_{ch}+R)}-\frac{4E_m}{3\pi}\,\frac{\cos\,(2\omega t-\phi)}{\sqrt{[(R_{ch}+R)^2+4\omega^2L^2]}} \qquad ...(5)$$

The ripple factor γ is defined as

$$\gamma=\frac{\text{r.m.s value of the a.c. component of current}}{\text{d.c. component of the current}}$$

$$=\frac{\dfrac{4E_m}{3\pi\sqrt{2}}\cdot\dfrac{1}{\sqrt{[(R_{ch}+R)^2+4\omega^2L^2]}}}{\dfrac{2E_m}{\pi\,(R_{ch}+R)}}$$

$$=\frac{2}{3\sqrt{2}}\,\frac{(R_{ch}+R)}{\sqrt{[(R_{ch}+R)^2+4\omega^2L^2]}}$$

$$=\frac{\sqrt{2}}{3}\,\frac{1}{\sqrt{[1+4\omega^2L^2/(R_{ch}+R)^2]}}$$

If $\dfrac{\omega L}{R_{ch}+R}$ is large, the ripple factor reduces to

$$\gamma=\frac{\sqrt{2}}{3}\,\frac{(R_{ch}+R)}{2\omega L}.$$

and if $R_{ch}<<<R$, then

$$\gamma=\frac{\sqrt{2}}{3}\,\frac{R}{2\omega L}=0{\cdot}236\,\frac{R}{\omega L}.$$

The equation shows that the ripple decreases with the increase of inductance.

The d.c. output voltage will be due to the d.c. component of current only *i.e*,

$$E_{d.c.}=I_{d.c.}\times R=\frac{2E_m}{\pi}\frac{R}{(R_{ch}+R)}=\frac{2E_m}{\pi\left(1+\frac{R_{ch}}{R}\right)}$$

If $R_{ch} << R$

$$E_{d.c.}=\frac{2E_m}{\pi} \quad \ldots(7)$$

Equation (7) is independent of R and hence shows a perfect voltage regulation. However, this expression is derived on the assumption that plate resistance of the tube, choke resistance and transformer reactance are zero, while they are not. Due to this fact the output voltage decreases with the increase of load.

(ii) L section filter or choke input filter : The two types namely series inductor filter and shunt capacitor filter are not much efficient to provide a low value of ripple at all loads. In choke input filter we use a combination of series inductor and a shunt capacitor. The capacitor, shunting the load, bypasses the harmonics currents while inductor offers a high impedance to the harmonic terms.

The capacitor filter has low ripple at heavy load while inductor filter at small loads. Therefore a combination system of these two should give low values of the ripple at all loads.

We choose elements of the filter such that the series impedance is high while shunt impedance is low, *i.e.* choke impedance is higher than the effective parallel impedance of the capacitor and load resistor. The Fourier series of the voltage waveform at the input of the filter has the form

$$e_t=\frac{2E_m}{\pi}-\frac{4E_m}{3\pi}\cos 2\omega t-\frac{4}{15}\frac{E_m}{\pi}\cos 4\omega t, \quad \ldots(1)$$

where E_m is the peak input voltage across each diode.

The first term $(2E_m/\pi)$ represents the d.c. voltage applied to the input of the filter. As the plate resistance of the diode and d.c.

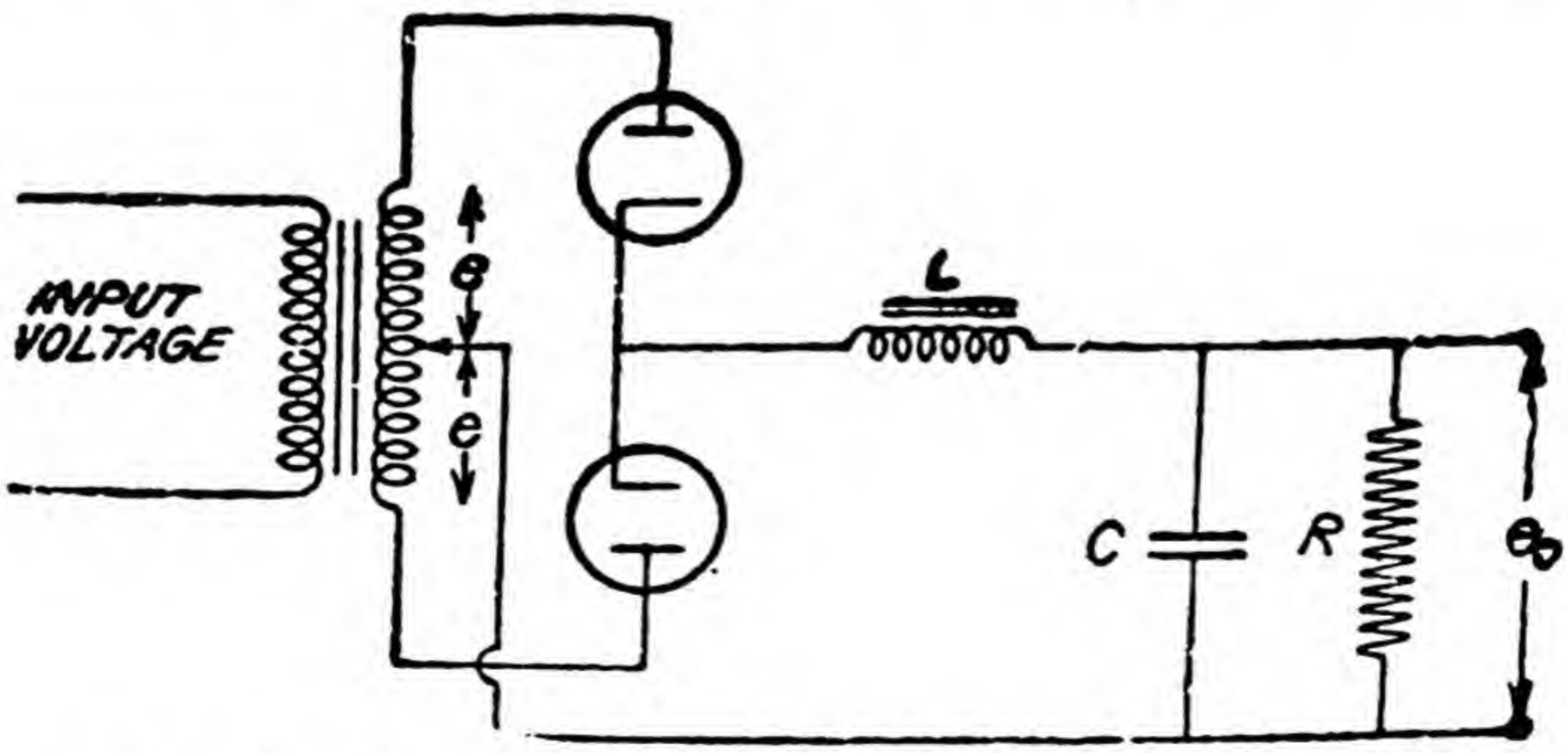


Fig. 37. Rectifier circuit with choke input filter.

resistance of inductance L are assumed far smaller than R, the first term $(2E_m/\pi)$ also represents the d.c. output voltage across R.

$$\therefore \quad E_{d.c.}=\frac{2E_m}{\pi} \qquad \ldots(2)$$

and

$$I_{d.c.}=\frac{2E_m}{\pi R}$$

where $I_{d.c.}$ is the direct current through R.

The second and third terms represent the second and fourth harmonics respectively. The filter is used to get rid of the harmonics.

To calculate the second harmonic current, consider the impedance offered by the filter circuit at the second harmonic frequency 2ω, which is given by

$$Z_{2f}=j\,2\omega L+\frac{R/\,j\,2\omega C}{R+\dfrac{1}{j\,2\omega C}}. \qquad \ldots(3)$$

Now second harmonic current $I_{2f}=\dfrac{\text{Second harmonic term}}{Z_{2f}}$

$$\therefore \quad I_{2f}=\frac{4E_m}{3\pi\left[j\,2\omega L+\dfrac{R/j\,2\omega C}{\left(R+\dfrac{1}{j\,2\omega C}\right)}\right]} \qquad \ldots(4)$$

The filter elements are chosen so that the ripple voltage appears across the inductor instead across the load or condenser C and thus make the filter effective. In practice load R is usually high

$$\text{Thus } R \gg 2\omega L \gg \frac{1}{2\omega C}.$$

The approximate value of Z_{2f} and I_{2f} are

$$Z_{2f}=2\omega L$$

and

$$I_{2f}=\frac{4E_m}{3\pi\,2\omega L}. \qquad \ldots(5)$$

As the reactance offered by the capacitor C is quite small, compared with load resistance R, almost all a.c. component of the second harmonic current passes through C. Thus the approximate value of the second harmonic voltage E_{2f} across R is

$$E_{2f}=I_{2f}\times X_c=\frac{4E_m}{3\pi\,2\omega L}\cdot\frac{1}{2\omega C}. \qquad \ldots(6)$$

Similarly, the impedance offered by the filter circuit at the fourth harmonic frequency is

$$Z_{4f}=j\,4\omega L+\frac{R/j\,4\omega C}{\left(R+\dfrac{1}{j\,4\omega C}\right)}\simeq j\,4\omega L. \qquad \ldots(7)$$

The fourth harmonic voltage across the load is

$$E_{4f}=\frac{4E_m}{15\pi}\times\frac{1}{Z_{4f}}\times\frac{1}{j\,4\omega C}$$

$$=\frac{4E_m}{15\pi}\times\frac{1}{j\,4\omega L}\times\frac{1}{j\,4\omega C}=\frac{4E_m}{15\pi\times 4\omega L\times 4\omega C}. \quad ...(8)$$

The ratio between the fourth harmonic and the second harmonic voltage across the load resistance R is

$$\frac{E_{4f}}{E_{2f}}=\frac{4E_m}{15\pi\times 4\omega L\times 4\omega C}\times\frac{3\pi\times 2\omega L\times 2\omega C}{4E_m}$$

$$=\frac{1}{20}$$

$$=5\%. \quad ...(9)$$

As the fourth harmonic is only 5% of the second harmonic, the higher harmonics above second harmonic may be neglected in calculating the ripple factor.

The ripple factor is defined as the ratio betwen the a.c. and d.c. components of the voltage across the load; hence

$$\gamma=\frac{E'_{r.m.s.}}{E_{d.c.}}=\frac{E_{2f}/\sqrt{2}}{E_{d.c.}}$$

$$=\frac{\left(\frac{4E_m}{3\pi\times 2\omega L\times 2\omega C}\right)\Big/\sqrt{2}}{2E_m/\pi}$$

$$=\frac{1}{6\sqrt{2\omega^2 LC}}. \quad ...(10)$$

It is obvious from above relation that ripple is independent of load resistance. By choosing approximate values of L and C ripple can be reduced to less than one percent.

Voltage Regulation : The smoothing effect due to inductor is shown in fig. 38. It is seen that current does not fall to zero during any portion of voltage variation cycle because it releases its stored energy during the period when voltage drops to zero. The a.c. component are bypassed through the capacitor and, therefore, a ripple-free smooth voltage may be obtained across the load.

A continuous current must flow in inductor for

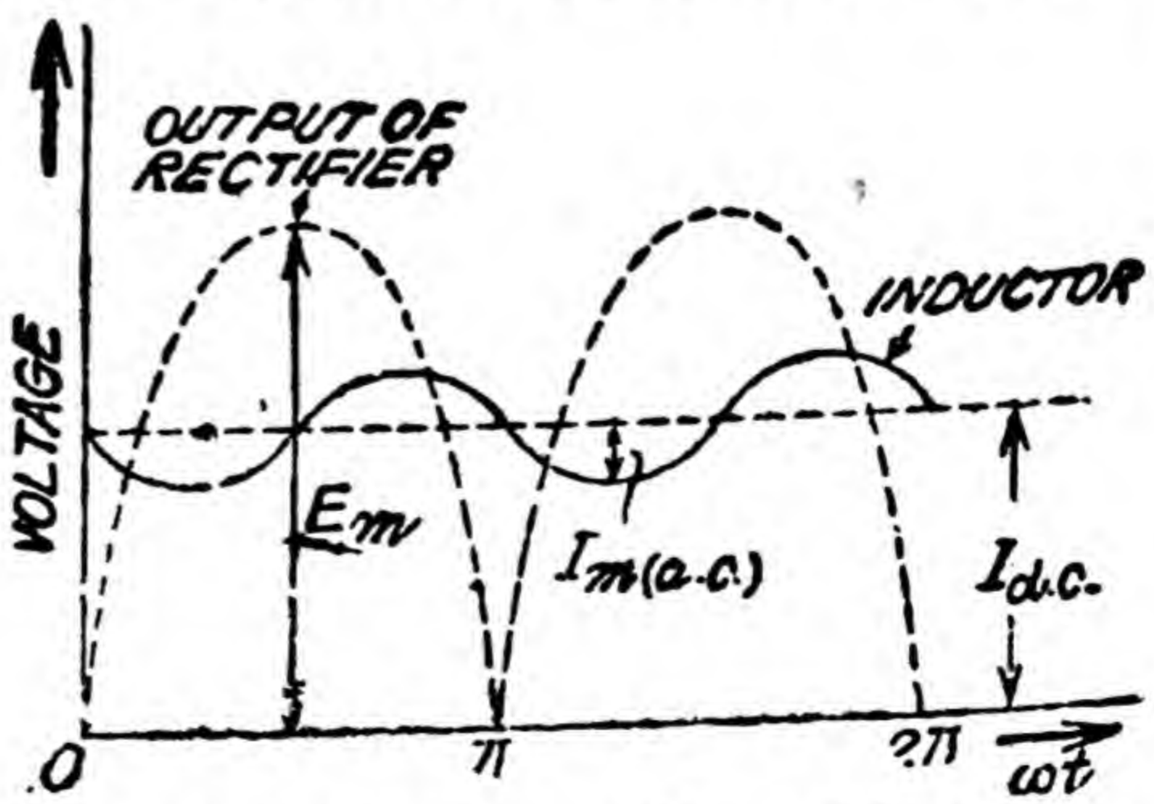


Fig. 38. Smoothing effect of inductor.

which I_m (a.c.) $\leqslant I_{d.c.}$ is necessary. This condition in turn depends on relative values of L and R. The critical inductance L_c required to give sufficient energy storage for a continuous flow of current is given by

$$L_c = \frac{R}{3\omega} \text{ for a given load } R. \qquad \ldots(1)$$

R is in ohms and L_c in henrys.

A regulation curve in direct output voltage against load current for a filter employing a given value of inductance is shown in figure 39.

It is clear from the graph Fig. 39 that below critical current I_o, output voltage variation is large, while for the load current greater than I_o output voltage is essentially independent of load ; the reason being that when load current is greater than I_o inductance current flows continually, which is necessary in the action of choke input filter.

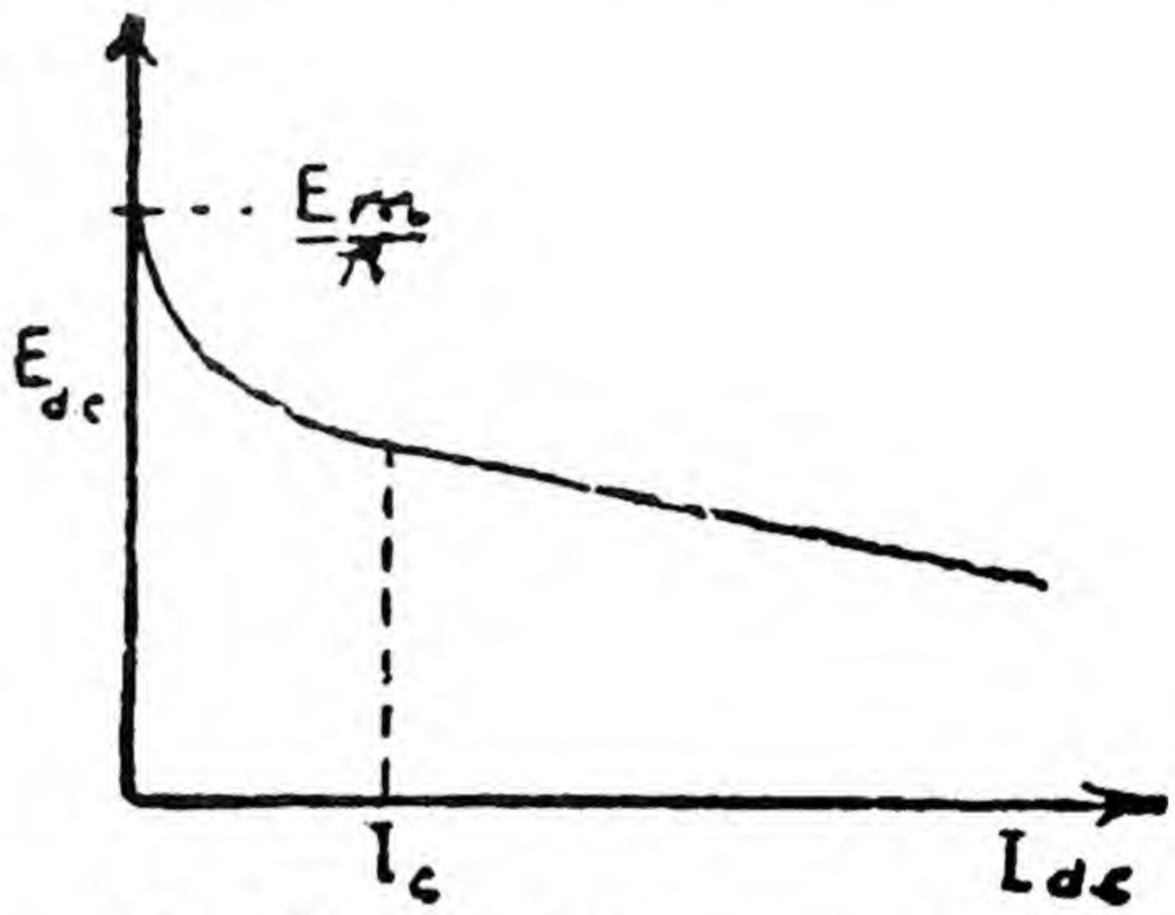


Fig. 39. Regulation current for L section filter.

(v) Capacitor input Filter or π Section Filter : A circuit of full wave rectifier with π section filter is shown in fig. 40.

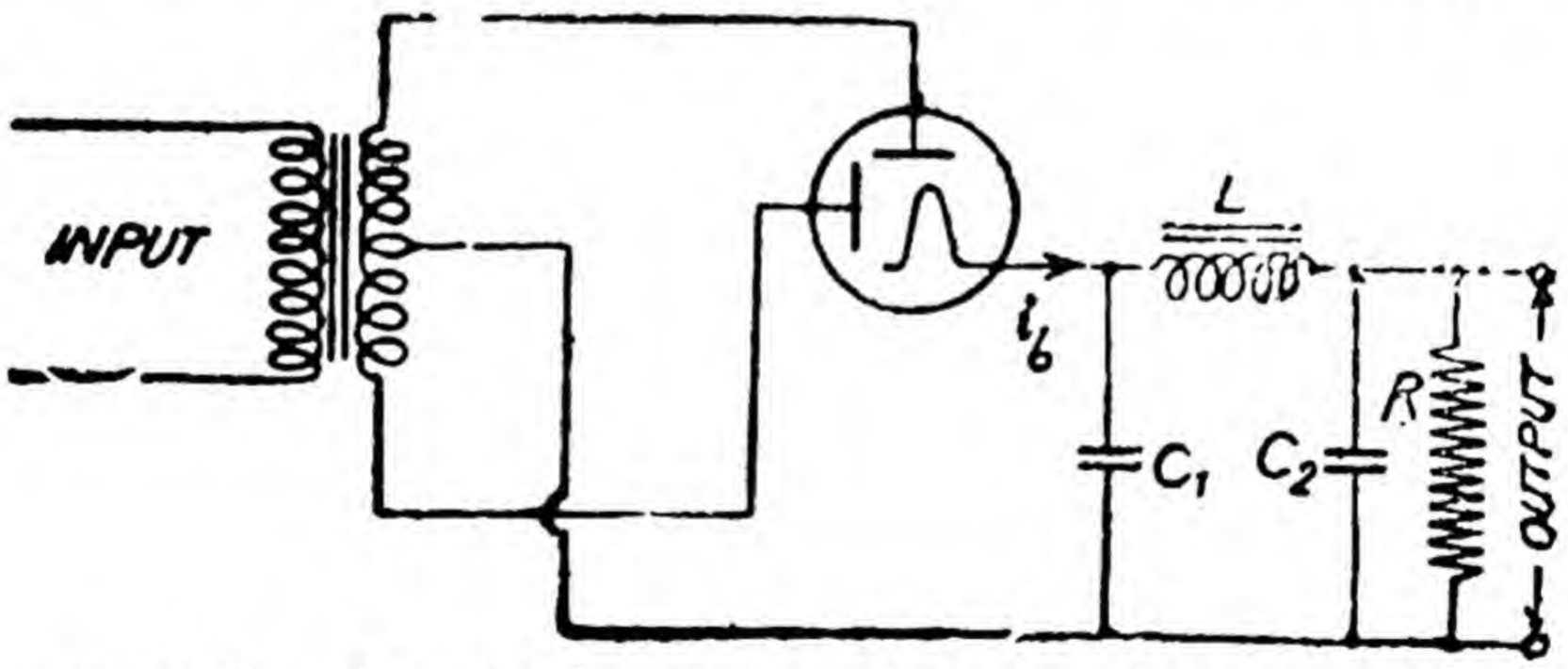


Fig. 40. Circuit of rectifier with capacitor input filter.

When higher output voltage at light loads is desired, an input capacitor is added to L section filter to form a π section filter.

The capacitor C_1 draws a charging current pulse twice each cycle and charges to E_1 before being disconnected from the source by a diode. The capacitor then discharges through the inductance L and load R until the second diode connects the source on the next half cycle.

As shown in figure 41, during each conduction interval capacitor C_1 charges almost to the peak value of applied alternating voltage. After this neither tube conducts nor capacitor begins to

discharge. The inductance serves to maintain an almost constant discharge current. During the time in which the capacitor voltage reaches to an appropriate value, one of the tube conducts and thus the capacitor is again charged.

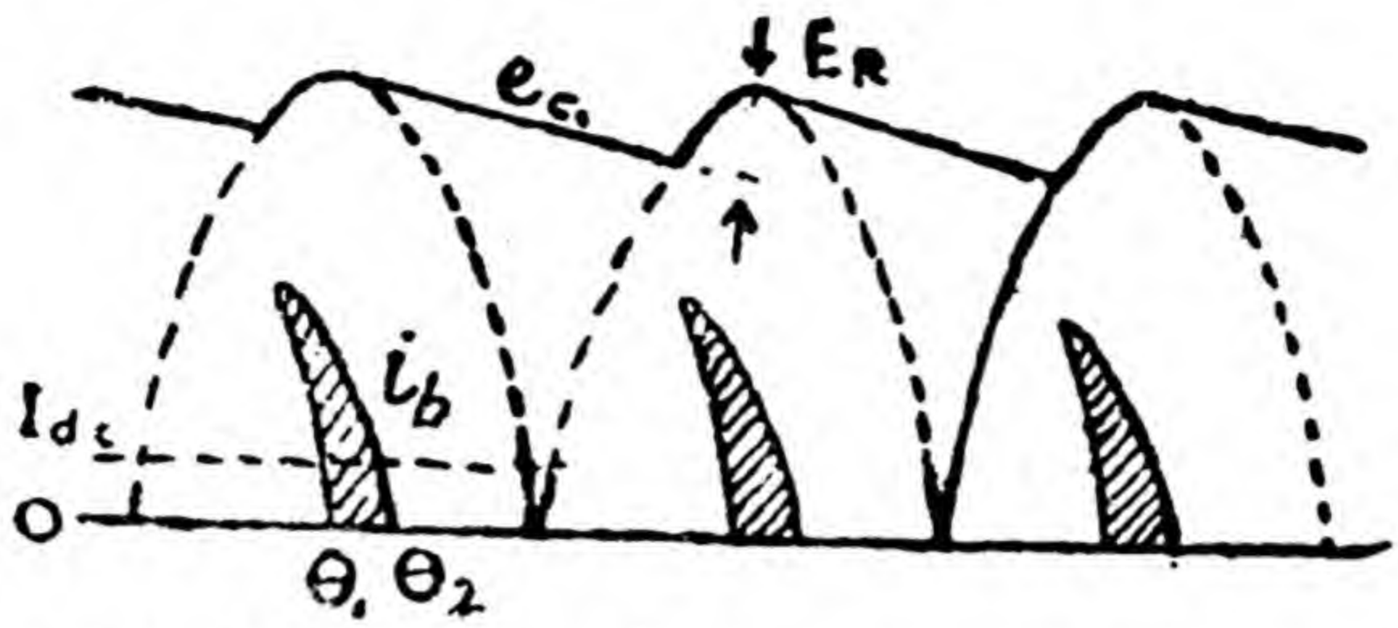


Fig. 41. Waveforms of current and voltage on C_1.

Output voltage : In this analysis it is assumed that C_1 is large enough to supply a constant current to the load during the non-conduction period of the diode (it is charged to the peak value of supply voltage E_m during the positive half cycle). The current flowing in L is constant and the ripple component is small. The d.c. voltage across load is

$$E_{d.c.}=E_m-\frac{E_R}{2}-I_{d.c.}R_c \qquad ...(1)$$

where R_c is choke resistance.

E_R is peak to peak magnitude which may be found by assuming that C_1 is large and the conduction interval is zero.

If the capacitor C_1 is assumed to discharge at a constant rate through L and load for $T/2$ seconds, the total quantity of charge discharged is

$$Q=I_{d.c.}\frac{T}{2} \text{ coulombs,}$$

and the fall of voltage across C_1 is

$$E_R=\frac{Q}{C_1}=\frac{I_{d.c.}T}{2C_1} \qquad ..(2)$$

we know that $T=1/f$, where f is the frequency of a.c. cycle

$$\therefore \quad E_R=\frac{I_{d.c.}}{2fC_1}=\frac{\pi I_{d.c.}}{2\pi fC_1}=\pi\frac{I_{d.c.}}{\omega C_1}.$$

Thus output voltage becomes

$$E_{d.c.}=E_m-I_{d.c.}\left(\frac{\pi}{2\omega C_1}+R_c\right) \qquad ...(3)$$

Voltage Regulation : The above expression for output voltage shows that the regulation depends upon the value of C_1. The regulation is good for higher values of C_1. In fact, the major effect of C_1 is on regulation while that of C_2 is on the ripple.

Ripple Factor : If the ripple waveform be considered triangular, the ripple is given by

$$e_R = \frac{E_R}{\pi} \sin 2\omega t - \frac{E_R}{2\pi} \sin 4\omega t + \frac{E_R}{3\pi} \sin 6\omega t. \qquad \ldots(4)$$

If we consider the current in inductance L mainly due to second harmonic component and taking

$$\frac{1}{2\omega C_2} \ll R,$$

then the portion of r.m.s. second harmonic voltage appearing across C_2 or R is

$$E_2 = \frac{E_R}{\sqrt{2}.\pi} \cdot \frac{-j/2\omega C_2}{j2\omega L - j/2\omega C_2} \cdot$$

$$= \frac{E_R}{\sqrt{2}.\pi} \cdot \frac{1}{4\omega^2 LC_2 - 1} \cdot$$

Putting $E_R = \frac{I_{d.c.}\ \pi}{\omega C_1}$,

$$| E_2 | = \frac{I_{d.c.}}{\sqrt{2}\omega C_1} \cdot \frac{1}{4\omega^2 LC_2 - 1}$$

$$= \text{Ripple voltage across load.} \qquad \ldots(5)$$

$$\text{Ripple factor} = \frac{| E_2 |}{E_{d.c.}} = \frac{1}{\sqrt{2}\omega C_1 R\ (4\omega^2 LC_2 - 1)} \cdot$$

The above expression shows that ripple factor is a function of load. The ripple factor decreases with the increase of load resistance or decreasing current output.

When $4\omega^2 LC_2$ far greater than 1, the ripple factor reduces to

$$\gamma = \frac{\sqrt{2}}{8\omega^3\ C_1 C_2 LR} \cdot$$

where C_1 and C_2 are in μF, and L in henrys.

Comparison of L Section and π Section Filters :

L Section	*π Section*
1. Ripple is independent of load resistance.	1. Ripple factor varies inversely as load resistance.
2. Voltage regulation is fair.	2. Voltage regulation is poor.
3. Output voltage is low.	3. Output voltage is higher in comparison to L section at light load.

8·14. REGULATION OF POWER SUPPLY OUTPUT VOLTAGE :

In section 8·5 we have discussed the filter system. The filters are used to reduce the ripple and help in regulation of voltage. The good voltage regulation means that the output voltage remains constant.

Voltage regulation is defined as the percentage change in output voltage when the load is removed, *i.e.*,

$$\%\text{regulation} = \frac{E_0 \text{ (no load)} - E_0 \text{ (full load)}}{E_0 \text{ (full load)}} \times 100,$$

where E_0 represents output voltage.

In electronic applications, to avoid any confusion, no load and full load conditions are referred to minimum and maximum load current limits for which the regulation is desired

Example. The output voltage from a power supply drops from 320 volts at 20 mA. to 280 volts at 120 mA. Find the percentage of regulation over this range.

$$\text{Percentage of regulation} = \frac{E_{NL} - E_{FL}}{E_{FL}} \times 100$$

$$= \frac{320 - 280}{280} \times 100$$

$$= 14{\cdot}3\%.$$

Causes for poor regulation : The poor voltage regulation is due to the decrease in the output voltage of power supply on account of the voltage drops in any portion of the circuit that has resistance *e.g.* resistance in the primary and secondary windings of the transformer, resistance in the series filter element (L or R), resistance in the rectifier tube. Line voltage variations are also responsible for poor regulation. Therefore, to improve regulation of power supply voltage :

(i) chokes and power transformers should be wound with larger sized wire (to minimise winding resistances).

(ii) vacuum diodes, because of their high internal resistance, should be avoided and in place of these semi-conductor diodes, which have much lower resistance are most suited.

(iii) R.C. filter, because of its high series resistance, and π type filters, because of varying discharge time constant, should not be used.

In order to maintain the stability of output voltage over large operating range, voltage regulator circuits are employed between the filter and the load resistance in the power supply system.

We shall discuss the following voltage regulator systems :

8·10. VOLTAGE STABILIZATION BY GAS DISCHARGE TUBES (V.R. TUBES)

The principle is based on the fact that in a cold cathode gaseous discharge tube the voltage across the discharge tube remains constant over a considerable range of current. When the tube is connected as shown in figure 42, the voltage drop across the load remains more or less constant, even when there is variation in the applied input voltage (E_{in}). The tube bypasses then

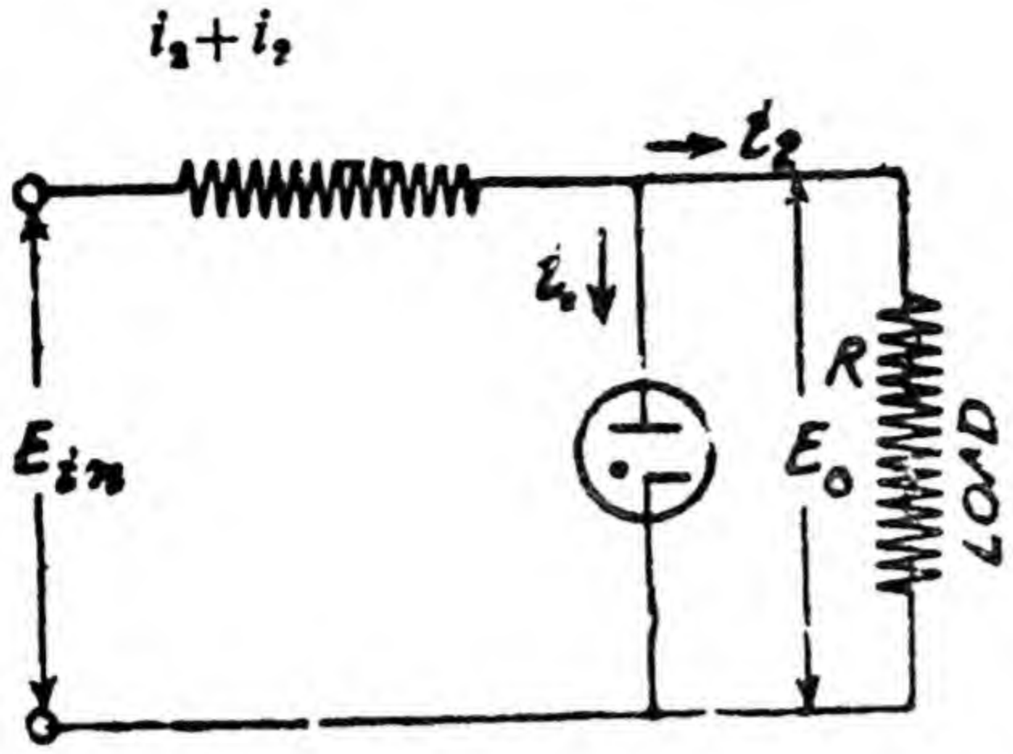


Fig. 42. Voltage regulator tube circuit.

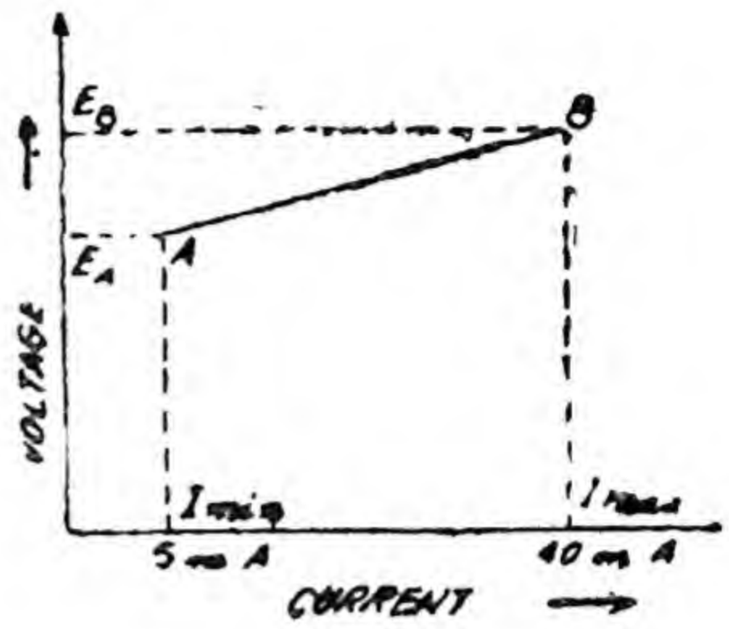


Fig. 43. Volt-ampere characteristics of a V.R tube.

current variations produced by variation in input voltage and maintains the current constant through the load *i.e.* it keeps the output voltage constant across the load. Secondly, if the input voltage is constant and load varies, then output voltage tends to vary but the tube bypasses current variation through it and maintains the current through R constant. The output voltage across the load then remains constant.

The tubes are known as V-R tubes. The voltage across the tubes is determined by the material of the cathode, nature of gas and gas pressure. When we are required to have the output voltage greater than a single V-R tube, then several V-R tubes are connected in series.

Figure 43 shows the volt-ampere characteristic of the V-R tube. The curve is almost linear over the wide range of operation and hence it can be represented by

$$E=m\ I_{max.}+C. \qquad (\text{as } y=mx+C)$$

where $$m=\frac{E_B-E_A}{I_{max.}-I_{min}}$$

and $$C=E-m_A\ I_{min.}$$

E_A and E_B are the voltages corresponding to points A and B.

From figure 4 , we have

$$E_{in}=R_1\ (i_1+i_2)+E_0 \qquad \ldots(1)$$

$$E_0=i_2R=m\ i_1+C. \qquad .(2)$$

From equation (2), we have

$$i_2=\frac{E_0}{R} \text{ and } i_1=\frac{E_0-C}{m}.$$

Substituting these values in equation (1) we have

$$E_{in}=\left(\frac{E_0}{R}+\frac{E_0-C}{m}\right)R_1+E_0.$$

Solving it we have

$$E_0 = \frac{mR\,E_{in} + CR_1R}{RR_1 + mR_1 + mR} \qquad \ldots(3)$$

Equation (3) gives a relationship between regulated output E_0, unregulated input E_{in} and the circuit parameters.

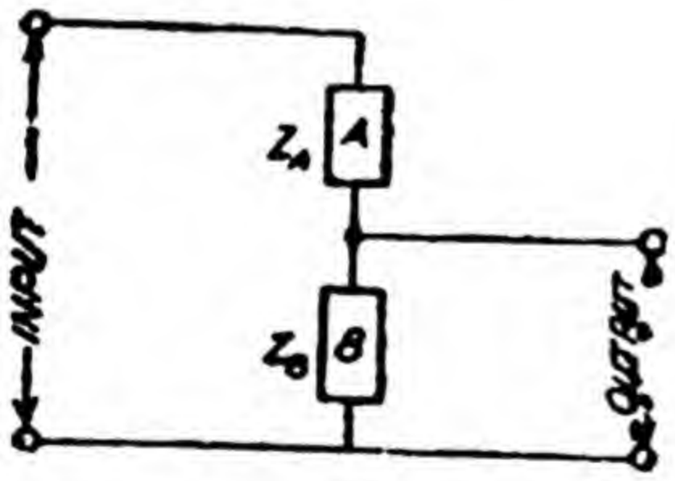
The stabilization ratio is given by

$$S = \frac{d(E_{in})}{d\,E_0} = \frac{R\,(R_1+m)+mR_1}{mR} = \frac{mE_{in}+CR_1}{mE_0} \qquad \ldots(4)$$

which shows that for stabilization, E_0 and m must be small while C, E_{in} and R_1 should be large.

8·16. ELECTRONIC POTENTIAL REGULATOR :

Principle : The input voltage to be stabilized is fed to two elements A and B and the output is taken from the element B as shown in fig. 44 (*a*). If Z_A and Z_B are the impedances of elements A and B respectively, then the fraction of input voltage fed to the output is $Z_B/(Z_A+Z_B)$. This ratio may be made to change in such a way that the changes in the input voltage may be compensated. This can be done by making either or both Z_A and Z_B non-linear. Now we shall consider the case in which Z_B is linear and Z_A is non-linear utilizing the beam resistance of a triode valve The circuit arrangement is shown in fig. 44 (*b*).



Ftg. 44 (*a*) Basic potential divider.

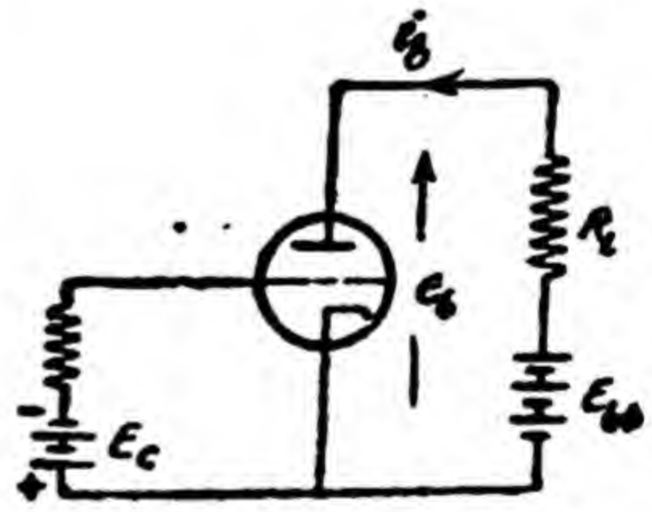


Fig. 44 (*b*) Triode circuit to explain beam resistance.

The beam resistance is different from the plate resistance (r_p) and is defined as the ratio of the plate voltage e_b to the plate current i_b *i.e.* $r_b = e_b/i_b$. The value of plate voltage e_b is given by

$$e_b = E_{bb} - i_b\,R_L \qquad \text{[from fig. 44 (}b\text{)]} \qquad \ldots(1)$$

Equation (1) shows that when plate current increases, the plate voltage e_b decreases and when plate current decreases it increases. We also know thate plate current depends upon grid voltage. When grid voltage is zero or slightly positive, plate, current is maximum and plate current is zero for cut off grid voltage. In this way, when grid voltage is zero, plate current i_b is large and beam resistance is small. Similarly, when grid voltage is negative, the plate

current decreases and beam resistance increases. Thus, as the grid voltage decreases, the beam resistance increases. For a properly designed circuit, the change in beam resistance is of such a nature as to give the output voltage which remains substantially constant.

Circuit arrangement : The circuit arrangement is shown in figure 45. Let the voltage across the load be at a desired value. Then the cathode is positive with respect to ground by a potential E_0. By adjusting R, the grid is made positive with respect to ground by a potential E_g, of course, E_g is always less than E_0. The difference of E_g and E_0 is equal to grid cathode potential. The value of E_g is adjusted, such that the tube is biased in the active region and passes adequate current through R_1 to maintain the output constant.

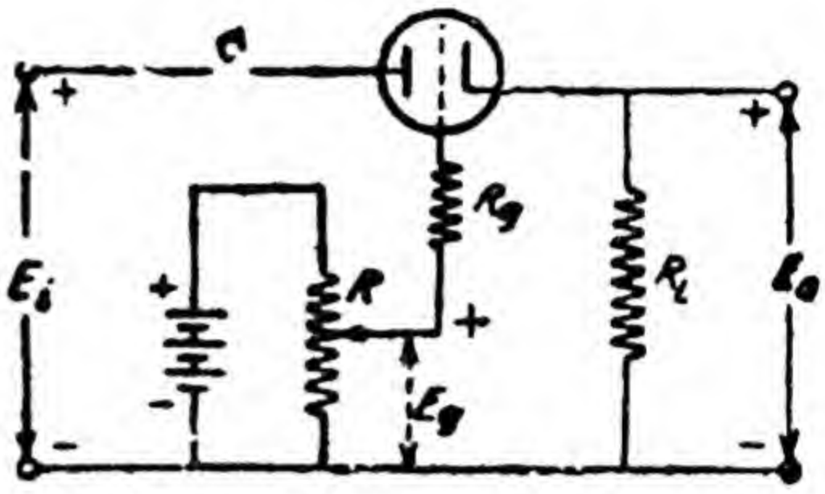


Fig. 45. Simple vacuum tube regulator.

Let for some reason E_i be increased. Due to the increase in potential, more current flows through the tube and hence E_0 will have a tendency to increase. As the output potential increases, the grid cathode potential have a tendency to decrease because E_g is constant. With this decrease of grid potential, the plate current decreases and thus beam resistance increases. The ultimate effect is that E_0 have a tendency to decrease. Any increase in output is counteracted by the circuit arrangement and E_0 is brought back to its original desired value.

Similarly consider that input E_i decreases. Now E_0 has a tendency to decrease and thus grid cathode potential has a tendency to increase. Due to the increase in grid potential, plate current increases and hence beam resistance decreases. Thus the output potential tends to increase. The output potential is brought back to its original value.

If the input voltage is removed a high grid current may flow and may damage the tube and hence to limit this current, R_g is placed in series grid.

In practice, the battery is replaced by V. R. tube as shown in fig. 46.

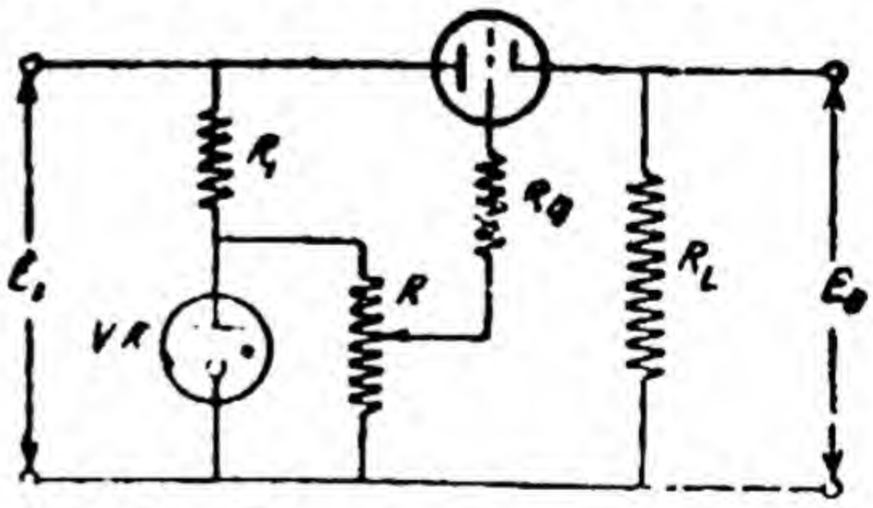


Fig. 46. Simple vacuum tube regulator with V-R. tube instead of battery.

8·17. TWO STAGE DEGENERATING VOLTAGE STABILISER (ELECTRONIC VOLTAGE STABILISER) :

Figure 47 shows the circuit of electronic voltage stabiliser. The

unregulated d.c. voltage from rectifier and filter is applied to the input terminals of regulator circuit and same is obtained at the output terminals reduced by voltage equal to the voltage drop in series control tube T_1. Tube T_2 is a d.c. amplifier and its cathode is maintained at a positive potential with respect to ground by a constant potential E_s. This is done by a V-R tube or by a battery. A voltage E_2 is applied to the control grid of amplifier tube T_2. The grid of cathode voltage of tube T_2 is equal to the difference of voltages E_2 and E_s. This difference of voltages is amplified by d.c. amplifier with tube T_2 and load resistance R_2 in plate. The amplified voltage developed across R_2 appears between plate and grid of valve T_1.

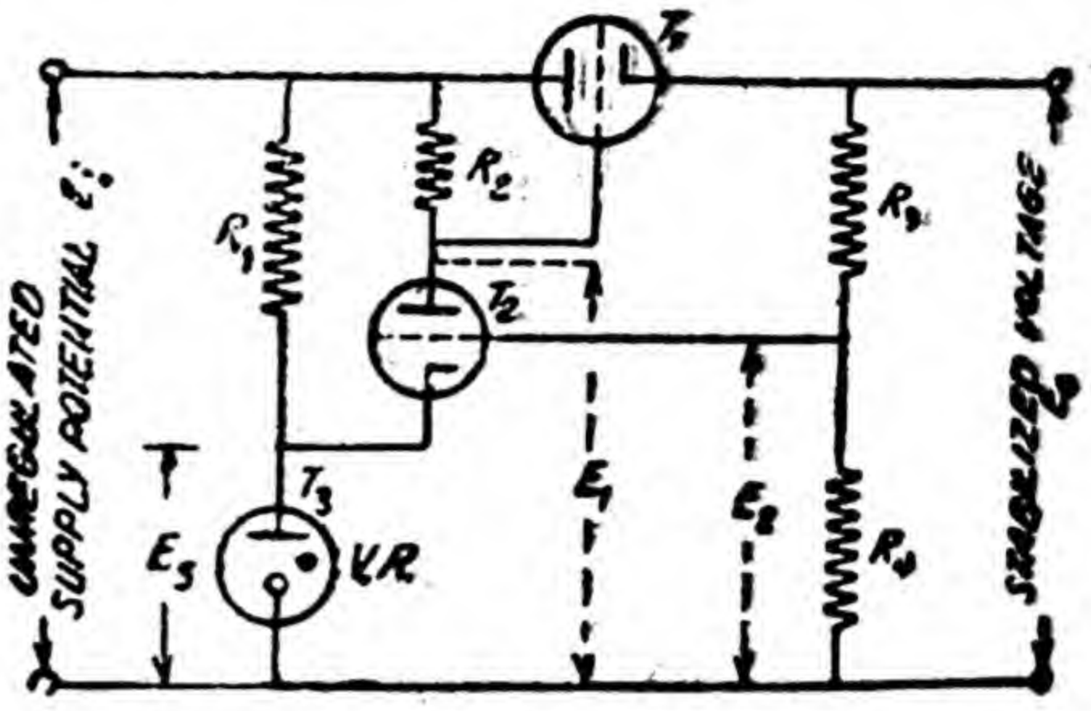


Fig. 47. Electronic voltage stabilizer.

Let us consider that for some reason, the unregulate d.c voltage increases. Due to this increment, the output voltage E_0 tends to increase. The voltage E_2, which is fraction of output voltage E_0 also tends to increase and thus the difference $(E_2 - E_s)$ increases. It means that plate current of d.c. amplifier tends to increase and its plate voltage tends to decrease. As the plate of T_2 is directly connected to the grid of T_1, the grid cathode voltage of T_1 decreases and thus the plate current tends to decrease. When the plate current decreases, the beam resistance increases and the net result is that the output tends to decrease. Thus due to the circuit arrangement, E_0 remains constant at the desired value.

Similarly we can explain that the decrease in input is also counteracted such that output is brought back to the desired value.

8·17. ANALYSIS OF ELECTRONIC VOLTAGE STABILIZER :

Here we shall consider the following two cases :

(*a*) Varying input voltage, (*b*) Varying load current.

(a) Varying input voltage :

(i) R_2 connected to the input side of stabilizer :

Let g_{m1} and r_{p1} be the mutual conductance and dynamic plate resistance of tube T_1 respectively. The variation in the plate current I_{p1} of tube T_1 comes due (i) small variation of plate voltage e_{p1} and (ii) small variation in grid voltage e_{g1}. Thus

$$dI_{p1} = g_{m1}\, e_{g1} + \frac{e_{p1}}{r_{p1}} \qquad \ldots(1)$$

From figure 47, we find that

$$dE_{g1} = e_{g1} = dE_1 - dE_0$$

$$dE_{p1} = e_{p1} = dE_i - dE_0$$

$$dI_{p1} = i_{p1} = \frac{dE_0}{R_3 + R_4} = \frac{dE_0}{R_L},$$

where R_L is the load resistance and is assumed to be constant. Substituting these values in equation (1), we get

$$\frac{dE_0}{R_L} = g_{m1}(dE_1 - dE_0) + \frac{(dE_i - dE_0)}{r_{p1}}$$

or $$dE_0\left[\frac{1}{R_L} + g_{m1} + \frac{1}{r_{p1}}\right] = g_{m1}\, dE_1 + \frac{dE_i}{r_{p1}}$$

or $$dE_0\left[\frac{r_{p1}}{R_L} + r_{p1}\, g_{m1} + 1\right] = g_{m1}\, r_{p1}\, dE_1 + dE_i$$

or $$dE_0\left[\frac{r_{p1}}{R_L} + \mu_1 + 1\right] = \mu_1 dE_1 + dE_i$$

$$\therefore \quad dE_0 = \frac{\mu_1\, dE_1 + dE_i}{\left[\frac{r_{p1}}{R_L} + \mu_1 + 1\right]} \qquad \ldots(2)$$

In order to find the ratio of $\frac{dE_i}{dE_0}$, the term dE_1 should be replaced in terms of dE_i, dE_0 and circuit constan's. The change dE_1 occurs due to two effects : (i) voltage variation dE_i, and (ii) voltage variation dE_2.

The effect of first factor is $\frac{r_{p2}\, dE_i}{r_{p2} + R_2}$, where r_{p2} is the dynamical plate resistance of tube T_2. This effect is due to series circuit comprising to R_2 and plate resistance r_{p2} of tube T_2. Of course, the voltage E_s across V-R tube is assumed to be constant and hence does not come in picture.

The second component is the amplified effect of input voltage applied between the grid and cathode of tube T_2. If dE_0 is the change in output potential, then the change in the grid voltage of this tube is $[R_3/ R_3 + R_4)]\, dE_0$. Let A_2 be the voltage gain of the amplifier using tube T_2, then the change in dE_1 due to amplified signal of T_2 is equal to $[R_4/(R_3 + R_4)]\, dE_0 \times A_2$.

Thus the total effective change dE_1 due to both effect is

$$dE_1 = \frac{r_{p2}\, dE_i}{r_{p2} + R_2} + \frac{R_4 dE_0 A_2}{(R_3 + R_4)}$$

or $$dE_1 = \alpha_2\, dE_i + \beta A_2\, dE_0, \qquad \ldots(3)$$

where $\alpha_2 = \frac{r_{p2}}{R_2 + r_{p2}}$, $\beta = \frac{R_4}{R_3 + R_4}$ and $A_2 = \frac{-\mu_2 R_2}{r_{p2} + R_2}$.

Substituting the value of dE_1 from equation (3) in equation (2), we have

$$dE_0 = \frac{\mu_1(\alpha_2\, dE_i + \beta\, A_2\, dE_0) + dE_i}{(r_{p1}/R_L + \mu_1 + 1)}$$

or $$dE_0=\frac{dE_i(\mu_1\alpha_2+1)+\mu_1\beta A_2\, dE_0}{(r_{p1}/R_L+\mu_1+1)}$$

or $$dE_0\left[\left(\frac{r_{p1}}{R_L}+\mu_1+1\right)-\mu_1\beta A_2\right]=dE_i\,(\mu_1\alpha_2+1). \quad ...(4)$$

Hence the stabilization ratio S is given by

$$S=\frac{dE_i}{dE_0}=\frac{r_{p1}/R_L+1+\mu_1\,(1-\beta A_2)}{(\mu_1\,\alpha_2+1)}. \quad ...(5)$$

Under normal conditions $\beta A_1 > > 1$

and $\mu_1\,(1-\beta A_2) > > \frac{r_{p1}}{R_L}+1$, and hence we may write

$$S=\frac{-\beta\mu_1 A_2}{\mu_1\alpha_2+1} \quad ...(6)$$

The effectiveness of a circuit is determined by quantity S. Larger is the value of S, better is the regulation. With a simple d.c. amplifier S varies from 500 to 1000. Thus to improve S further two stages of d.c. amplifier may be used.

(ii) R_2 connected to the output side of the stabilizer :

Under this condition

$$dE_1=\beta A_2\, dE_0+\alpha_2\, dE_0. \quad ...(7)$$

But $\alpha_2 << \beta A_2$, hence

$$dE_1=\beta A_2\, dE_0. \quad ...(8)$$

Substituting equation (8) in equation (2), we ge

$$dE_0=\frac{\mu_1\,\beta A_2\, dE_0+dE_1}{r_{p1}/R_L+\mu_1+1}$$

or $$\left(\frac{r_{p1}}{R_L}+\mu_1+1\right)dE_0-\mu_1\beta A_2\, dE_0=dE_1$$

or $$\frac{dE_1}{dE_0}=S=\frac{r_{p1}}{R_L}+1+\mu_1(1-\beta A_2). \quad ...(9)$$

As $\beta A_2 \gg 1$ and $\mu_1\,(1-\beta A_2) > > 1+\frac{r_{p1}}{R_L}$, we have

$$S=-\mu_1\beta A_2. \quad ...(10)$$

In this case, the value of S is greater than previous case ; however, the d.c. amplifier is more linear and gives more gain in previous case. Hence there is no appreciable difference in results of two cases :

(*b*) Varying load current :

(i) R_2 connected to the input side of stabilizer :

Now consider the case in which input voltage remains constant but there is a change in the output potential due to the change in load current *i.e.* stabilizer has internal resistance which produces output voltage drop with increase of load current. Thus the internal resistance of the stabilizer should be as small as possible. The internal resistance of the stabilizer can be derived in the following way :

We know that

$$dI_{p1}=g_{m1}\, e_{g1}+\frac{e_{p1}}{r_{p1}}$$

$$dE_{g1}=e_{g1}=dE_1-dE_0$$

and $$dE_{p1}=e_{p1}=dE_l-dE_0$$

In this case $dE_l=0$, *i.e.*, e_{p1} is very small and hence $\frac{e_{p1}}{r_{p1}}$ can be neglected as compared to $g_{m1}\, e_{g1}$.

$$\therefore \quad dI_{p1}=g_{m1}\, e_{g1}=g_{m1}\,(dE_1-dE_0).$$

Now from equation (3).

$$dE_1=\beta A_2\, dE_0 \text{ as } dE_i=0. \qquad \ldots(12)$$

From equations (11) and (12), we get

$$dI_{p1}=g_{m1}\,(\beta A_2-1)\, dE_0$$

$$\therefore \quad R_0=\frac{dE_0}{dI_{p1}}=\frac{1}{g_{m1}\,(\beta A_2-1)}.$$

As $\beta A_2 \gg 1$, the effective internal resistance of regulated power supply is given by

$$R_0=\frac{1}{\beta A_2 g_{m1}}. \qquad \ldots(13)$$

(ii) R_2 connected to the output side of the stabilizer :

In this case $$dE_1=\beta A_2\, dE_0+\alpha_2\, dE_0$$
$$=(\beta A_2+\alpha_2)\, dE_0.$$

$$\therefore \quad dI_{p1}=g_{m1}\,(\beta A_2+\alpha_2-1)\, dE_0$$

or $$\frac{dE_0}{dI_{p1}}=R_0=\frac{1}{g_{m1}\,(\beta A_2+\alpha_2-1)}$$

$$=\frac{1}{g_{m1}\,\beta A_2} \text{ as } \beta A_2 >> (\alpha_2-1).$$

The value of S_0 may be as small as 0·2 ohm.

8·17. MODIFIED FORM OF ELECTRONIC VOLTAGE STABILIZER :

Figure 48 shows a typical arrangement *i.e.* modified form of simple electronic voltage stabilizer. In the circuit, the triode working as d.c. amplifier is changed by a pentode because a.c. resistance of a pentode is much higher than a triode. High a.c. resistance or grid impedance is desirable for good regulation and for utilizing high gain capacity. Another modification is the use of condenser C_1. C_1 increases the fraction of output voltage applied between grid and cathode of tube T_1. The condenser also helps in

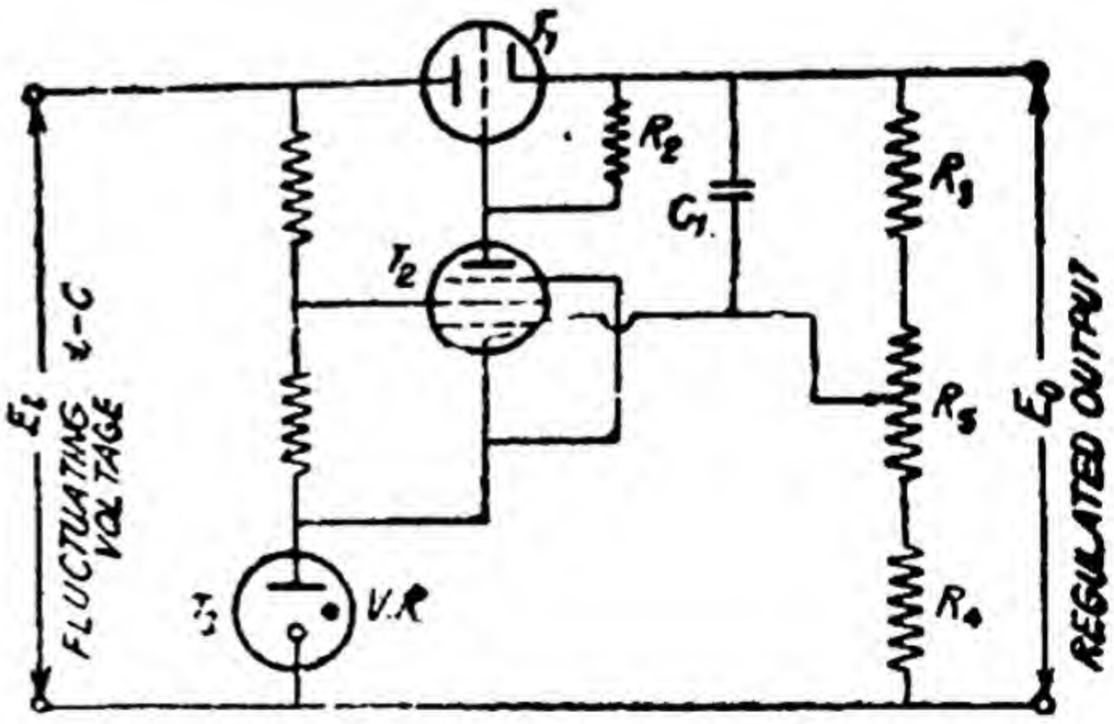


Fig. 48. Modified circuit of an electronic regulator.

reducing the ripple. For ideal regulation and ripple free output two or more stages of amplification are used.

8·18. VOLTAGE DOUBLER :

A circuit of voltage doubler is shown in fig. 49.

The magnitude of direct output voltage obtained by using a voltage doubler is approximately twice the peak value of input a.c. voltage. These are often used in power supplies in which alternating voltage is directly applied to rectifier (without step up transformer).

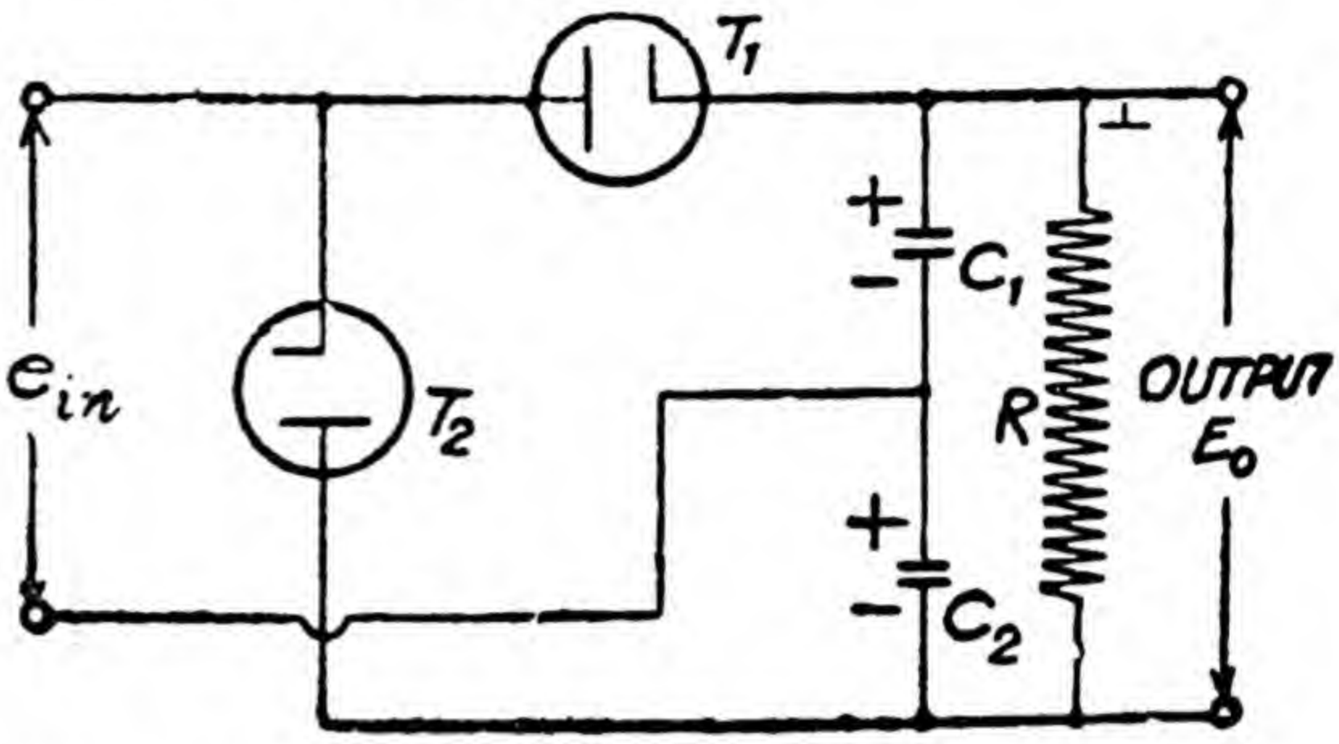


Fig. 49. Cir cuit of voltage doubler.

The two capacitors, C_1 and C_2 are connected in series across the output terminals. During one half cycle of input voltage, T_1 conducts and charges C_1 to peak value of input a.c. voltage. On other half cycle T_2 conducts and charges C_2 to peak value of input a.c. voltage. C_1 and C_2 supply the load current in series, such that the output voltage (if we assume load to be absent) approximates twice the peak voltage of applied voltage.

For voltage regulation the capacitors C_1 and C_2 should be large enough to maintain a constant voltage across the output during non-conducting intervals of the tubes which is possibly only when the energy stored by them during conduction interval is sufficiently large

EXPERIMENT

To design the power supply unit and to determine

(*i*) *Voltage regulation.*

(*ii*) *Percentage of ripple.*

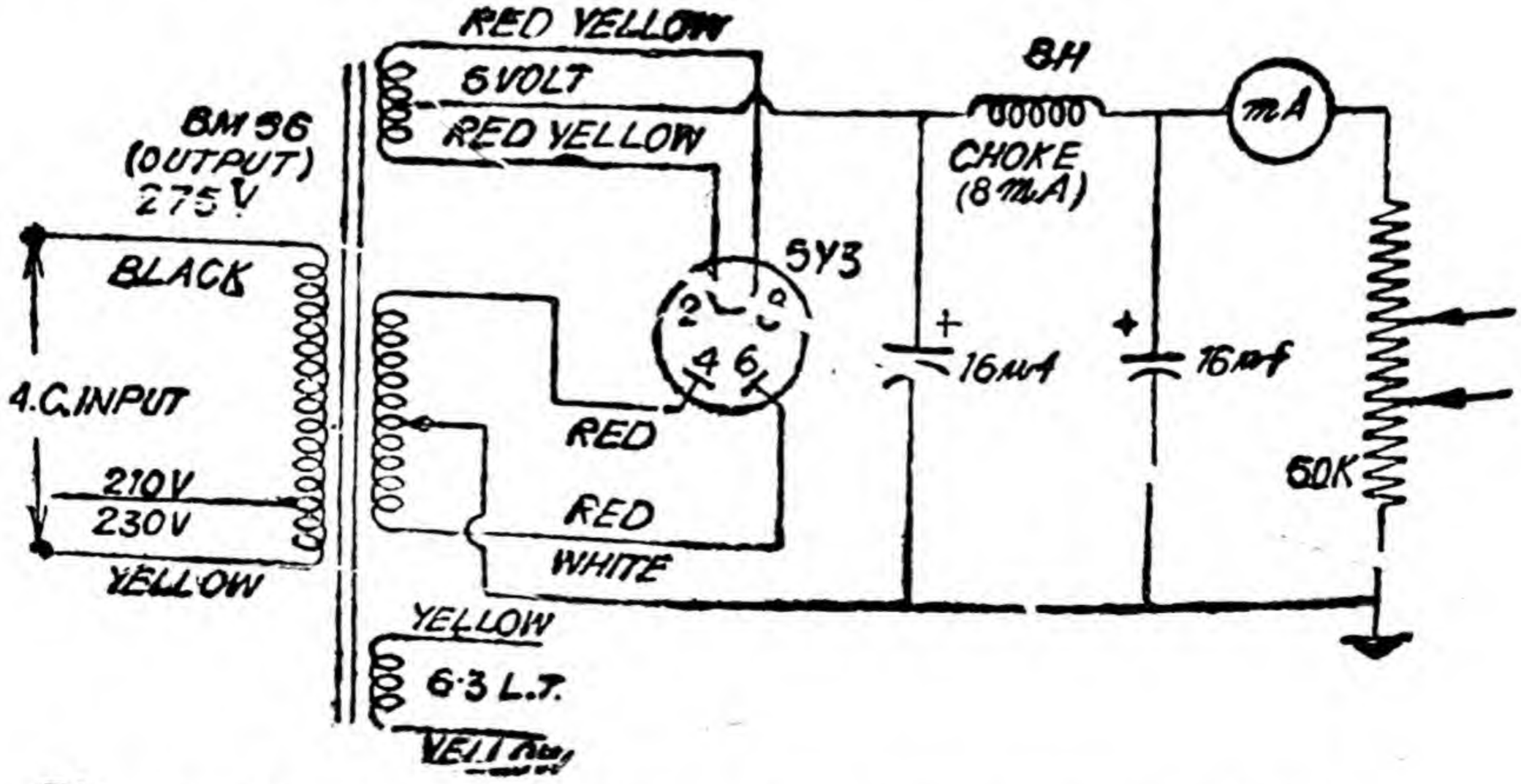


Fig. 50. Circuit of a power supply unit with π section filter.

Procedure : A simple power supply circuit using full wave rectifier with π section filter is shown in fig. 50.

The choice of various circuit elements depends upon the output voltage, voltage regulation and ripple factor. Their values can be estimated from the relations deduced in previous articles.

(*i*) *Voltage Regulation* : We know that percentage of regulation is

$$= \frac{\text{Voltage at no load} - \text{Voltage at full load}}{\text{Voltage at full load}} \times 100.$$

With different values of load, the effect on regulation can be shown. In the case of π section filter, the shape of the curve called operational characteristic of power supply unit, drawn in output d.c. voltage at filter input and d.c. load current in milliamperes is shown in fig. 51.

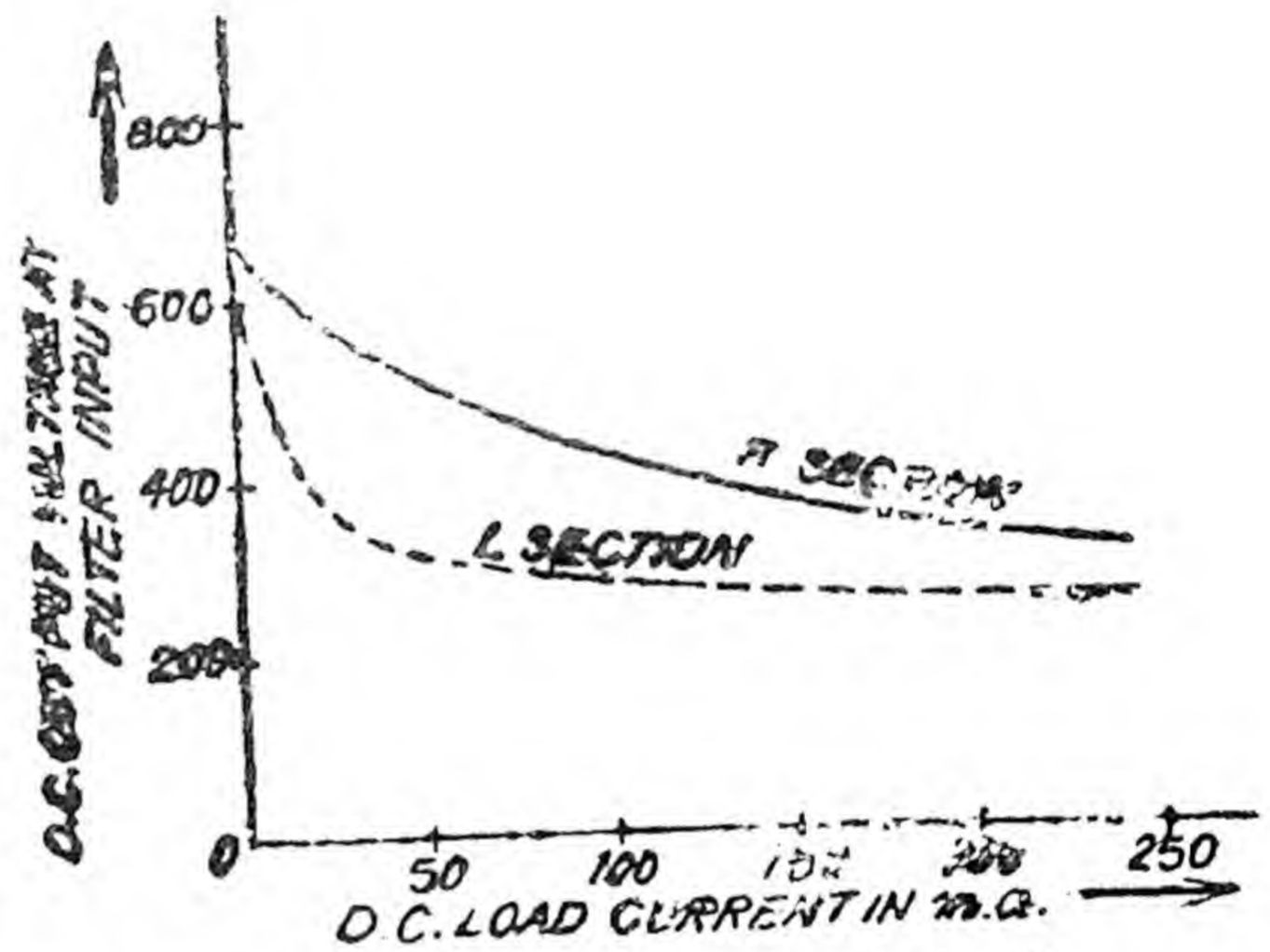


Fig. 51. Operational characteristic of power supply unit

It can also be seen, that if we use L section filter, then voltage regulation of π section filter is poor in comparison to L section filter.

(*ii*) *Percentage of Ripple* : In a π section filter, ripple factor is a function of load resistance and varies inversely. For the calculation of ripple voltage first cathode ray oscillograph is calibrated by giving different known a.c. voltages to one of the deflecting plates. A calibration curve in deflection lengths and a.c. voltages applied is drawn.

To determine the ripple present in the output voltage of power supply, an arrangement shown in fig. 52 is made. The d.c. voltage is blocked by using a capacitor. The deflection on C.R.O. screen is measured with varying loads. The corresponding d.c. load current is also noted. From the calibration curve the value of ripple voltage is calculated and then percentage of ripple is given by

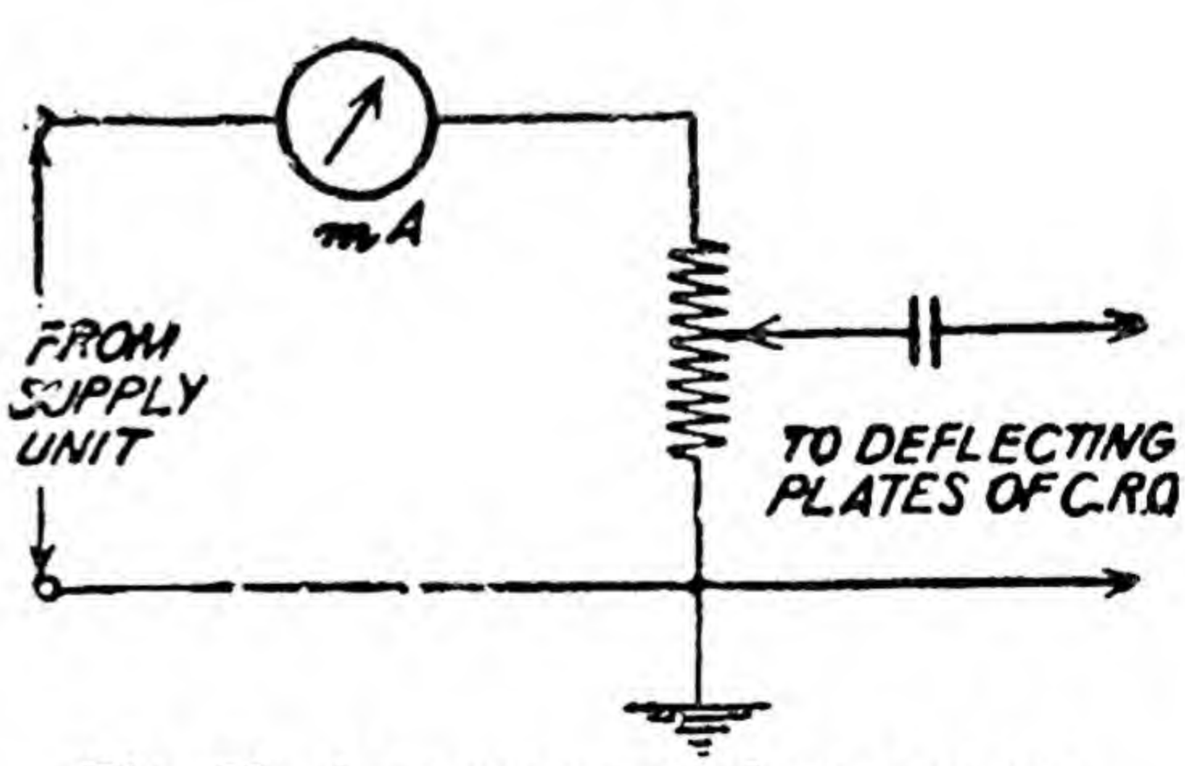


Fig. 52. Arrangement of measur ng ripple voltage.

$$\% \text{ of ripple} = \frac{\text{a.c. voltage across output}}{\text{d.c. voltage across output}} \times 100.$$

A graph is plotted in percentage of ripple and d.c. load current showing the dependence of ripple on load resistance as shown in figure 53.

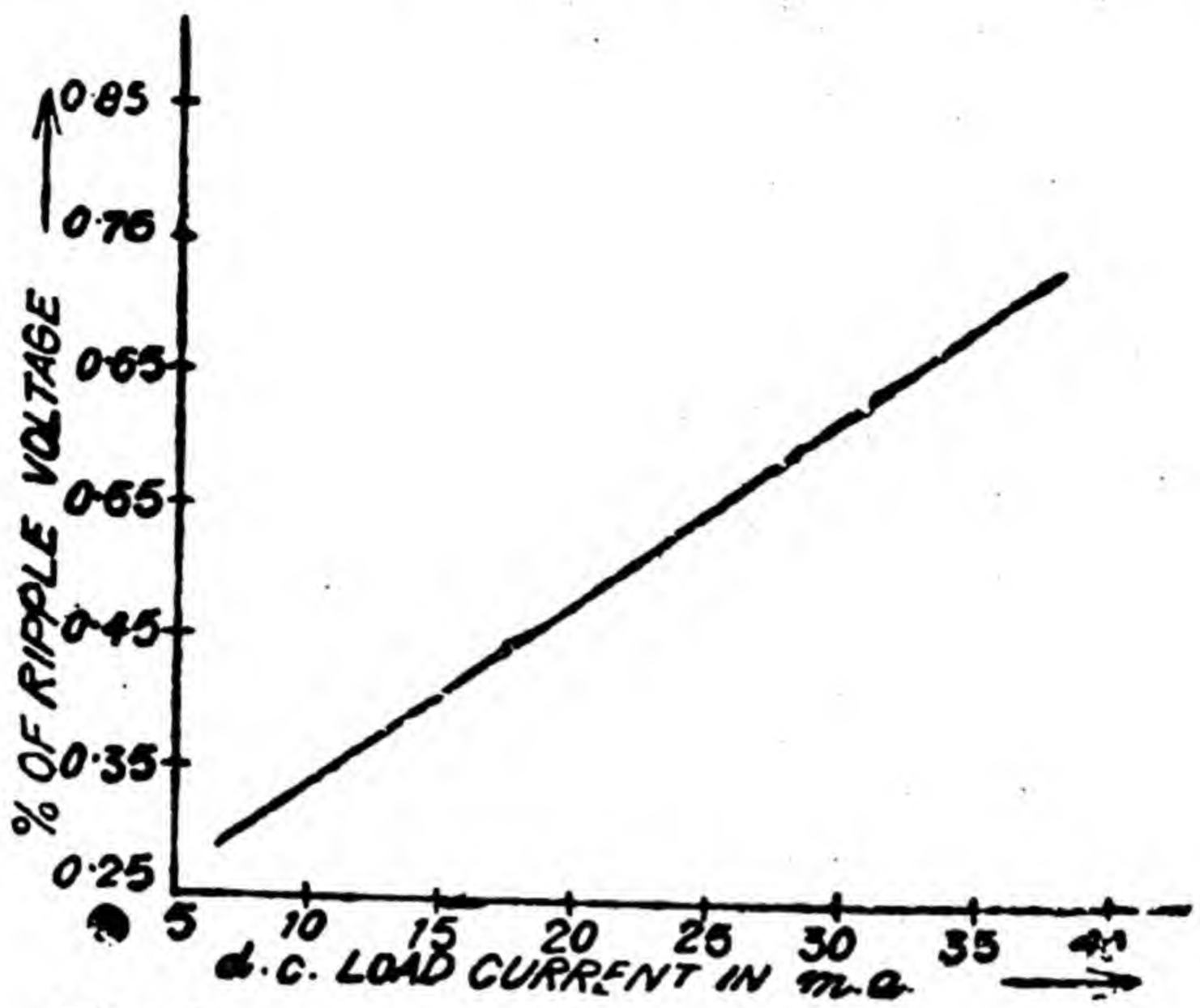


Fig. 53. Variation of ripple percentage with increasing load current.

If we use L section in place of π section, ripple voltage can be shown to be independent of load resistance.

High Voltage Regulated Power Supply : For the design purposes, a circuit of high voltage regulated power supply is given below :

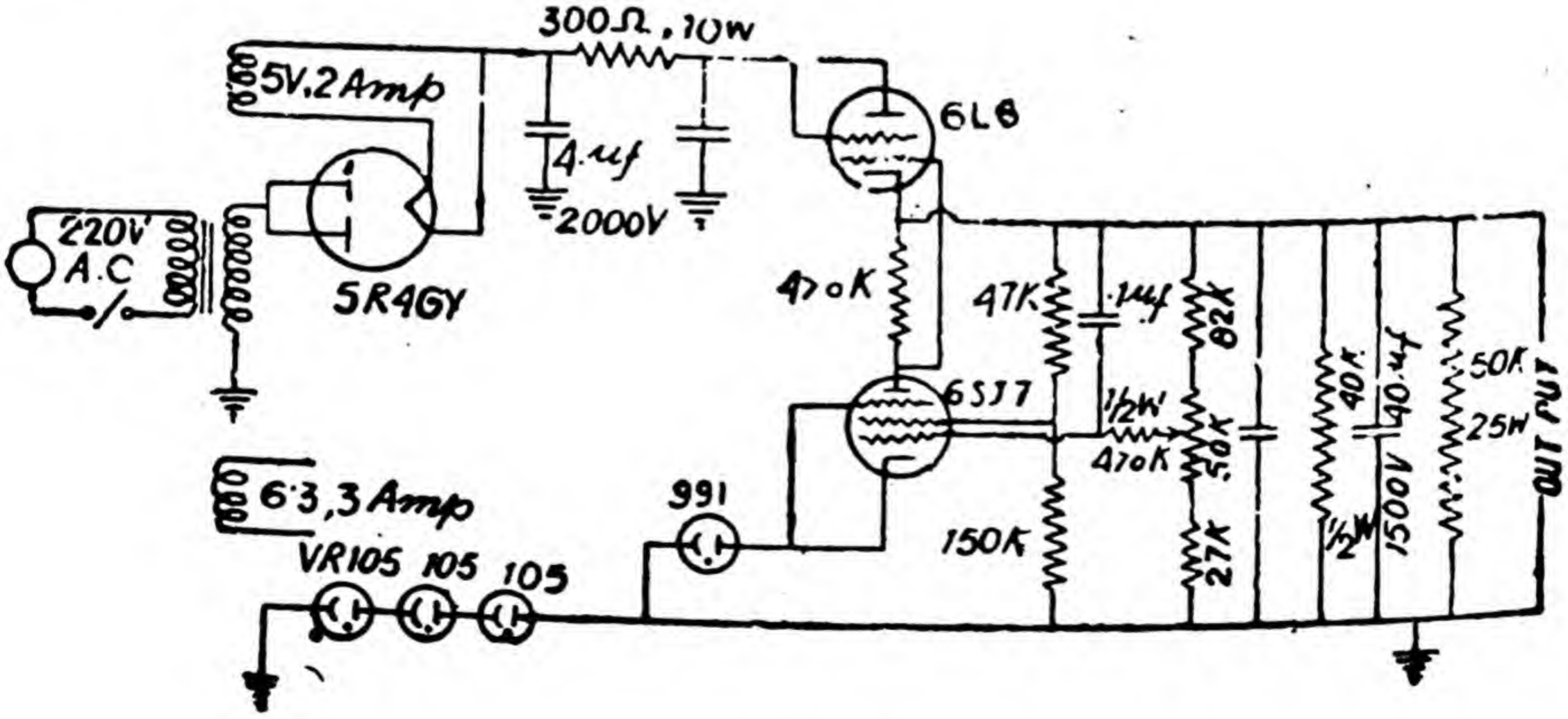


Fig. 54

NUMERICAL EXAMPLES

1. *A ripple voltage having the shape shown in fig. 55 has the deak to peak value equal to* E_R. *Find its r.m.s. value.*

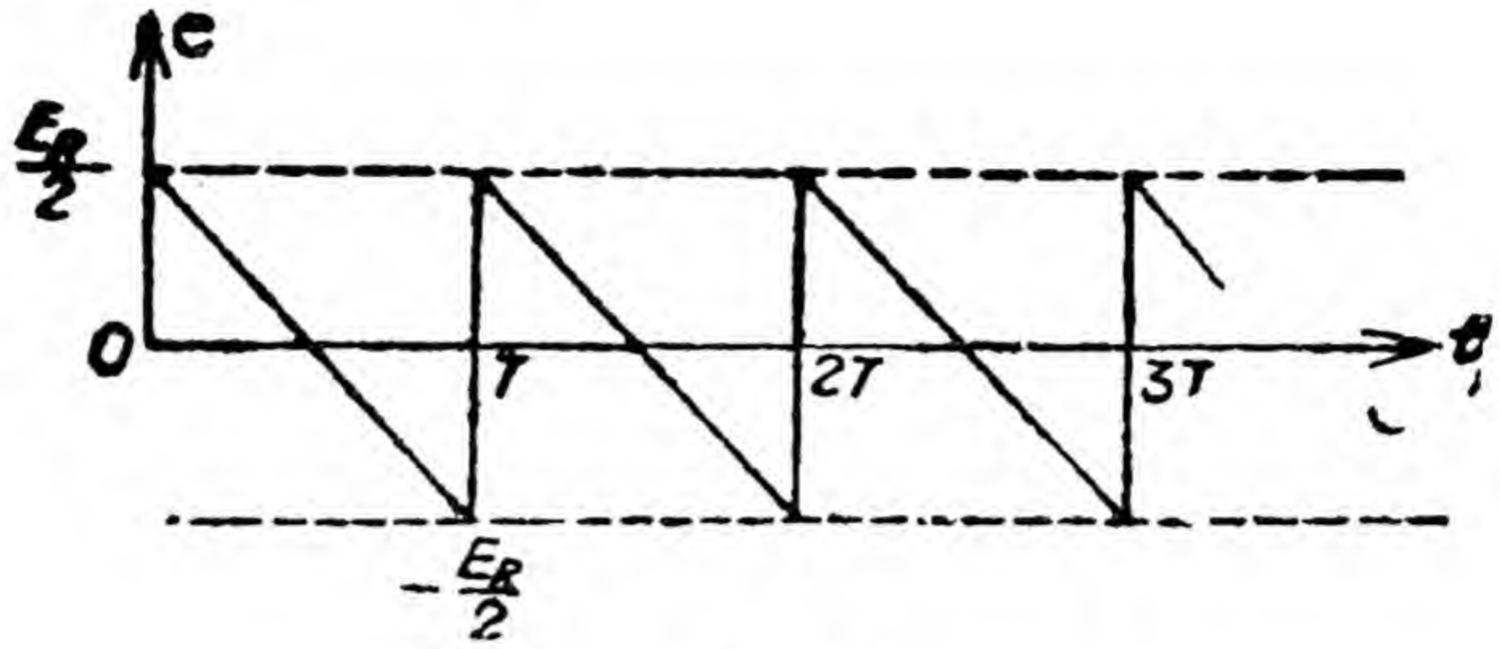


Fig. 55.

The triangular wave is represented by

$$e = A + Bt,$$

where A and B are constants. The values of A and B can be obtained by applying the following boundary conditions :

$$\text{when } t=0,\ e=-\frac{E_R}{2}$$

and $$\text{when } t=T,\ e=-\frac{E_R}{2}.$$

Applying these conditions

$$A=\frac{E_R}{2} \quad \text{and} \quad -\frac{E_R}{2}=\frac{E_R}{2}+BT$$

or $$B=-\frac{E_R}{T}$$

Hence $$e=\frac{E_R}{2}-\frac{E_R}{T}\ t.$$

$$E^2_{r\cdot m\cdot s\cdot}=\frac{1}{T}\int_0^T\left(\frac{E_R}{2}-\frac{E_R}{T}.t\right)^2 dt$$

$$=\frac{1}{T}\int_0^T E_R^2\left[\frac{1}{2}+\frac{t^2}{T^2}-\frac{t}{T}\right] dt$$

$$=\frac{E_R^2}{T}\left[\frac{t}{4}+\frac{t^3}{3T^2}-\frac{t^2}{2T}\right]_0^T$$

$$=\frac{E_R^2}{T}\left[\frac{T}{4}+\frac{T^3}{3T^2}-\frac{T^2}{2T}\right]=\frac{E_R^2}{12}$$

or $$E_{r\cdot m\cdot s\cdot}=\frac{E_R}{2\sqrt{3}}.$$

2. *In a half wave rectifier using high vacuum diode the d.c. output voltage is* 400 *volts on open circuit when a* 230*V a.c. supply is connected across the primary of the transformer. Calculate*

(*i*) *the ratio of the transformer.*

(*ii*) *the peak inverse voltage, neglecting the valve resistance.*

Assuming the effective valve resistance $r_p=500\Omega$, *determine the voltage regulation with a specified full load d.c. current of* 100 *mA.*

(*i*) $E_{d.c.}=\frac{E_m}{\pi}=400$ volts

$\therefore \quad E_m=400\times\pi=1256$ volts.

The R.M.S. value of secondary voltage $=\frac{1256}{\sqrt{2}}=887$ volts.

$\therefore$ Turn ratio $n=\frac{887}{230}=3{\cdot}86.$

(*ii*) Peak inverse voltage $E_m=1{,}256$ volts.

(*iii*) Tube drop at full load $=I_{d.c.}\ r_p$

$=100\times10^{-3}\times500=50$ volts

$E_{d.c.}$ (at no load) $=400$ volts

$E_{d.c.}$ (at load) $=400-50=350$ volts

$\therefore$ voltage regulation $=\frac{400-350}{350}=\frac{1}{7}$

% voltage regulation $=\frac{100}{7}=14{\cdot}3\%.$

3. *A voltage* 200 *sin* 100πt *is applied to a half wave vacuum tube rectifier with filter. The diode resistance of valve* $r_d=1000$ *ohms. The resistance load is* 10^4 *ohms. Calculate*

(a) Maximum current flowing, (b) average value, (c) r.m.s. value, (d) a.c. power input, (e)) d.c. power ourput, (f) ripple factor.

(*a*) Maximum current $I_m=\frac{E_m}{r_d+R}=\frac{200}{100+10^4}$

$=18\ 8\times10^{-3}$ amperes.

(*b*) Average value $I_{d.c.}=\frac{I_m}{\pi}=\frac{18{\cdot}8\times10^{-3}}{3{\cdot}1416}=5{\cdot}71$ mA.

(*c*) The r.m.s. value $I_{r.m.s.}=\frac{I_m}{2}=\frac{18{\cdot}18\times10^{-3}}{2}$

$=9{\cdot}09$ mA.

(*d*) The a c. power input $P_m=(I_{r\ m.s}\ )^2\ (r_d+R)$

$=(9\ 09)^2\ (1000+10^4).10^{-6}$

$=0{\cdot}909$ watt.

(*e*) The d.c. power output $P_{d.c.}=I_{d.c.}{}^2\ R$

$=(5{\cdot}78)^2\times10^4\times10^{-6}$

$=0{\cdot}334$ watt.

(*f*) Ripple factor $\gamma=\sqrt{\left\{\frac{(I_{r.m.s.})^2}{(I_{d\ o.})^2}-1\right\}}$

$=\sqrt{\left\{\frac{(I_m/2)^2}{(I_m/\pi)^2}-1\right\}}$

$=\sqrt{\left\{\left(\frac{\pi}{2}\right)^2-1\right\}}$

$=\sqrt{\{(1{\cdot}57)^2-1\}}$

$=1{\cdot}21.$

4. ***In a full wave vacuum tube rectlfier without filter, the load resistance is of*** 400 ***ohms.*** *Each diode has idealized characteristics having slope corresponding to a resistance of* 800 *ohms,* ***voltage applied to each diode is*** 240 *sin* $100\pi t$. *Calculate*

(*a*) *peak, average and r.m.s. values of current,*

(*b*) *d.c. power output and total power input,*

(*c*) ***rectifier efficiency,***

(*d*) *form factor,*

(*e*) ***ripple factor*** .

(*a*) Maximum value of current $I_m = \dfrac{E_m}{r_d + R}$

$$\therefore \quad I_m = \frac{240}{800 + 4000} = 50 \times 10^{-3} \text{ amp}$$

$$= \mathbf{50\ mA.}$$

Average value of current $I_{d\,c.} = \dfrac{I_m}{\pi/2} = \dfrac{50 \times 2}{3{\cdot}14}$

$$= \mathbf{31{\cdot}84\ mA.}$$

The r m.s. value of current $I_{r\ m\ s.} = \dfrac{I_m}{\sqrt{2}} = \dfrac{50}{1{\cdot}414}$

$$= \mathbf{35{\cdot}36\ mA.}$$

(*b*) The d.c. power output $P_{d.c.} = I_{d.c.}^2 \times R$

$$= (31{\cdot}84)^2 \times 10^{-6} \times 4000$$

$$= \mathbf{4{\cdot}056\ watts.}$$

Total power input $P_{in} = (I_{r.m.s.})^2 \times (r_d + R)$

$$= (35{\cdot}36 \times 10^{-3})^2 \times (800 + 4000)$$

$$= \mathbf{6{\cdot}001\ watts}$$

(*c*) Rectifier efficiency $\eta = \dfrac{P_{d\ c.}}{P_{in}}$.

$$\therefore \quad \eta = \frac{4{\cdot}056}{6{\cdot}001} \times 100 \text{ per cent} = \mathbf{67{\cdot}59\ per\ cent.}$$

(*d*) Form factor $= \dfrac{I_{r.m.s.}}{I_{d.c.}} = \dfrac{33{\cdot}36}{31{\cdot}48} = \mathbf{1{\cdot}11}$

(*e*) Ripple factor $\gamma = \sqrt{\left\{ \dfrac{(I_{r.m.s.})^2}{(I_{d.c.})^2} - 1 \right\}}$

$$= \sqrt{\{(1{\cdot}11)^2 - 1\}}$$

$$= \mathbf{0{\cdot}48.}$$

5. ***A single phase full wave rectifier is to supply*** 100 *mA. at* 300 *V with a ripple that must be less than* 10*V. Specify the elements of the rectifier using a single L section filter that will provide the desired result.*

We require 100 mA. at 300*V*, *i.e.*,

$$I_{d.c.} = 100 \text{ mA.},$$

$$E_{d.c.} = 300 \text{ volts}$$

Hence resistive load $R=\frac{E_{d\cdot c}}{I_{d\cdot c.}}=\frac{330\times10^3}{100}=3000\Omega.$

We know that

$$\gamma=\frac{1}{6\sqrt{2\omega^2LC}}=\frac{E_{r\cdot m\cdot s.}\text{ ripple}}{E_{d\cdot c.}}$$

$$=\frac{10}{300}=\frac{1}{30} \qquad \text{as } E_{r\cdot m\cdot s.}=10V$$

$$\text{Now } LC=\frac{1}{6\sqrt{(2)}\ \omega^2\gamma}=\frac{1}{6\sqrt{(2)}\ 4\pi^2.\ f^2.\gamma}.$$

$$=\frac{1}{6\sqrt{2.4\pi^2\times2500\times\gamma}} \text{ for } f=50 \text{ c/s.}$$

$$LC=\frac{30}{6\sqrt{2\times40\times2500}}. \qquad \text{.. (1)}$$

We also know that

$$L_{critical}=\frac{R}{3\omega}=\frac{R}{3.2\pi.\ f},$$

$$L_c=\frac{R}{942\cdot8} \text{ (for the frequency} = 50 \text{ c/s.)}$$

$$L=\frac{3000}{942\cdot8}=3\cdot18\text{H.}$$

From the equation (1),

$$C=\frac{30}{6\sqrt{2\times40\times2500\times L}}$$

$$=\frac{30\times10^6}{6\sqrt{2\times40\times2500\times3\cdot18}}\ \mu F$$

$$=11\cdot26\ \mu\text{F.}$$

Hence **R**$=3000\Omega$, **L**$=3\cdot18$**H**, and **C**$=11\cdot26\mu$**F**.

6. (*a*) *A full wave rectifier with inductor input filter is fed with* 50 *c/s. supply. Calculate the ripple factor* (*i*) *single section* (*ii*) *two identical sections if L and C of the filter are* 10 *H and* 20μF *respectively.*

(*b*) *Show that in a full wave rectifier with choke input filter the ripple fact or for a* 50 *c/s supply is given by an approximate expression* $\gamma=100/LC_{\mu F}\%$, *where L is in henrys and* $C_{\mu F}$ *is capaciiance in* μF.

(*a*) (*i*) We know that in full-wave rectifier with inductor input filter, for a single section

$$\gamma=\frac{0\cdot47}{(4\omega^2LC-1)}$$

$$=\frac{0\cdot47}{\{4(2\pi50)^2\times10\times20\times10^{-6}-1\}}=\frac{0\cdot47}{[79-1]}$$

$$=\cdot00603=0\cdot603\%.$$

(ii) For two sections filter $\gamma = \dfrac{0{\cdot}47}{(4\omega^2 LC-1)^2}$

$$\frac{0{\cdot}47}{[79-1]^2} = 0{\cdot}0000773$$

$$= \mathbf{0{\cdot}00773\%}.$$

(*b*) We know that $\gamma = \dfrac{0{\cdot}47}{(4\omega^2 LC-1)} \times 100\%$ when $4\omega^2 LC > > > 1$

$$\gamma \simeq \frac{0{\cdot}47}{4\omega^2 LC} \times 100\%$$

$$\simeq \frac{0{\cdot}47 \times 100}{4 \times 4\pi \times 50 \times 50 \times L \times C_{\mu F} \times 10^{-6}}$$

$$\simeq \frac{0{\cdot}47 \times 100}{0{\cdot}126\ LC_{\mu F}}$$

$$\simeq \frac{373}{\mathbf{LC_{\mu F}}}\%.$$

7. *A full wave vacuum tube rectifier has an inductance input filter with inductance L equal to* 30 *henrys and condenser C equal to* 10 *micro farad. Calculate the ripple factor assuming that* (*i*) *the current flows through the cycle,* (*ii*) *all harmonic terms except the second harmonic may be neglected,* (*iii*) *reactance* X_L *of inductor is much greater than the impedance of the parallel combination of C and* R_L, (*iv*) *reactance of C is much smaller than* R_L. *The applied voltage frequency is* 50 *c/s. Calculate also the critical value of the inductance. Given that the load resistance is* 8 *kilo-ohms.*

The rectified voltage is given by

$$e_0 = \frac{2E_m}{\pi} - \frac{4E_m}{3\pi} \cos 2\omega t$$

$$\therefore \quad E_{d.c.} = \frac{2E_m}{\pi}$$

The r.m.s. value of the current through the circuit is given by

$$I_{a.c.} = \frac{\text{R.M.S. voltage value of second harmonic of supply frequency}}{\text{choke impedance of the circuit}}$$

$$= \frac{\frac{4}{3\pi}\cdot\frac{E_m}{\sqrt{2}}}{X_L} = \frac{\sqrt{2}}{3X_L} E_{d.c.}$$

The a.c. potential across the load or condenser is given by

$$E_{d.c.} = I_{a.c.}\ X_C = \frac{\sqrt{2}}{3X_L}\ E_{d.c.}\ X_C.$$

Hence ripple factor

$$\gamma = \frac{E_{a.c.}}{E_{d.c.}} = \frac{\sqrt{2}\ X_C}{3\ \ X_L} = \frac{\sqrt{2}}{3 \times 2\omega L} \times \frac{1}{2\omega C}$$

$$= \frac{\sqrt{2}}{3\,(4\pi.50)^2}\ \frac{1}{(30\times 10\times 10^{-6})}$$

$$= 0{\cdot}004.$$

For critical value of inductance, the peak a.c. current equals the d.c. component of current so that

$$\frac{E_{d.c.}}{2\pi R_L} = \frac{4E_m}{2\pi R_L} = \frac{2E_{d.c.}}{3X_L}$$

or $$X_L = \frac{2R_L}{3}$$

Hence the critical inductance

$$L_C = \frac{2R_L}{3\times 2\omega} = \frac{2\times 8\times 10^3}{3\times 2\times (2\pi\times 50)} = 8{\cdot}5 \text{ Henrys.}$$

8. ***Design power supply using a π section filter to give d.c. output 250 V at 50 mA. with a ripple factor not to exceeding 0·1 percent.***

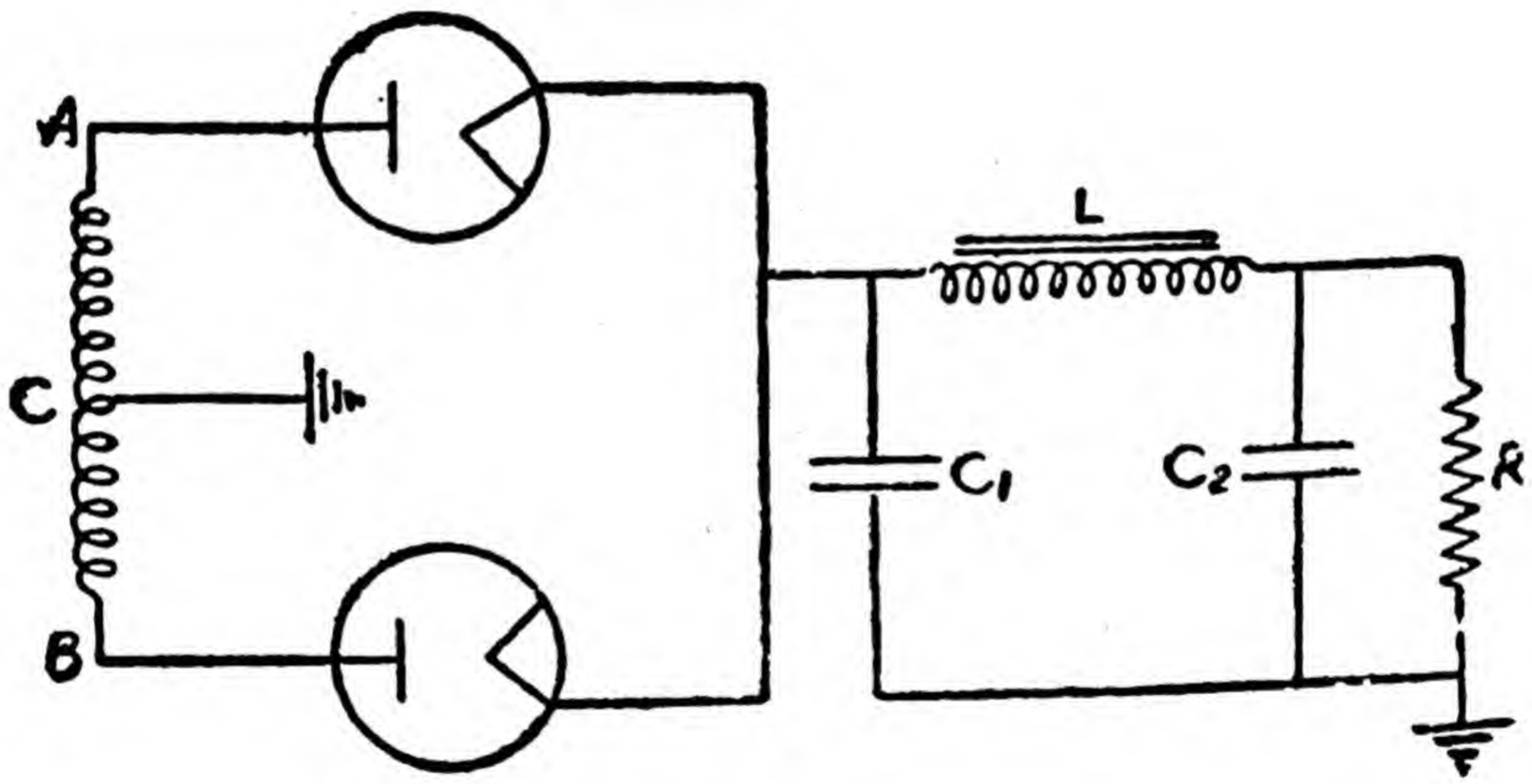


Fig. 56.

We know that resistive load $R = \frac{E_{d.c.}}{I_{d.c.}} = \frac{250\times 100}{5}$ ...(1)

$$= 5\times 10^3 \Omega.$$

γ(percentage ripple factor) $= \frac{3300\times 100}{C_1C_2\times L\times 5\times 10^3}$

$$= \frac{66}{LC_1C_2}.$$

According to the question,

$$\frac{66}{LC_1C_2} = 0{\cdot}01. \qquad (2)$$

$$\frac{66}{LC_1^2} = 0{\cdot}01, \text{ as in practice } C_1 = C_2.$$

$$LC_1^2 = \frac{66}{0{\cdot}01} = 6600. \qquad \ldots(3)$$

To find L and C_1, we shall have to substitute the value of either L or C_1.

Suppose $L = 20H$ and its ohmic resistance $R_L = 500\Omega$.

$$\therefore \quad C_1^2 \times 20 = 6600.$$

$$C_1^2 = \frac{6600}{20} = 330\mu f,$$

$$C_1 = 18{\cdot}1 = \mathbf{20\mu F \text{ (say)}}$$

Voltage drop in choke $= I_{d.c.} \times R_L$

$$= \frac{50 \times 500}{1000} = \mathbf{25V.}$$

Rectifier output voltage before filtration

$$= 250 + 25 = \mathbf{275V.}$$

$$E_{d.c.} = \sqrt{2}E_s - I_{d.c.}\left(\frac{1}{4fC_1} + R_L\right),$$

$$250 = \sqrt{2}E_s - \frac{50}{1000}\left(\frac{10^6}{4 \times 60 \times 20} + 500\right)$$

$$= \sqrt{2}E_s - 56{\cdot}25,$$

$$\sqrt{2}E_s = 250 + 56{\cdot}25$$

$$= 306{\cdot}25$$

$$E_s = 216{\cdot}5 = 200V \text{ (say)}.$$

E_s is the effective value of the voltage supplied between AC and CB. Thus transformer secondary should develop across it.

$E_{AC} = 200V$ and $E_{BC} = 300V$.

Peak inverse voltage (P.I.V.) of a full wave rectifier

$= 2 \times 275 = 550V$.

$$\text{Wattage of transformer secondary} = \frac{400 \times 50}{100}$$

$$= \mathbf{20\ W.}$$

If transformer efficiency is assumed to be 80%, then tranformer primary wattage $= \frac{20}{0{\cdot}8} = \mathbf{25W.}$

$$\text{Primary current } I_p = \frac{25 \times 1000}{220}$$

$$= \mathbf{114mA.}$$

9. *Calculate the regulation and efficiency of rectification of a diode* ($r_d = 500\Omega$) *used as half wave rectifier. The rated output current is* 85 *m.A., the transformer secondary voltage is* 230 *volts r.m.s.*

We know that

$$E_{d.c.} = I_{d.c.} \times R_L \qquad \ldots(1)$$

$$I_{d.c.}=\frac{E_m}{\pi(R+r_d)} \qquad ...(2)$$

$$\frac{E_{d.c.}}{I_{d.c.}}=R. \qquad ...(3)$$

Substituting equation (1) in equation (2), we have

$$I_{d.c.}=\frac{E_m}{\left(\frac{E_{d.c.}}{I_{d.c.}}+r_d\right)}$$

or $$I_{d.c.}\times\pi\left(\frac{E_{d.c.}}{I_{d.c.}}+r_d\right)=E_m,$$

$$\pi E_{d.c.}+\pi r_d.I_{d.c.}=E_m$$

$$\pi E_{d.c.}=E_m-\pi r_d.I_{d.c.}$$

$$E_{d.c.}=\frac{E_m-\pi I_{d.c.}\,r_d}{\pi}=\frac{E_m}{\pi}-I_{d.c.}\,r_d. \qquad ...(4)$$

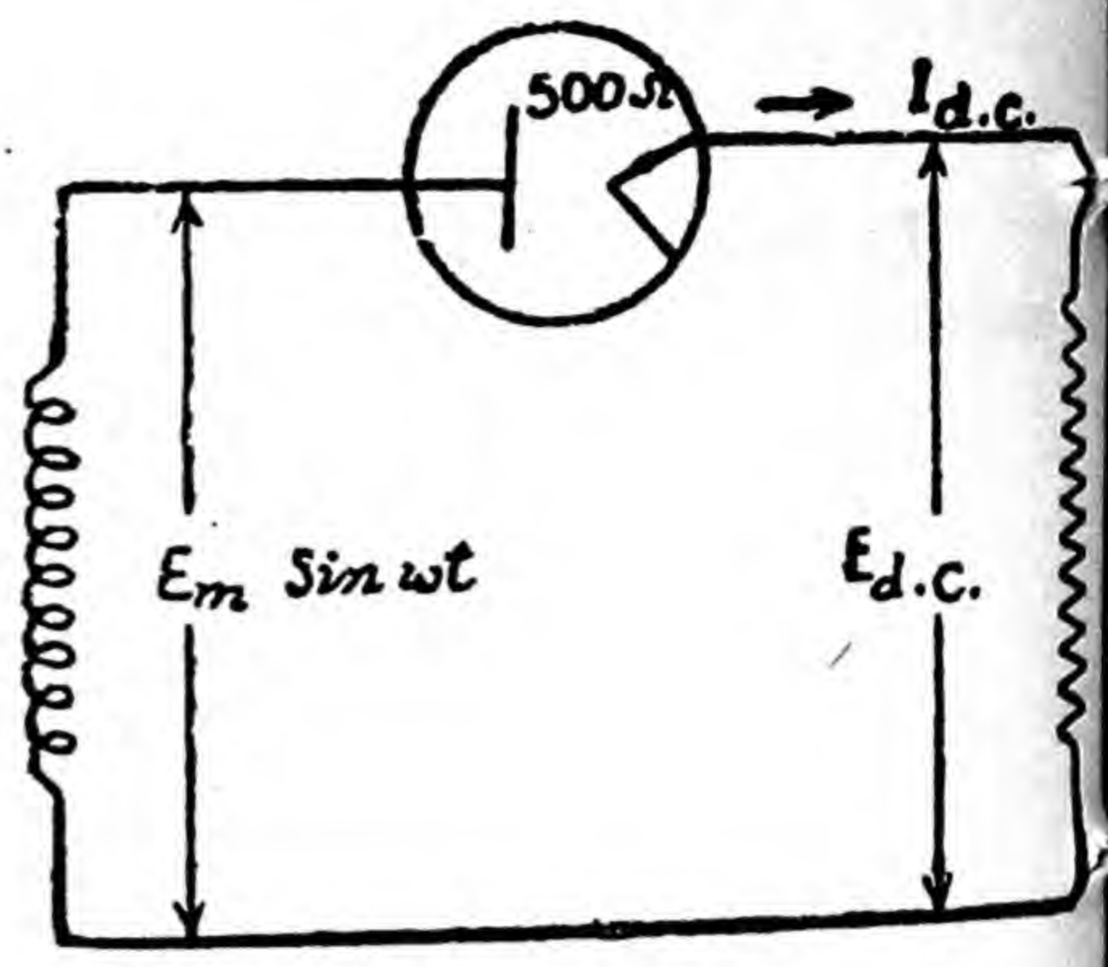


Fig. 57.

This is the voltage output on full load.

$\frac{E_m}{\pi}$ is the maximum voltage under no load condition which is applied to the input of the rectifier or the maximum voltage which the transformer secondary develops..

Voltage regulation

$$=\left[\frac{E_{d.c.\ (no\ load)}-E_{d.c.\ (full\ load)}}{E_{d.c.\ (full\ load)}}\right]\times 100,$$

$$E_{d.c.\ (no\ load)}=\frac{E_m}{\pi}=\frac{\sqrt{2}E_{r.m.s.}}{\pi},$$

where $E_{r.m.s.}$ = r.m.s. value of input voltage = 230 V,

$$E_{d.c.\ (no\ load)}=\frac{\sqrt{2}E_{r.m.s.}}{\pi}=\frac{1{\cdot}414\times 230}{3{\cdot}142}=107 \text{ volts.}$$

$$E_{d.c.\ (full\ load)}=\frac{E_m}{\pi}-I_{d.c.}\,r_d$$

$$=107-\frac{85\times 500}{1000}$$

$$=107-64{\cdot}5=\mathbf{64{\cdot}5\ V.}$$

$$\text{Regulation}=\left(\frac{107-64{\cdot}5}{64{\cdot}5}\right)\times 100=65{\cdot}8\%.$$

Rectification efficiency

$$\eta=\frac{40{\cdot}6}{\left(1+\frac{r_d}{R}\right)}\%=\frac{40{\cdot}6}{1+\frac{I_{d.c.}\,r_d}{I_{d.c.}\,R}}=\frac{40{\cdot}6}{1+\frac{I_{d.c.}\,r_d}{E_{d.c.}}}$$

$$= \frac{40{\cdot}6}{1+\dfrac{r_d.I_{d\cdot c\cdot}}{\dfrac{E_m}{\pi}-I_{d\cdot c\cdot}\,r_d}} = \frac{40\,6\left[\dfrac{E_m}{\pi}-I_{d\cdot c\cdot}\,r_d\right]}{\dfrac{E_m}{\pi}-I_{d\cdot c\cdot}\,r_d+I_{d\cdot c\cdot}\,r_d}$$

$$= \frac{40{\cdot}6\left[\dfrac{E_m}{\pi}-I_{d\ c\cdot}\,r_d\right]}{\dfrac{E_m}{\pi}} = 40{\cdot}6\left(1-\frac{I_{d\cdot c\cdot}\ r_d.\pi}{E_m}\right)$$

$$= 40{\cdot}6\left(1-\frac{42{\cdot}5\times 3{\cdot}14}{1{\cdot}414\times 230}\right) \text{ since } E_m=\sqrt{2}E_{r\cdot m\cdot s\cdot}$$

$$= 40{\cdot}6\ (1-0{\cdot}41) = 40{\cdot}6\times 0{\cdot}59$$

$$= \mathbf{23{\cdot}954 \text{ percent.}}$$

10. *The alternating voltage applied to a hot cathode gas diode full wave rectifier using no filter is* 400 *sin* $(2\pi.50.t.)$. *The resistance load in the cathode circuit is* 4000 *ohms. The constant voltage drop across each diode during conduction is* 10 *volts.*

Calculate :

(*a*) *maximum value of the current,*

(*b*) *d.c. currents,*

(*c*) *d.c. power output,*

(*d*) *rectifier efficiency,*

(*a*) Maximum value of the current I_m.

$$I_m = \frac{E_m-E_0}{R} = \frac{400-10}{4000} = \frac{390}{4000} = \mathbf{97{\cdot}5 \text{ mA.}}$$

(*b*) The d.c. current $I_{d\cdot c\cdot} = 2\,\dfrac{E_m}{\pi R}\left(1-\dfrac{\pi}{2}\,\dfrac{E_0}{E_m}\right)$

$$= 2.\ \frac{400}{3{\cdot}14\times 4000}\left(1-\frac{3{\cdot}14\times 10}{2\times 400}\right)$$

$$= \frac{2}{31{\cdot}4}\left(1-\frac{3{\cdot}14}{80}\right)$$

$$= \frac{2}{31{\cdot}4}\left(\frac{76{\cdot}86}{86}\right)$$

$$= \frac{2}{31{\cdot}4}\times 0{\cdot}96$$

$$= \mathbf{61 \text{ mA.}}$$

(*c*) The d.c. power output

$$= P_{d\cdot c\cdot} = I_{d\cdot c\cdot}{}^2\times R = (61)^2\times 10^{-6}\times 4000 \text{ watts}$$

$$= \mathbf{2{\cdot}684 \text{ watts.}}$$

(*d*) The rectifier efficiency

$$\eta = 81{\cdot}2\left(1-1{\cdot}87\,\frac{E_0}{E_m}\right)\% = 8{\cdot}12\left(1-1{\cdot}87\times\frac{10}{400}\right)$$

$$=81\cdot2\left(1-\frac{1\cdot87}{40}\right)=81\cdot2\ (1-0\cdot046)$$
$$=81\cdot2\times0\cdot954$$
$$=77\cdot4\%.$$

11. *In the electronic voltage stabilizer of fig. 47, the d.c. voltage source has internal resistance of* 100Ω, *series control tube* T_1 *has amplification factor 5 and* $r_{p_1}=500\Omega$. *The load resistance is* 1000Ω. *The gain of d.c. amplifier is* -40, *minus sign indicates the phase reversal. Dynamical plate resistance of d.c. amplifier tube is* 300 *kilo ohms. Resistances* R_2, R_3 *and* R_4 *are* 300, 200 *and* 20 *kilo-ohms respectively. Calculate the stabilization ratio and the effective internal resistance of the stabilizer. Also calculate the variation in stabilized output voltage if the load current is increased by* 200 *mA. and no load input d.c. voltage to the stabilizer is increased by* 10 *volts.*

The stabilization ratio is given by

$$S=\frac{\mu_1\ (1-\beta A_2)+1+r_{p_1}/R_L}{(\mu_1\alpha_2+1)},$$

where
$$\beta=\frac{R_4}{R_3+R_4}=\frac{200\times10^3}{200\times10^3+200\times10^3}=0\cdot5$$
$$\alpha_2=\frac{r_{p_2}}{r_{p_2}+R}=\frac{300\times10^3}{300\times10^3+300\times10^3}=0\cdot5$$
$$A_2=-40 \text{ and } \mu_1=5$$
$$\therefore\quad S=\frac{5\ (1+0\cdot5\times40)+1+500/1000}{(5\times0\cdot5+1)}=30\cdot4$$

Effective internal impedance of the stabilizer is given by

$$R_0'=\frac{R_S}{S}+\frac{1}{g_{m1}\ (1-\beta A_2)}$$
$$=\frac{100}{30\cdot4}+\frac{1}{5/500\ (1+0\cdot5\times40)}$$
$$=3\cdot2+4\cdot7=7\cdot9 \text{ ohms.}$$

Change in output voltage is given by

$$dE_0=\frac{dE_i}{S}-di_1\ R_0'$$
$$=\frac{10}{30\cdot4}-0\cdot2\times7\cdot9=0\cdot32-1\cdot58=-1\cdot26 \text{ volt.}$$

The output voltage is thus decreased by 1·26 volts.

12. *In a half wave thyratron rectifier circuit, a voltage source of* $230V$ *r.m.s. is connected in series with a load of* 400Ω *in the plate circuit. The grid voltage is adjusted such that the conduction starts at* 60° *after the start of each cycle. Calculate the readings of the following meters :*

(*i*) *a.c. ammeter in seies with the load,*

(*ii*) *an a.c. voltmeter across the tube,*

(*iii*) *a wattmeter connected to read the total power in the plate circuit. Assume that the tube drop is* 15 *volts during conduction period.*

According to the given problem

$$E_m = 230\sqrt{2} = 326V \text{ and } E_0 = 5V$$

Since $E_m \gg E_0$, there will be very little error if conduction is assumed to continue upto the end of each positive half cycle. The instantaneous plate current during conduction is

$$i_b = \frac{230\sqrt{2}\sin\omega t - 15}{400} \quad \pi/3 \leqslant \omega t \leqslant \pi$$

$$= 0{\cdot}8125 \sin \omega t - 0{\cdot}0375.$$

The a.c. ammeter will read the r.m.s. value of current which is

$$I_{r\ m{\cdot}s{\cdot}} = \left[\frac{1}{2\pi}\left\{\int_{\pi/3}^{\pi} (0{\cdot}8125 \sin \omega t - 0{\cdot}0375)^2\, d(\omega t)\right\}\right]^{1/2}$$

$$= \left[\frac{1}{2\pi}\int_{\pi/3}^{\pi} (0{\cdot}660 \sin^2\omega t - 0{\cdot}061 \sin \omega t + 0{\cdot}00141)\, d(\omega t)\right]^{1/2}$$

$$= \left[\ 0{\cdot}134 - 0{\cdot}0146 + 0{\cdot}0005\ \right]^{1/2}$$

$$= \mathbf{0{\cdot}35\ amps.}$$

If voltage drop is negligible, then

$$I_{r\cdot m\cdot s\cdot} = \left[\frac{1}{2\pi}\left\{\int_{\pi/3}^{\pi} (0{\cdot}8125 \sin \omega t)^2\, d(\omega t)\right\}\right]^{1/2}$$

$$= \{0{\cdot}133\}^{1/2} = \mathbf{0{\cdot}364\ amps.}$$

The a.c. voltmeter will read the r.m.s. value of voltage which is given by

$$E_{r\cdot m\cdot s\cdot} = \left[\frac{1}{2\pi}\left\{\int_{0}^{\pi/3} (230\sqrt{2}\sin\omega t)^2\, d(\omega t) + \int_{\pi/3}^{\pi} (15)^2\ d(\omega t) + \int_{\pi}^{2\pi} (200\sqrt{2}\sin\omega t)\, d(\omega t)\right\}\right]^{1/2}$$

Solving it we have

$$E_{r\cdot m\cdot s\cdot} = 178 \text{ volt}$$

The instantaneous power $= e_s i_b$

$$\therefore \quad P = \frac{1}{2\pi}\int_{\pi/3}^{\pi} e_s\, i_b\ (d(\omega t)$$

$$= \frac{1}{2\pi}\int_{\pi/3}^{\pi} (230\sqrt{2}\sin\omega t)\,(0{\cdot}81 \sin \omega t - 0{\cdot}0375)\, d(\omega t,$$

$$= \mathbf{50\ 5\ watts.}$$

13. ***In a thyratron circuit the plate supply voltage has a peak value of*** 600 ***volts. Constant voltage drop in the tube during conduction is*** 12 ***volts. Load resistance is*** 5,000Ω. ***The critical grid voltage curve is then such that for grid bias of*** —2 ***and*** —3 ***volts, the angle*** θ_1 ***of initiation of current is*** 10 ***and*** 20 ***degrees. Calculate the average value of load current under these bias conditions. As angle*** θ_1 ***is increased, find the value of*** θ_1 ***which gives maximum value of average plate current.***

Instantaneous plate current i_b is given by

$$i_b=\frac{E_m \sin \omega t-E_0}{R_L}=\frac{600 \sin \omega t-12}{500},$$

and average load current is given by

$$I_{d.c.}=\frac{1}{2\pi}\int_{\theta_1}^{\pi-\theta_2} i_b \, d(\omega t)$$

$$\theta_2=\sin^{-1}\left(\frac{E_0}{E_m}\right)=\sin^{-1}\left(\frac{12}{600}\right)=1\cdot15^\circ$$

$$\theta_1=10 \quad \text{or} \quad 20 \text{ degree}$$

$$\therefore \quad I_{d.c.}=\frac{E_m}{2\pi R}\int_{\theta_1}^{\pi-\theta_2}\left(\sin\theta-\frac{E_0}{E_m}\right)d\theta$$

$$=\frac{E_m}{2\pi R_L}\left[\cos\theta_1+\cos\theta_2-\frac{E_0}{E_m}(\pi-\theta_2-\theta_1)\right]$$

(*i*) When $\theta_1=10$ degrees

$$I_{d.c.}=\frac{600}{2\pi\times5000}\left[0\cdot9848+0\cdot9398-\frac{12}{600}\left(\pi-\frac{11\cdot5\pi}{180}\right)\right]$$

Solving we have

$$I_{d.c.}=0\cdot0038 \text{ amps.}$$

(*ii*) When $\theta_2=20$ degrees

$$I_{d.c.}=\frac{600}{2\pi\times5000}\left[0\cdot9097+0\cdot9998-\frac{12}{600}\left(\pi-\frac{11\cdot5\pi}{180}\right)\right]$$

$$=0\cdot0368 \text{ amps.}$$

$I_{d.c.}$ is maximum when $\theta_1=0$

$$\therefore \quad (I_{d.c.})_{max.}=\frac{600}{2\pi\times5000}\left[1+0\cdot9998-\frac{12}{600}\left(\pi-\frac{11\cdot5\pi}{180}\right)\right]$$

$$=0\cdot037 \text{ amps.}$$

EXERCISES AND PROBLEMS

1. Draw the circuit diagram of a half wave vacuum tube rectifier. Explain the operation of the rectifier.
2. Derive expressions for rectification efficiency, ripple factor and voltage regulation of a half wave vacuum tube rectifier with resistive load.
3. Draw the circuit diagram of a full wave vacuum tube rectifier using a power transformer with centre tapped secondary for use of a radio rectifier and give waveforms of current and output-voltage.
4. Derive expression for d.c. or average value of current, R.M.S. value of

current and average voltage across the load for a full wave vacuum rectifier with resistive load.

6. Derive an approximate expression for a ripple factor in a full wave vacuum tube rectifier with resistive load.
6. Show that a full-wave rectifier is twice as efficient as a half wave rectifier. Calculate the ripple factor of single phase full wave rectifier.
7. What is the necessity of having filter in power supply ? Obtain the ripple factor of a full wave rectifier with shunt capacitor filter.
8. Derive an expression for the ripple factor in a full wave vacuum tube rectifier using inductance inpu filter.
9. Derive an expression for the output voltage in a full wave vacuum tube rectifier using capacitor input filter.
10. Explain the working of half wave and full wave volta e doublers and discuss their relative merits and demerits.
11. Draw a circuit of an electronic regulator and derive an expression for its regulation sensitivity.
12. Differentiate between half and full wave rectification. State where half wave rectification may be used and the purposes for which it is not well adopted.
13. In a full wave rectifier, the most negative point in the circuit is
(a) the centre tap on the high voltage secondary, (b) chassis ground,
(c) either cathode, (d) either plate. [Ans. : (a)]
14. What is the frequency of the ripple voltage at the output of a full wave rectifier operating from a 50 cycle supply ?
(a) 50 cycles, (b) 100 cycles, (c) 150 cycles, (d) 200 cycles. [Ans. : (b)]
15. Select the correct advantage of a full wave rectifier over a half wave rectifier
(a) the output voltage is lower with more ripple,
(b) the tube will conduct during both halves,
(c) ripple frequency is lower.
(d) each diode can cool off during half of each input cycle. [Ans. : (b)]
16. In comparison with an *RC* power supply filter, a filter with a choke provides
(a) less filtering action and larger d.c. voltage drop,
(b) more filtering action and smaller d.c. voltage drop,
(c) more filtering action and larger d.c. voltage drop,
(d) less filtering action and smaller d.c. voltage drop. [Ans. : (b)]
17. One advantage of using a choke input L.C. filter for a power supply is that the
(a) circuit has a good voltage regulation.
(b) d.c. output voltage of the filter is higher,
(c) a.c. input voltage is increased,
(d) direct load current is increased. [Ans. : (a)]

9

BASIC AMPLIFIER PRINCIPLES

One of the most important functions of electronic circuitry is amplification. In fact, were it not for this property, many other specific circuit function would not be possible. For example, oscillators to produce sine waves, square waves, pulse, or any other desired waveshapes, would not be possible except for the amplification properties of the circuit components.

A device by means of which amplification is affected is known as amplifier or an amplifier may be defined as a device that increases the voltage, current or power of an input signal with the aid of vacuum tubes or transistors by furnishing the additional power from a separate power source.

In the study of characteristics of vacuum tubes, (triode, tetrode and pentode), we have seen that the plate current depends upon the voltage of the control grid provided all the other voltages are kept constant. Any variation in the grid voltage produces corresponding change in the plate current. When a load resistance is connected in the plate circuit, the plate current variation will produce a corresponding variation of voltage across the load. Consider the case of a sinusoidal voltage, applied between the grid and cathode, then the sinusoidal voltage, changed in phase, will appear across the load. This output voltage is much greater than the input voltage. Thus the circuit gives the voltage amplification.

9·1. SYMBOLS FOR VOLTAGES AND CURRENTS :

(a) When no signal is applied, then the voltage and current are indicated by capital letters E and I respectively. In this case subscript b and c are used for plate and grid respectively. Thus

E_b, E_c—represent no-signal plate and grid voltages respectively.

I_b, I_c—represent no-signal plate and grid currents respectively.

(b) (i) Instantaneous values of varying voltage and current are represented by small letters e and i respectively. Subscripts g and p are used for grid and plate respectively. Thus

e_g, e_p—represent instantaneous values of varying components of grid and plate voltages respectively.

i_p i_p—represent instantaneous values of varying components of grid and plate currents respectively.

(*ii*) Instantaneous total values are represented by letters *e* and *i* but they are followed by subscripts *b* and *c* for plate and grid respectively. Thus

e_b, e_c—represent instantaneous total values of plate and grid voltages respectively.

i_b, i_c—represent instantaneous total values of plate and grid currents respectively.

(*a*) no signal component or d.c. component and,

(*b*) instantaneous value of time varying component.

9·2. BASIC VOLTAGE AMPLIFIER :

The basic circuit of a vacuum tube amplifier using a triode is shown in figure 1. The figure also shows the symbols for different electrode voltages and currents.

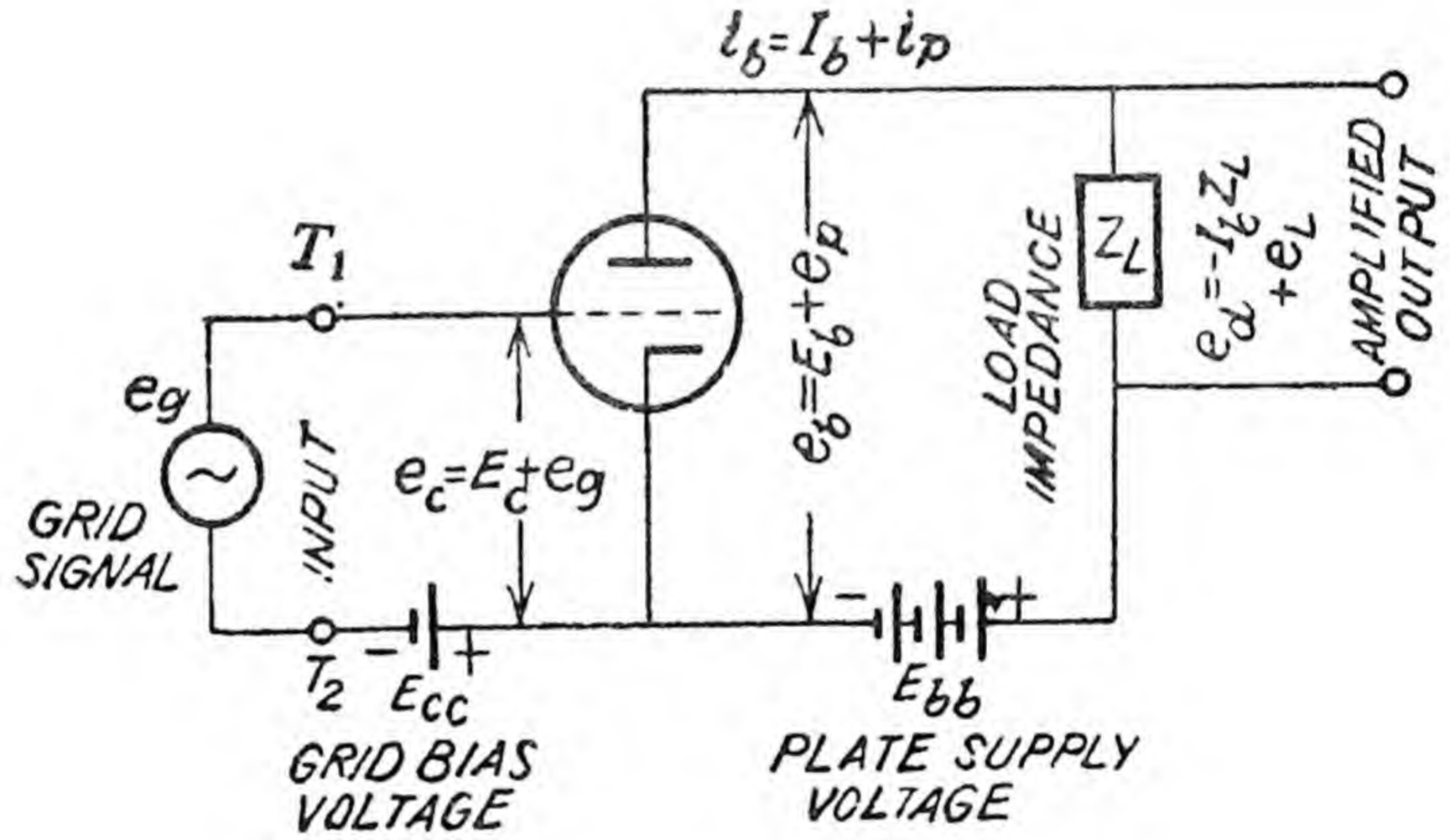


Fig. 1. Basic vacuum tube amplifier.

In the basic circuit of vacuum tube amplifier, the grid of the triode is always kept negative with respect to cathode by means of grid bias voltage E_{cc}. The voltage e_g to be amplified is connected at terminals T_1 and T_2 in series with E_{cc}. The voltage e_g is called signal voltage and in general is periodic. On the plate side, a d.c voltage E_{bb} is applied between plate and cathode through a load impedance. The voltage E_{bb} is called plate supply voltage and its positive terminal is connected to the plate. The load impedance may be a pure resistance, a complex impedance or a tuned circuit.

Now the question is : Why is the bias battery necessary in the grid circuit ? To explain this, let for a moment, it be not there and a.c. signal voltage e_g is applied alone to the grid. The grid will be alternately positive and negative inaccordance with the applied signal. Whenever input signal goes positive, a grid current will flow out from the grid, which not only distorts the linearity of the tube characteristic and thus leading to distorted amplification, but

also requires a certain amount of power, which is extracted from the input signal. This is an undesirable result. Hence the bias voltage of sufficient magnitude and of such a polarity as to maintain the grid negative or at zero voltage is applied. This is the function of grid bias in the grid circuit.

Now we shall see that what happens in the plate circuit when the grid to cathode voltage goes through one cycle of its operation. When the heater voltage is applied to the heater, the plate current will flow and will continue to flow as long as the a.c. signal voltage is not sufficiently large to drive the tube to plate current cut-off. *Hence the plate supply voltage E_{bb} is so chòsen that even for the most negative value of grid voltage the tube does not reach plate current cut off* This is the necessary condition for all voltage amplifiers to obtain linear distortionless amplification.

In the absence of signal voltage ($e_g=0$), the grid voltage is constant at value E_{cc} Now a constant plate current I_b flows in the plate circuit. Due to this plate current a voltage drop $I_b Z_t$ is produced across the load impedance, *i.e.*, voltage drop depends upon the plate current which, in turn, is controlled by grid voltage. On the amplification of signal voltage, plate current varies in-accordance with the variations in signal voltage. The a.c. component of plate current flowing through load impedance develops the amplified output voltage.

9·3. VACUUM TUBE AMPLIFIER AS A LINEAR DEVICE :

In the description of a vacuum tube amplifier, it is assumed that the a c. component of the plate current is proportional to the input signal voltage. This shows that the amplified output voltage is proportional to the input signal. If the signal amplitude is doubled or tripled, the amplitude of output also gets doubled or tripled respectively. Thus we can say that the vacuum tube amplifier is a linear device. This is only possible provided the charateristic curves of the tube are linear. As there exists considerable non-linearity in the tube characteristics, hence such a condition is not satisfied by any tube. When the grid signal, a.c. plate current and plate voltage are small in amplitude with corresponding d c. quantities the characteristic curves may be treated almost linear. Then the vacuum tube amplifier may be treated as a linear device.

9·4. ZERO SIGNAL OR QUIESCENT OPERATION OF VACUUM TUBE AMPLIFIER :

When no input signal is applied, the operation is said to be quiescent. Figure (2) shows the circuit arrangement of a vacuum triode under such conditio·. In this case, there is only grid bias voltage E_{cc} in the grid circuit and plate current and plate voltage have d c. components. Because no signal is applied $e_c=E_c=E_{cc}$. The supply voltage E_{bb} and grid bias voltage E_{cc} are so adjusted that there is some plate current $i_b=I_b$ flowing in the plate circuit.

The value of the plate current can easily be determined by graphical method using the static characteristics of the tube

To analyse the operation under zero signal condition, we apply Kirchoff's voltage law to the plate circuit, and get

$$e_b + i_b R_L = E_{bb}$$

or
$$i_b = \frac{E_{bb} - e_b}{R_L} \quad \ldots(1)$$

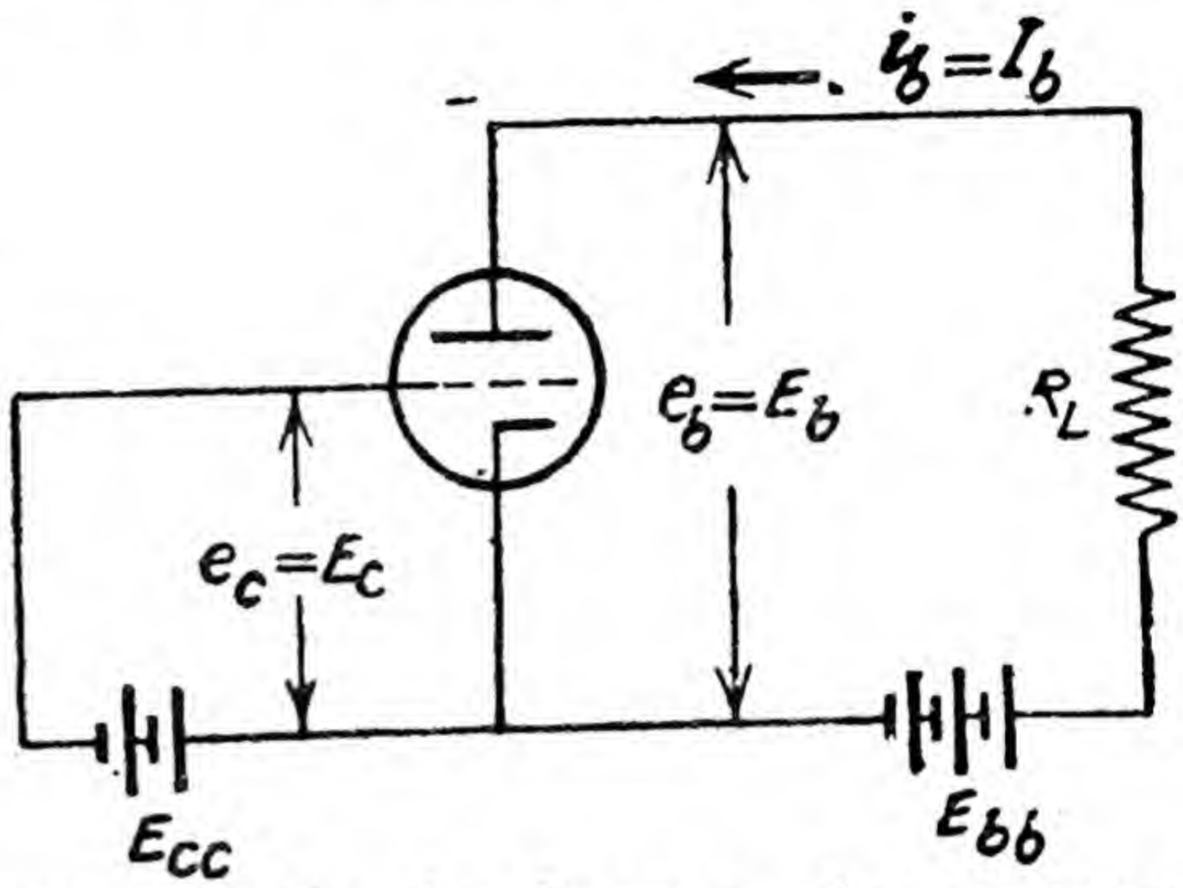


Fig. 2. Circuit diagram of a vacuum triode under zero signal operation.

In equation (1), there are two unknown quantities e_b and i_b. Therefore, this equation is insufficient to determine the plate current i_b and the plate voltage e_b. Another equation which relates these two quantities is required. Since the plate current is the function of the plate voltage and grid voltage, the plate characteristic of the tube may be represented by the following relation :

$$i_b = f(e_c, e_b) = f(E_{cc}, e_b). \quad \ldots(2)$$

From equations (1) and (2), we have

$$f(E_{cc}, e_b) = \frac{E_{bb} - e_b}{R_L} \quad \ldots(3)$$

It is now required to find the values of i_b and e_b that satisfy the equation (3). Since the nature of the variation of i_b with respect to e_b is not definitely specified by equation (2), the solution of simultaneous equations (I) and (2) is not possible. Therefore a graphical solution may be obtained by drawing the curves of these

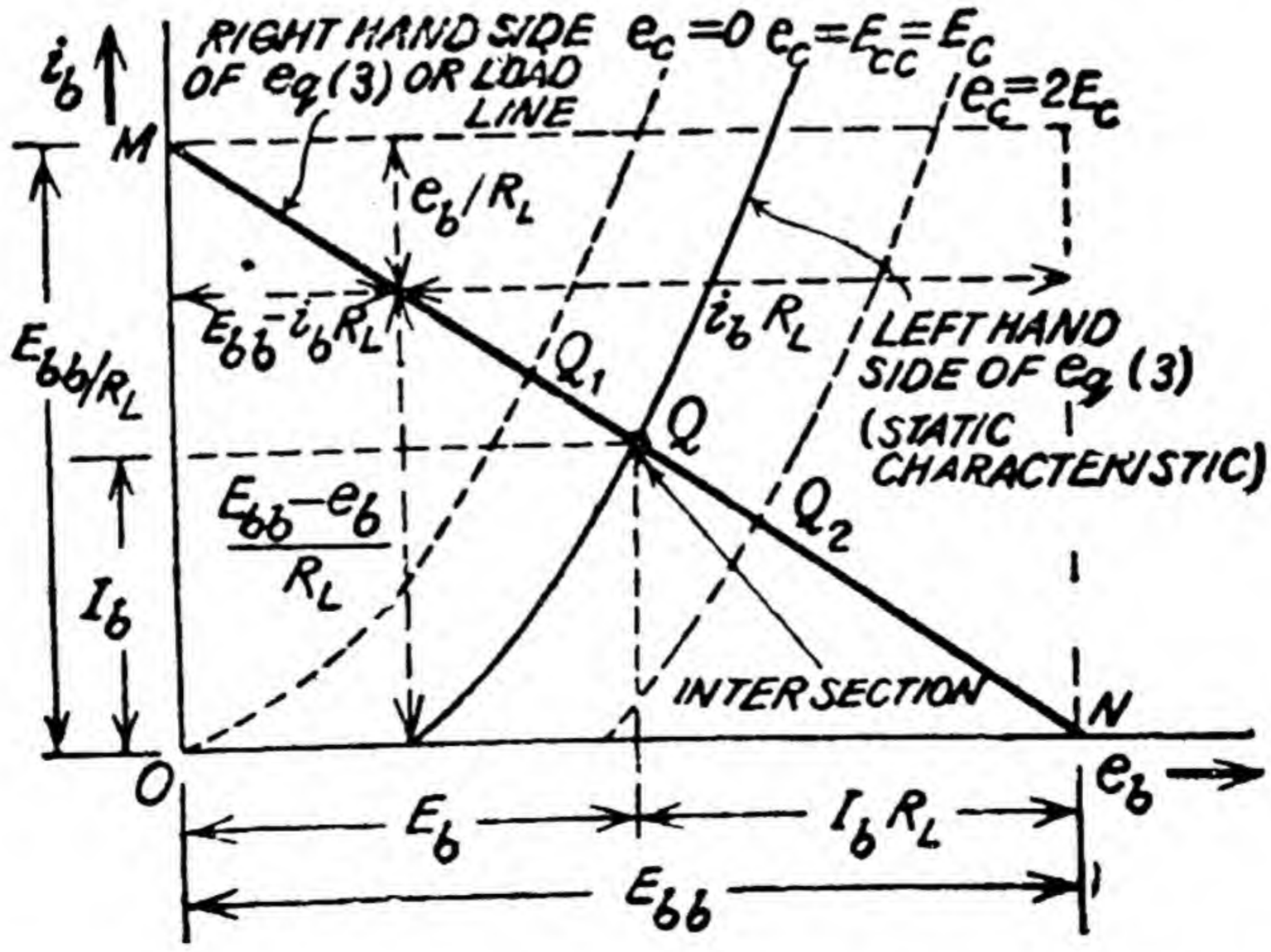


Fig. 3. Zero signal operation of a triode valve with resistive load.

equations and finding the point of intersection. This is shown in figure (3). The curve corresponding to equation (1) is a straight line which cuts the current and voltage axes at points M and N. The coordinates of these points may be obtained by putting $e_b=0$ and $i_b=0$ in equation (1). The coordinates are $(0, E_{ib}/R_L)$ and $(E_{bb}, 0)$ respectively. The straight line MN is known as load line and the reciprocal of the slope of this line equals the load resistance R_L. The point Q where the load line for R_L and the characteristic for $e_c=E_{cc}$ intersect is known as the *quiescent point* or the *operating point*. The ordinate of this point Q gives the zero signal plate current I_b for grid bias E_{cc}, plate supply voltage E_{bb} and load resistance R_L. The abscissa of the point Q gives the zero signal plate voltage E_b. The zero signal plate voltage E_b differs from the plate supply voltage E_{bb} because of voltage drop in the load.

If the grid supply voltage E_{cc} is changed to a new value (say $E_{cc}=0$), then the operating point moves along the resistance line to the new position at the intersection of the load line and the corresponding new plate characteristic as is shown by Q_1. If the load resistance R_L is changed to some value R_L', the load line will still pass through the point N but gets tilted up or down as to intersect the current axis at point E_{bb}/R_L'.

9·5. WORKING OF VACUUM TUBE AMPLIFIER WITH SMALL a.c. GRID SIGNAL :

The basic circuit of a vacuum tube rectifier with small a.c. grid signal is shown in figure (1). This circuit is different from the circuit of figure (2) as a voltage to be amplified, *e.g.* known as grid signal voltage, is applied at the input in series with E_{cc}. The voltages E_{bb}, and E_{cc} and the load resistance $Z_L=R_L$ are so chosen that the operating point lies over the linear region of the characteristics. Amplified output voltage appears across the load resistance. The objective of this article is to study the relation between the amplified voltage and grid signal voltage by graphical means.

The instantaneous grid signal voltage e_g may be expressed as

$$e_g=E_{gm}\sin\omega t. \qquad ...(1)$$

where ω is the angular frequency. $\omega=2\pi f$, where f is signal frequency in hertz E_{gm} is the amplitude of grid signal.

The total grid to cathode voltage e_c consists of two parts : (i) constant grid bias voltage $E_{cc}=E_c$, and (ii) a.c. signal voltage of instantaneous value e_g, *i.e.*

$$e_c=E_c+e_g. \qquad ...(2)$$

The grid signal voltage is thus an increment in e_c. It causes corresponding increments in total plate current i_b and plate voltage e_b denoted by i_p and e_p respectively. Thus the total instantaneous plate current i_b equals the no signal plate current I_b plus the incremental or a.c. plate current i_p, *i.e.*,

$$i_b=I_b+i_p. \qquad ...(3)$$

Similarly, total instantaneous plate voltage e_b is the sum of two parts : (*i*) zero signal plate voltage E_b, and (*ii*) incremental or a.c. plate voltage e_p *i.e.*

$$e_b = E_b + e_p. \qquad \text{.. (4)}$$

The graphical construction of the variation of plate current and plate voltage due to the application of grid signal is shown in

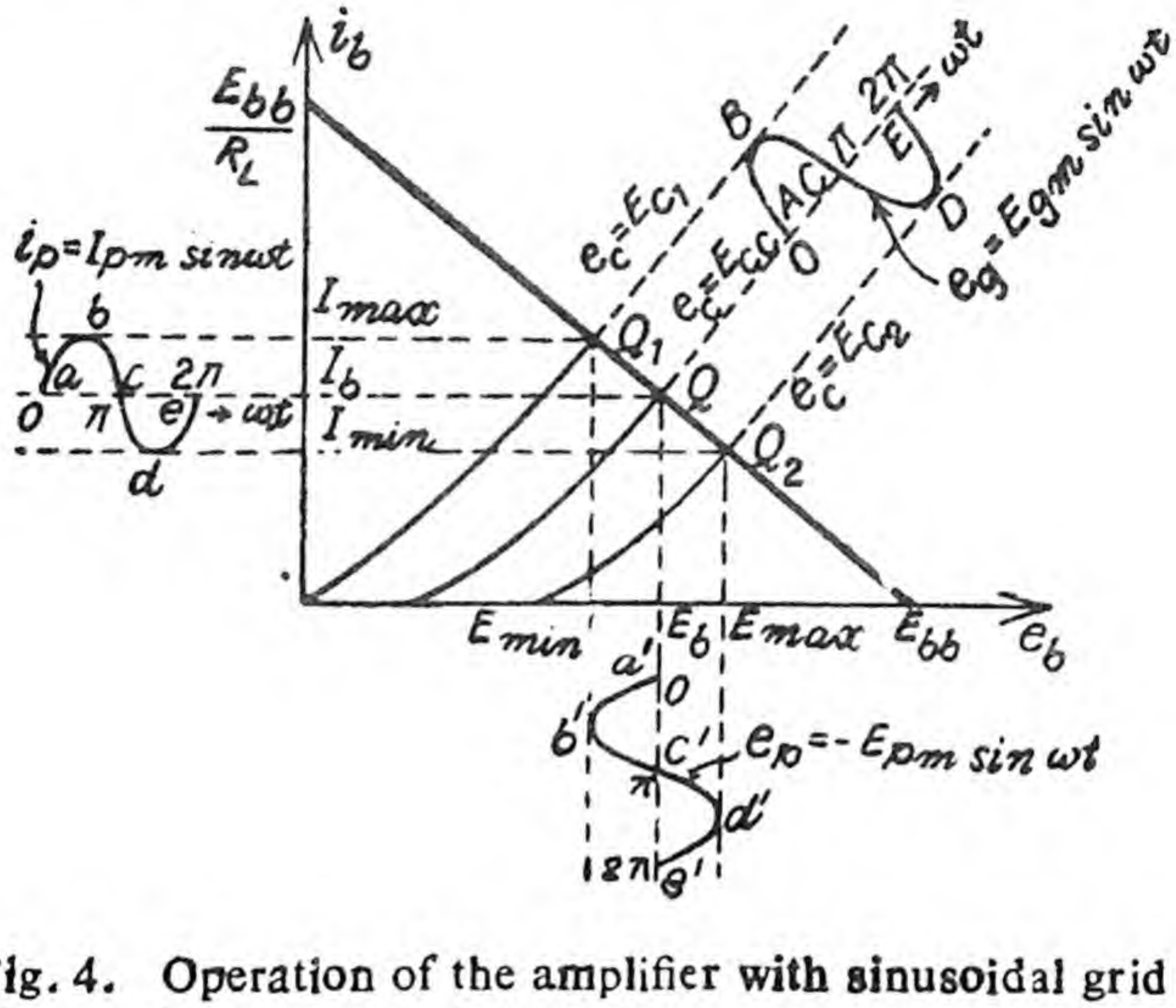


Fig. 4. Operation of the amplifier with sinusoidal grid signal.

figure (4). Hence it is assumed that the plate characteristics for different grid voltages are equally spaced for equal increments in the region cut by the load line. Under zero grid signal, Q is the point of operation and steady state grid bias $E_c = E_{cc}$. The points *a*, *b*, *c*, *d* and *e* of the output current and points *a'*, *b'*, *c'*, *d'* and *e'* of the output voltage waveform correspond respectively to the points *A*, *B*, *C*, *D* and *E* of the input grid signal waveform. At the positive peak value of the grid signal, the grid voltage

$$e_c = E_{c1} = E_{cc} + E_{gm}$$

and the operating point shifts to Q_1. The total plate current now becomes large while at the same time the plate voltage becomes smaller. At the negative peak value of the grid signal, the grid

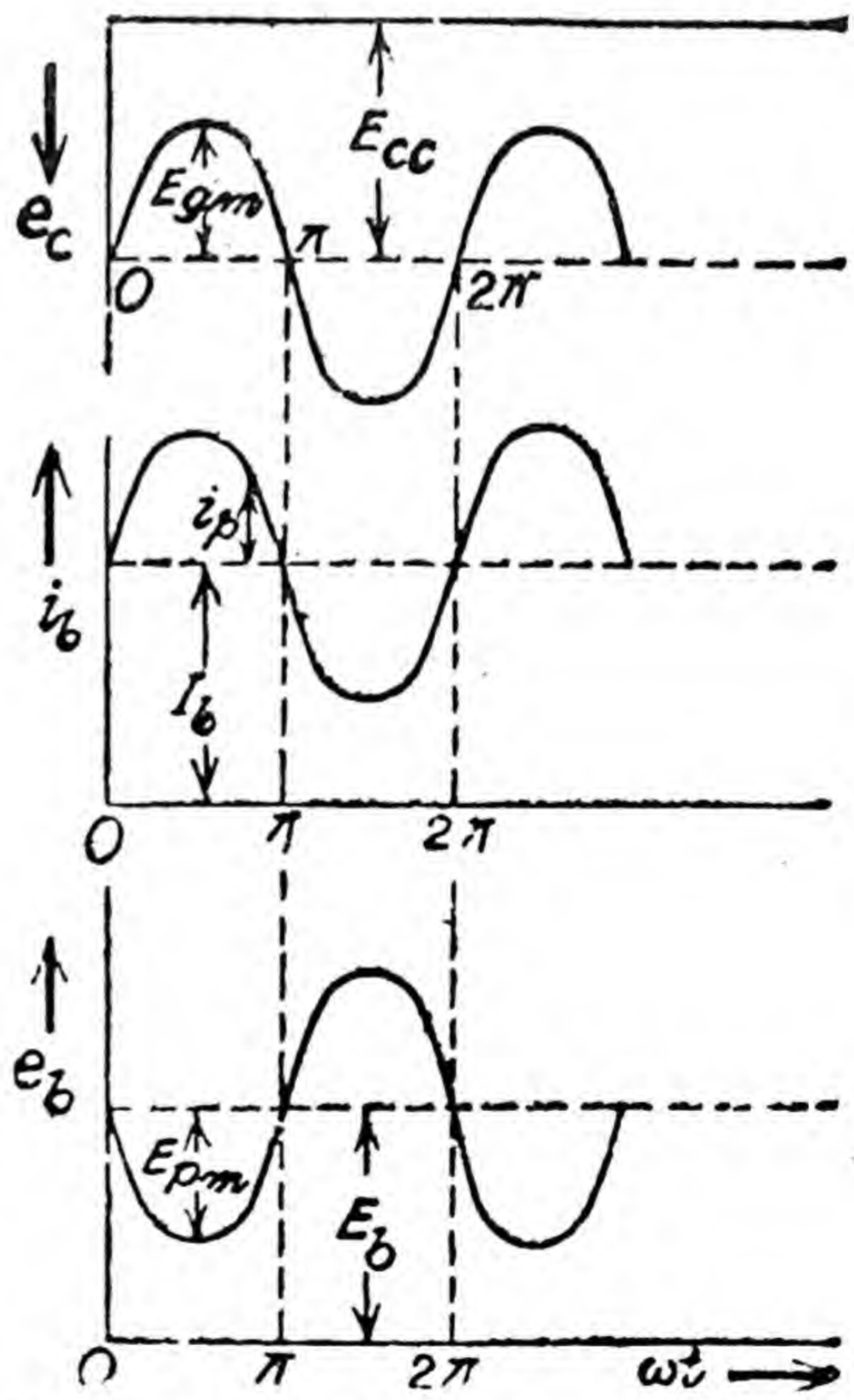


Fig. 5. Waveform of a.c. grid signal, a.c. plate current and a.c. plate voltage.

voltage $e_o = E_{o2} = E_{cc} - E_{gm}$ and the operating point shifts to Q_2. The total plate current now becomes smaller while at the same time the plate voltage becomes larger. It is clear from the figure *that plate current cycle is in phase with the grid voltage cycle where as the plate voltage cycle is* 180 *degrees out of phase.* Hence

$$i_p = I_{pm} \sin \omega t \quad \text{.. (5)}$$

and $$e_p = E_{pm} \sin (\omega t + 180) = -E_{pm} \sin \omega t \quad \text{...(6)}$$

This shows that whenever the a.c. grid voltage increases, the plate current increases but the plate voltage decreases. Figure (5) shows the grid input waveshape and corresponding plate current and plate voltage waveshapes.

The a.c. current and voltages enable us to determine the amplification property of the amplifier.

9·6. DYNAMIC CHARACTERISTIC :

In figure (4), we have seen that the load line cuts the plate characteristics in point corresponding to different values of e_g. These points are marked Q_1, Q and Q_2. The coordinates of these points give us the value of plate current i_b corresponding to different values of the grid voltage with a load resistance R_L. If we plot these values of the plate current against the corresponding values of grid voltage, the curve so obtained is called a dynamic characteristic of the tube circuit as shown in figure (6). The difference between static characteristic and dynamic characteristic is that static characteristic concerns with the tube alone while dynamic characteristic concerns with tube as well as associated circuit. For small values of plate current, the static characteristics show marked non-linearity while dynamic characteristic may be linear if along the load line the plate characteristics are equidistant for equal increment of grid voltage. The dynamic characteristic enables us to obtain directly the output current for a given in put signal.

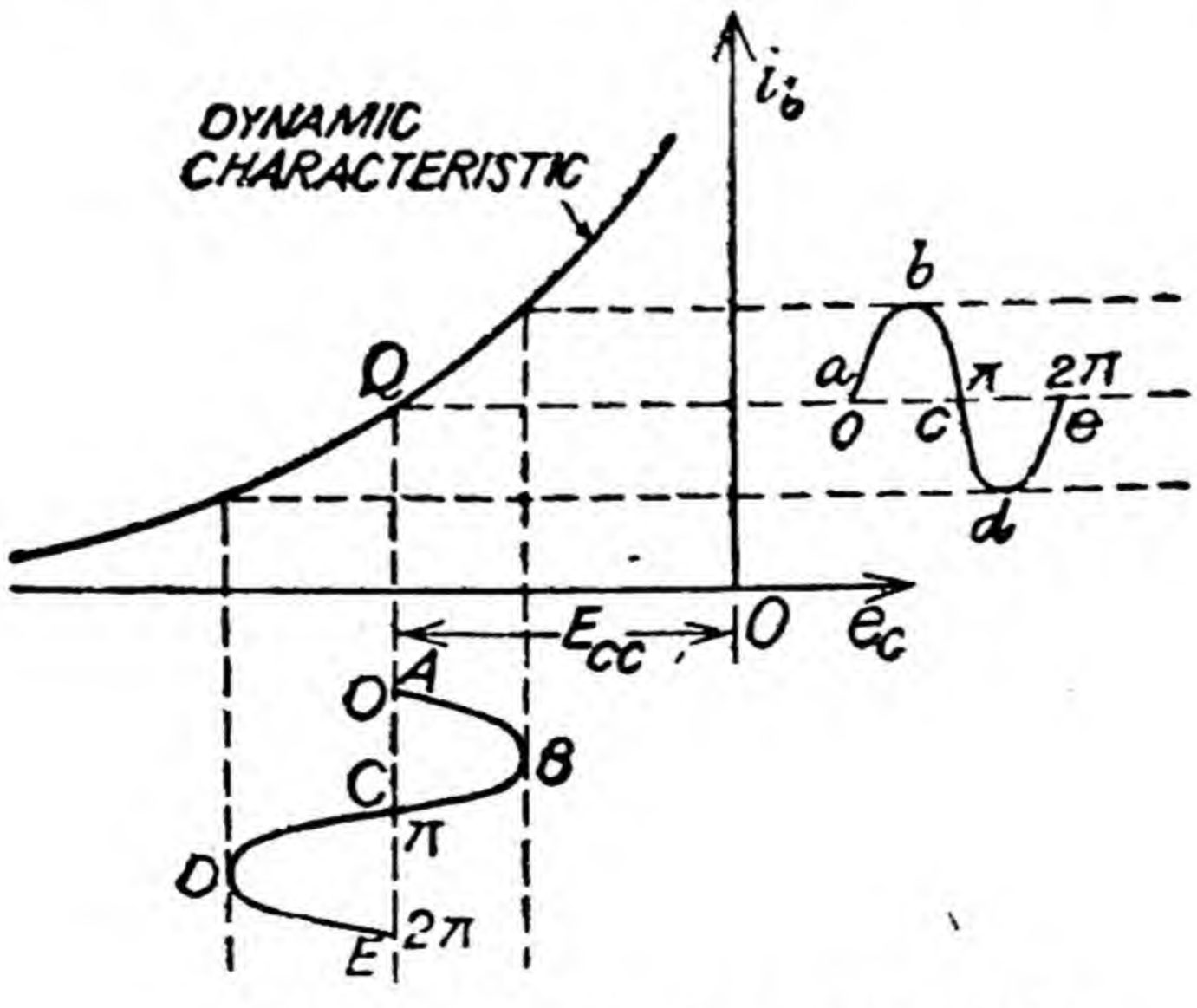


Fig. 6. Dynamic characteristics.

9·7. EQUIVALENT CIRCUIT REPRESENTATION FOR TRIODE AMPLIFIER :

In many electron tube problems, we generally require values of i_p (a.c. plate current), e_p (a.c. plate voltage) for a given e_g (the instantaneous grid voltage). From the static characteristic of the tube, such an information can be obtained with sufficient accuracy, if the potential variations are not small. But when the potential variations are small, the information, thus obtained, is not accurate. For small signal variations (as in most amplifier performance) the tube graphical characteristics can be taken as linear and in such linear regions one can assume the values of tube parameters μ, r_p and g_m to be almost constant.

Voltage source representation : Since the anode current is a function of e_c and e_b the instantaneous grid and plate potentials, we can write

$$i_b = f(e_c, e_b)$$

or
$$di_b = \frac{\partial i_b}{\partial e_c}.de_c + \frac{\partial i_b}{\partial e_b}.de_b.$$

For practical cases, change in i_b should be finite and let it be represented by Δ. Therefore,

$$\Delta i_b = \frac{\partial i_b}{\partial e_c}.\Delta e_c + \frac{\partial i_b}{\partial e_b}.\Delta e_b.$$

For the constancy of μ, r_p and g_m, it is essential that Δ changes be confined to linear region of the graphical tube characteristics and then we can put

$$\frac{\partial i_b}{\partial e_c} = g_m \text{ and } \frac{\partial i_b}{\partial e_b} = \frac{1}{r_p}$$

$$\therefore \quad \Delta i_b = g_m.\Delta e_c + \frac{1}{r_p}.\Delta e_b.$$

Now the components of e_b, e_c and i_b which vary are e_0 (a.c. out put) e_g and i_p. So

$$\Delta i_b = i_p, \qquad \Delta e_c = e_g \qquad \text{and } \Delta e_b = e_0$$

$$\therefore \quad i_p = g_m + \frac{1}{r_p}.e_0$$

or
$$-\mu e_g + i_p r_p = e_0$$

This equation suggests that the triode can be replaced by an equivalent voltage generator μe_g with positive terminal to the cathode plus as internal resistance r_p. The equivalent circuit is shown in fig. (7).

When the grid voltage is sinusoidal, the effective or r. m. s. value of the voltage and current may be used in place of the instantaneous values of the same. Using r.m.s. values

$$-\mu \mathbf{E}_g + \mathbf{I}_p r_p = \mathbf{E}_0$$

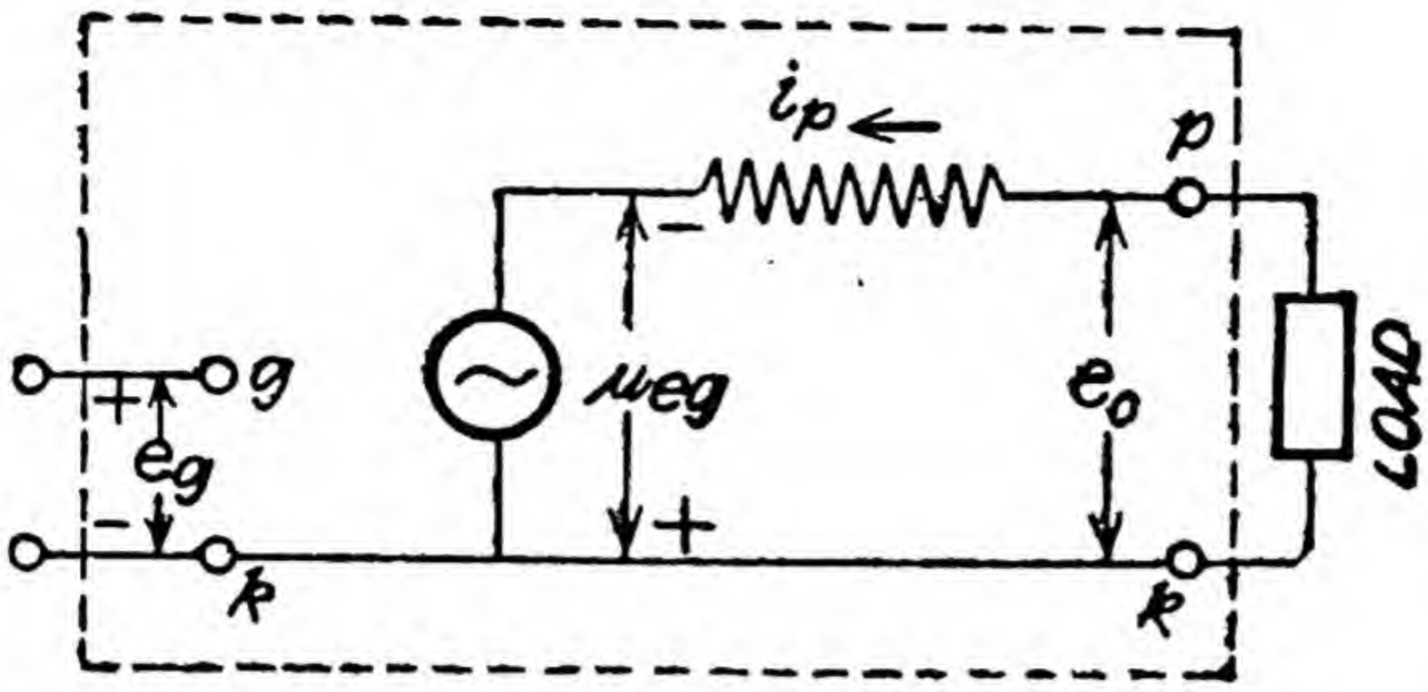


Fig. 7. Amplifier tube as a constant voltage generator.

The same equation can be obtained if a mesh [fig. 8] be constructed for current I_p. It, therefore, suggests that triode can be replaced by an equivalent voltage generator μE_g with positive terminal to cathode plus an internal resistance r_p [fig. 9 (*a*)]. The same representation is depicted in fig. 9 (*b* .

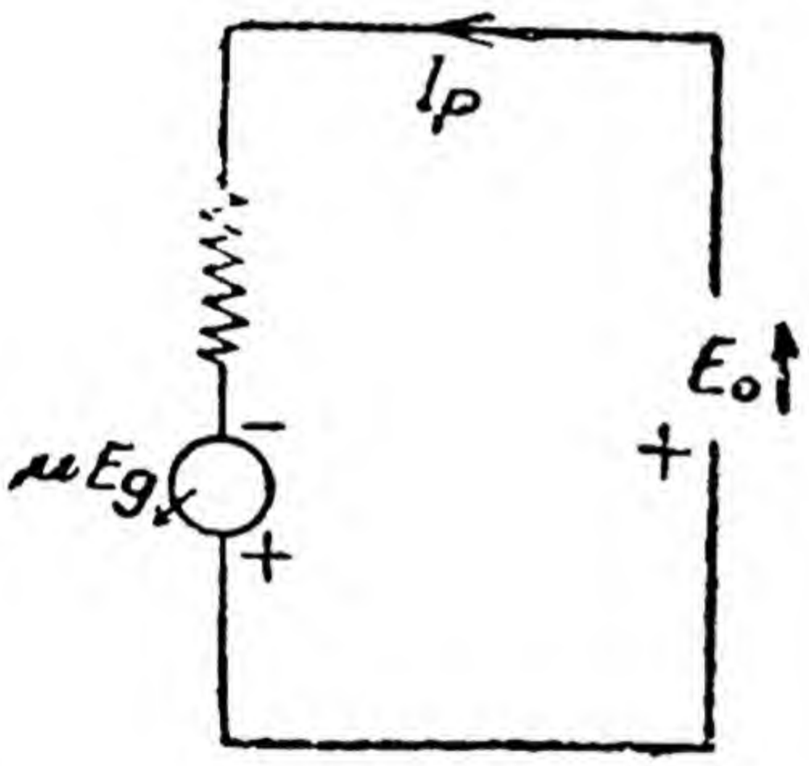


Fig. 8. Mesh Construction.

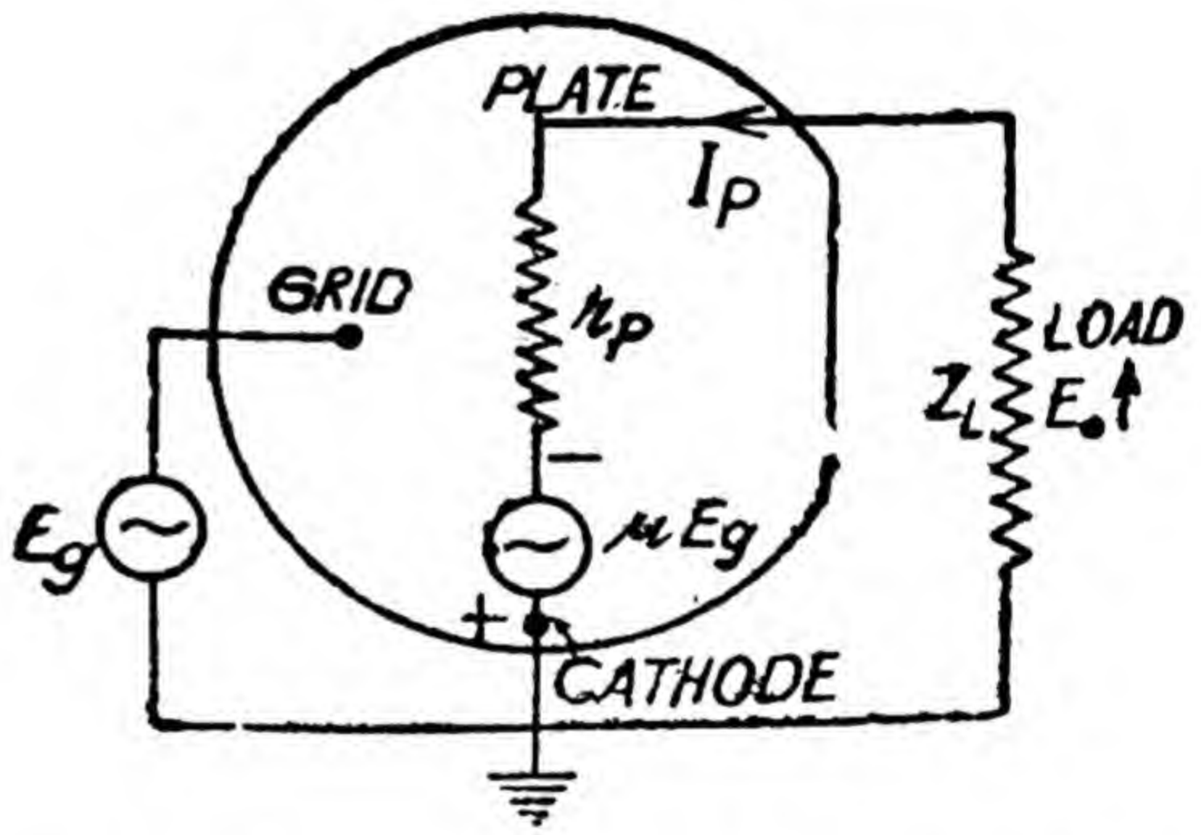


Fig. 9. (*a*) Equivalent circuit representation.

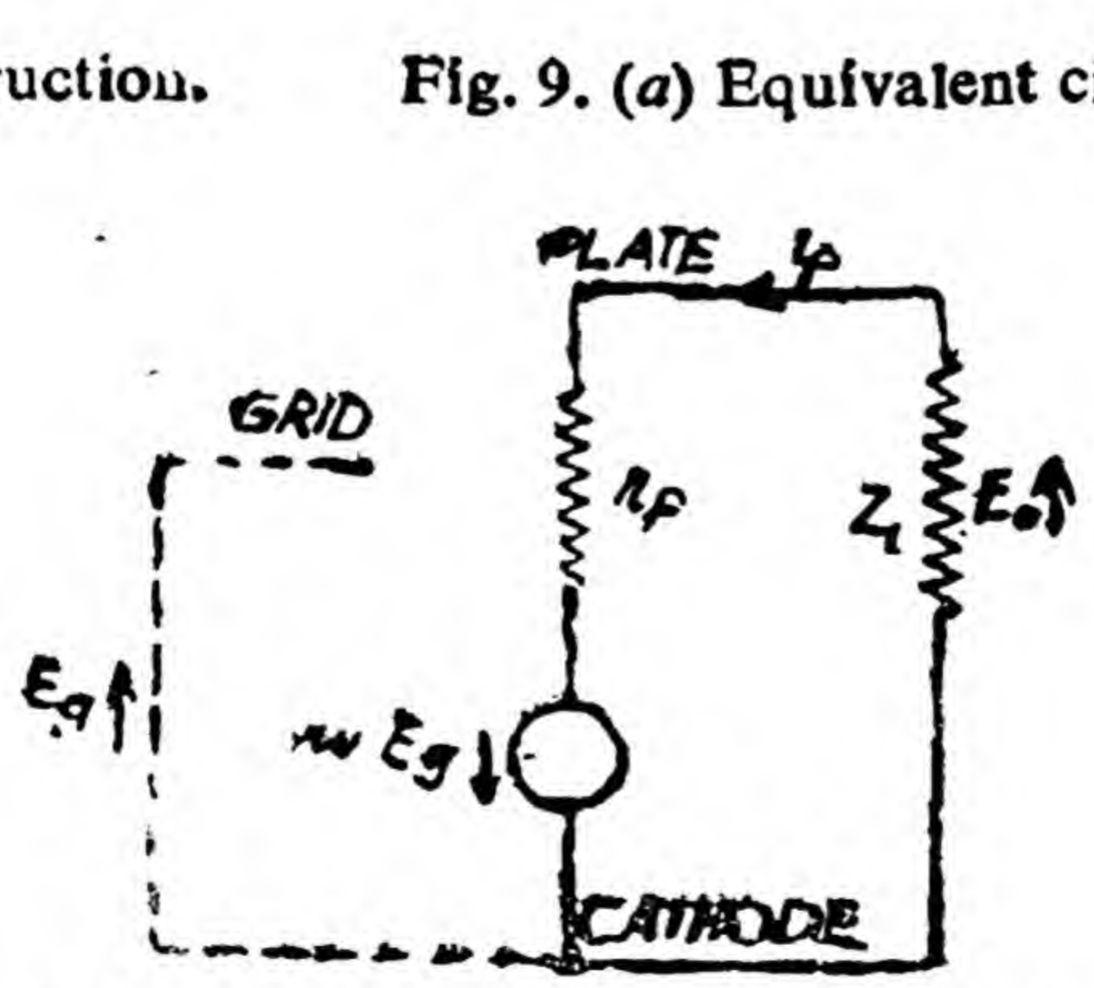


Fig. 9. (*b*) Equivalent circuit representation.

An alternative approach to this equivalent circuit is to replace the triode by a generator developing a voltage $-\mu E_g$ acting inside the tube from cathode toward the plate and internal resistance r_p shown in fig. 9 (*c*) and fig 9 (*d*).

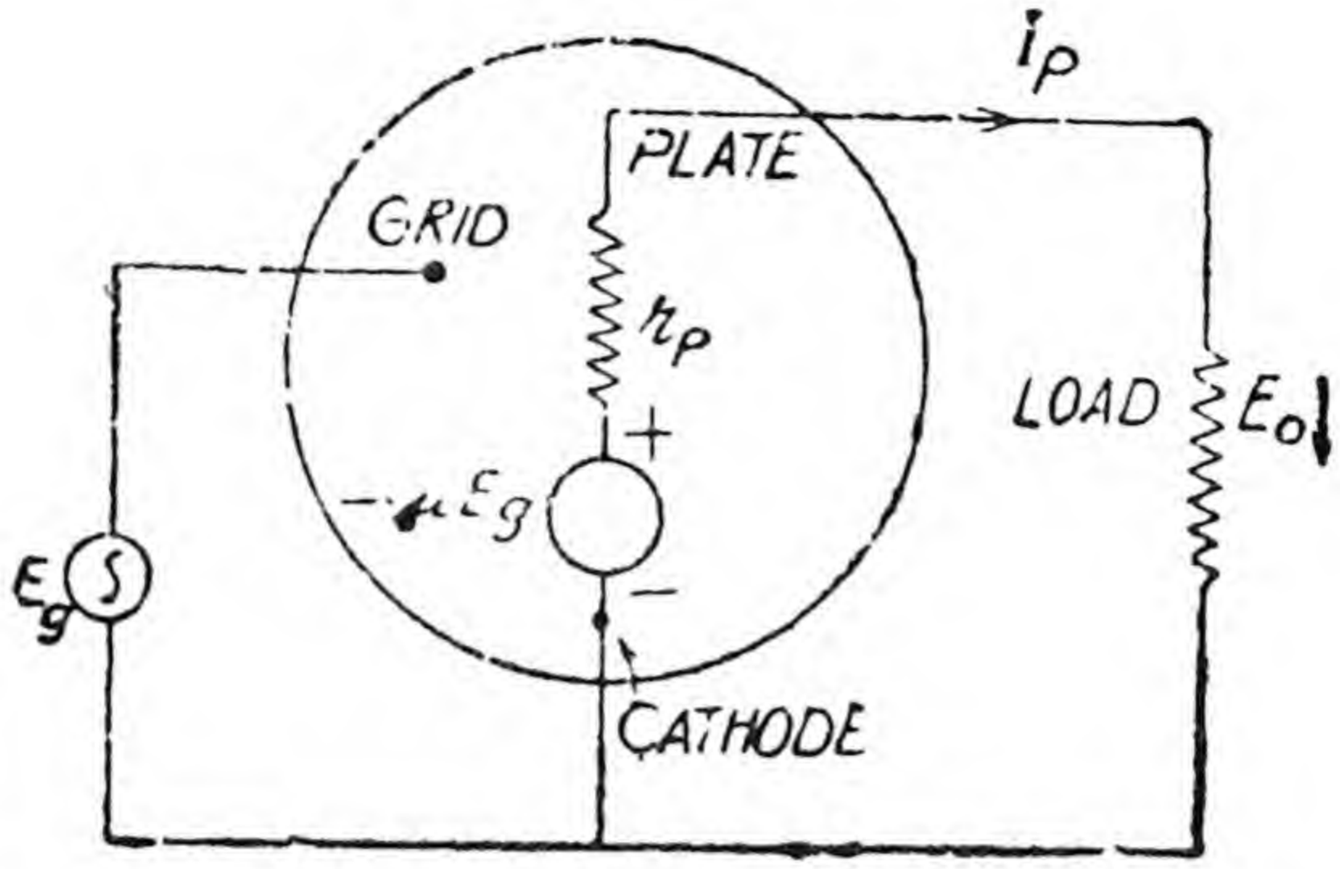


Fig. 9. (*c*) Equivalent circuit representation.

The latter representation very well accounts for the phase change of 180 degrees between the grid and anode potentials.

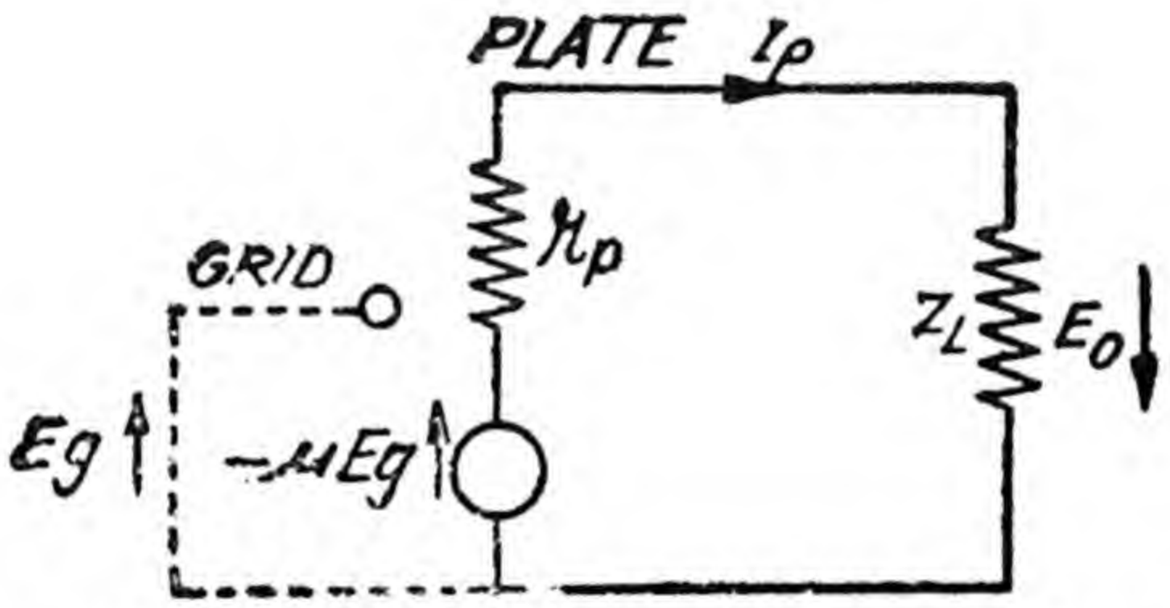


Fig. 9. (*d*) ·Equivalent circuit representation.

Thus for a region of graphical tube characteristic in which μ, r_p and g_m are constant, vacuum triode can be replaced with an equivalent voltage source circuit.

Current source representation : The equation

$$-\mu E_g + I_p r_p = E_0$$

can be put in the form

$$\frac{\mu}{r_p} . E_g = I_p - \frac{E_0}{r_p},$$

$$g_m E_g = I_p - \frac{E_0}{r_p}.$$

The equation, obviously, corresponds to a circuit shown in fig. (11) in which a current source $g_m E_g$ supplies two branches with current value of I_p and $-E_0/r_p$.

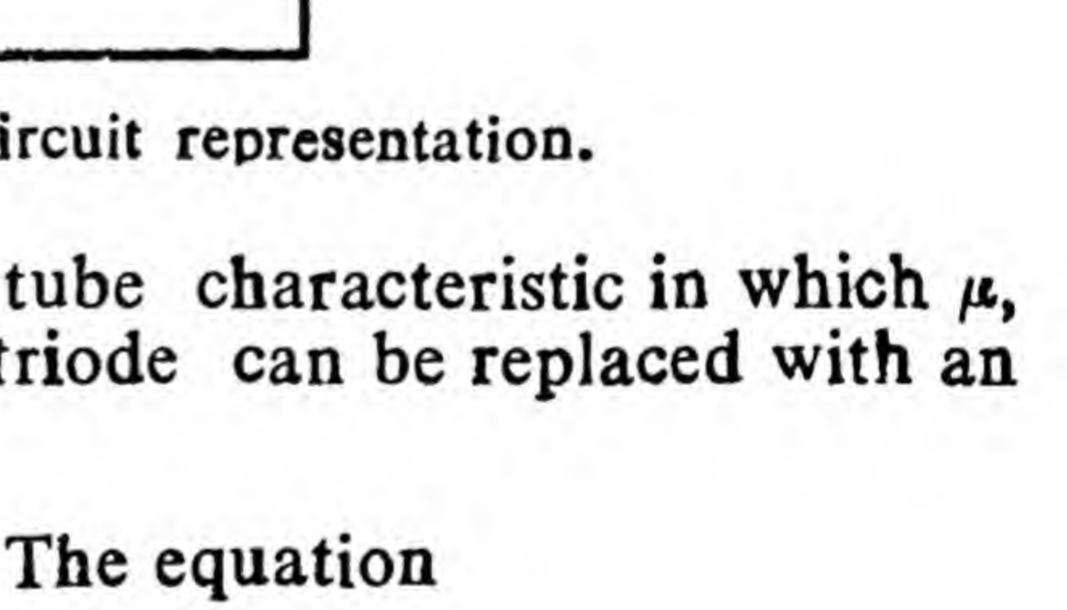


Fig. 10. Triode amplifier.

The tube is, thus, taken to be equivalent to a current source of value $g_m E_g$, in parallel with a resistor r_p. Direction of current

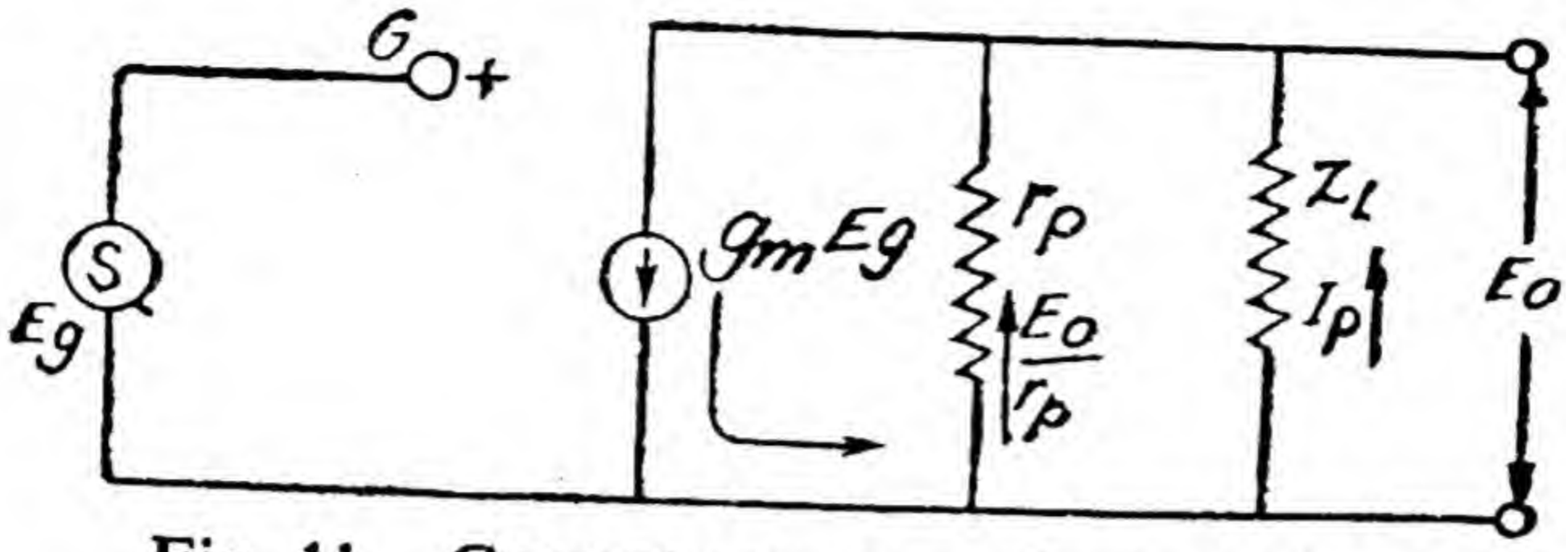


Fig. 11. Current source equivalent circuit.

flow is same as the direction of original voltage generator. This type of circuit equivalence is best suited for pentode tubes.

Rules to draw the equivalent circuit :

In order to draw the linear a.c. equivalent circuit of a vacuum tube circuit, the following simple rules s hould be followed :

(i) The actual circuit diagram of the vacuum tube circuit should be drawn neatly indicating the point *G, P* and *K* corresponding to grid, plate and cathode.

(ii) At the start of drawing the a.c. equivalent circuit, *G, P* and *K* should be located with relative positions of original diagram.

(iii) A voltage generator with potential μE_g and resistance r_p should be inserted between *P* and *K*. Here E_g denotes the r.m.s. value of the a.c. grid cathode voltage of the tube.

(iv) All the circuit elements except tube and d.c. sources should be transferred from the actual circuit to the equivalent circuit. It should be remembered that the relative positions of the transferred elements should not be changed.

(v) All d.c. sources should be replaced by their internal resistances, if any.

9·8. VOLTAGE GAIN OF A SINGLE STAGE UNTUNED LINEAR A_1 AMPLIFIER :

Figs. 12 (*a*) and (*b*) show the basic circuit and its equivalent a.c. circuit of single stage untuned linear A_1 amplifier using load impedance Z_L which, in general, is complex.

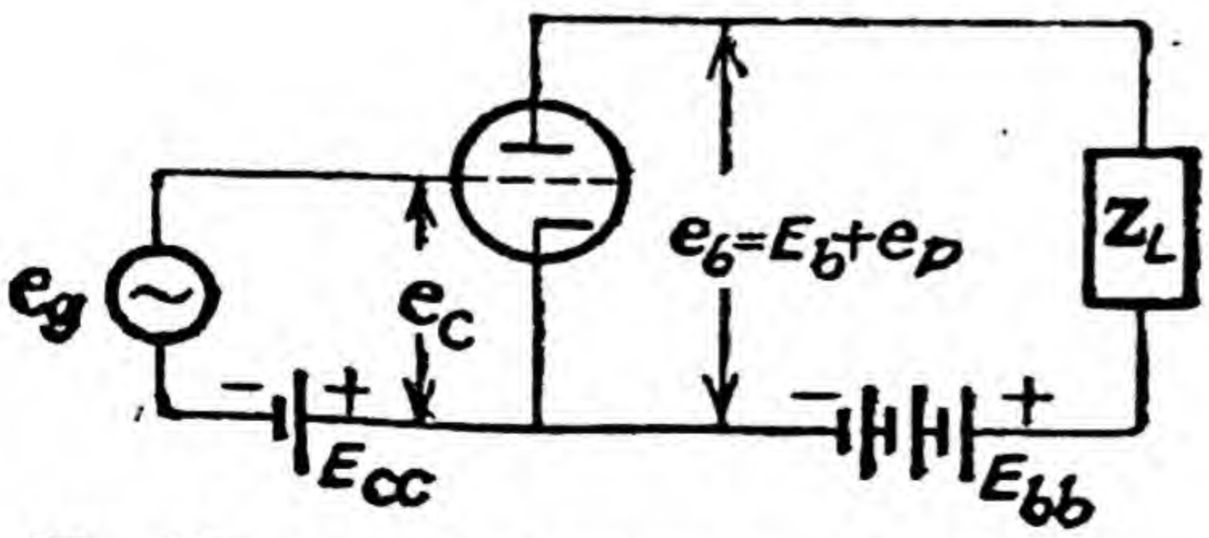


Fig. 12. (*a*) Basic circuit of a single stage untuned class A_1 amplifier.

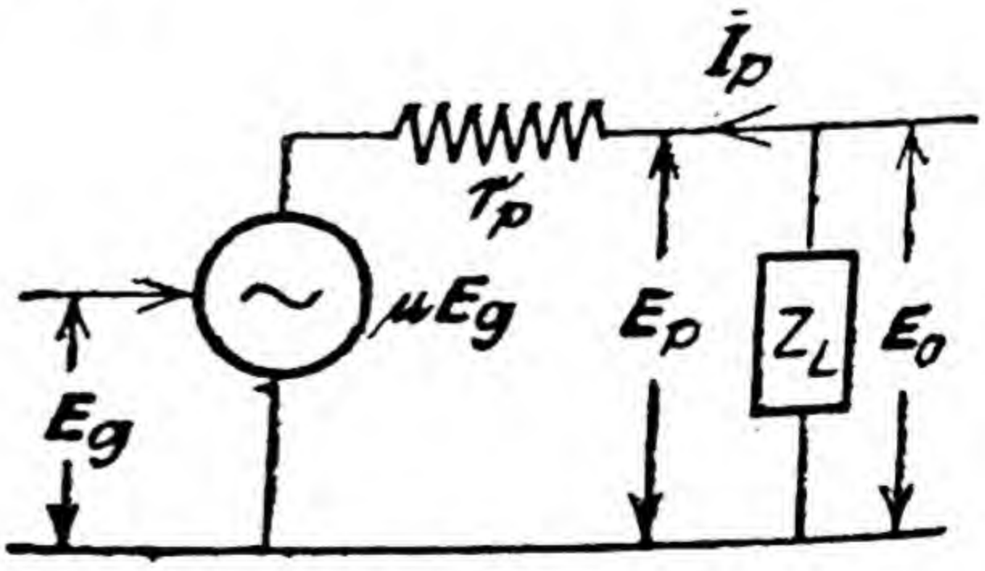


Fig. 12. (*b*) A.C. equivalent circuit.

If the grid signal voltage e_g is a small sinusoidal alternating voltage, then it can be expressed as

$$e_g = \sqrt{2}\, E_g . \sin \omega t, \qquad \ldots(1)$$

where e_g is the instantaneous value and E_g is the r.m.s. or effective value of the sinusoidal signal voltage of angular frequency ω. If the load impedance is linear and passive, it can be given as

$$Z_L = R_L + jX_L = Z_L \angle \theta. \qquad \ldots(2)$$

When the varying component of grid voltage is sinusoidal, the variations e_p and i_p are also sinusoidal and of the same frequency as e_g. In this case, instead of using the instantaneous values such as e_g, e_p, e_L. i_p etc. we may use the corresponding effective or r.m.s. complex values $\mathbf{E}_g$, $\mathbf{E}_p$, $\mathbf{E}_L$, $\mathbf{I}_p$ etc. The equivalent circuit for complex quantities is shown in fig. 12 (*b*). Using this circuit, we can calculate the voltage gain of the amplifier.

The effective value of plate current is given by

$$\mathbf{I}_p = \frac{\mu \mathbf{E}_g}{r_p + Z_L}, \qquad \ldots(3)$$

where $(r_p + Z_L)$ is the total impedance in the plate circuit.

This current flowing through the impedence load $\mathbf{Z}_L$ produces the output voltage $\mathbf{E}_0$, given by

$$\mathbf{E}_0 = -\mathbf{I}_p \mathbf{Z}_L. \qquad \ldots(4)$$

Substituting the value of $\mathbf{I}_p$ from equation (3) in equation (4), we have

$$\mathbf{E}_0 = -\frac{\mu \mathbf{E}_g}{r_p + \mathbf{Z}_L} . \mathbf{Z}_L = -\mu \mathbf{E}_g . \frac{\mathbf{Z}_L}{r_p + \mathbf{Z}_L}. \qquad \ldots(5)$$

The complex voltage gain of the amplifier is defined as the ratio of complex output voltage to the complex input voltage, *i.e.*,

$$\mathbf{A} = \frac{\mathbf{E}_0}{\mathbf{E}_g} = -\mu \frac{\mathbf{Z}_L}{r_p + \mathbf{Z}_L}. \qquad \ldots(6)$$

The complex voltage gain may be put as

$$\mathbf{A} = A \angle \alpha. \qquad \ldots(7)$$

where A is the magnitude of the complex voltage gain $\mathbf{A}$ and α is the phase angle of the complex voltage gain. From equation (5), it is clear that α has two components: (*i*) 180° phase difference due too negative sign and (*ii*) additional phase due to the complex nature of load impedance $\mathbf{Z}_L$.

Now we shall consider three types of load impedance $\mathbf{Z}_L$.

Case (1). Z is a complex inductive impedance :

Figure (13) shows a.c. equivalent circuit in this case.

Considering $\mathbf{E}_g$ as reference vector and positive angles measured counterclockwise, the following phase relationship exist :

$$\text{phase angle of load, } \theta = \tan^{-1} \frac{X_L}{R_L} = \tan^{-1} \frac{\omega L}{R_L}, \qquad \ldots(8)$$

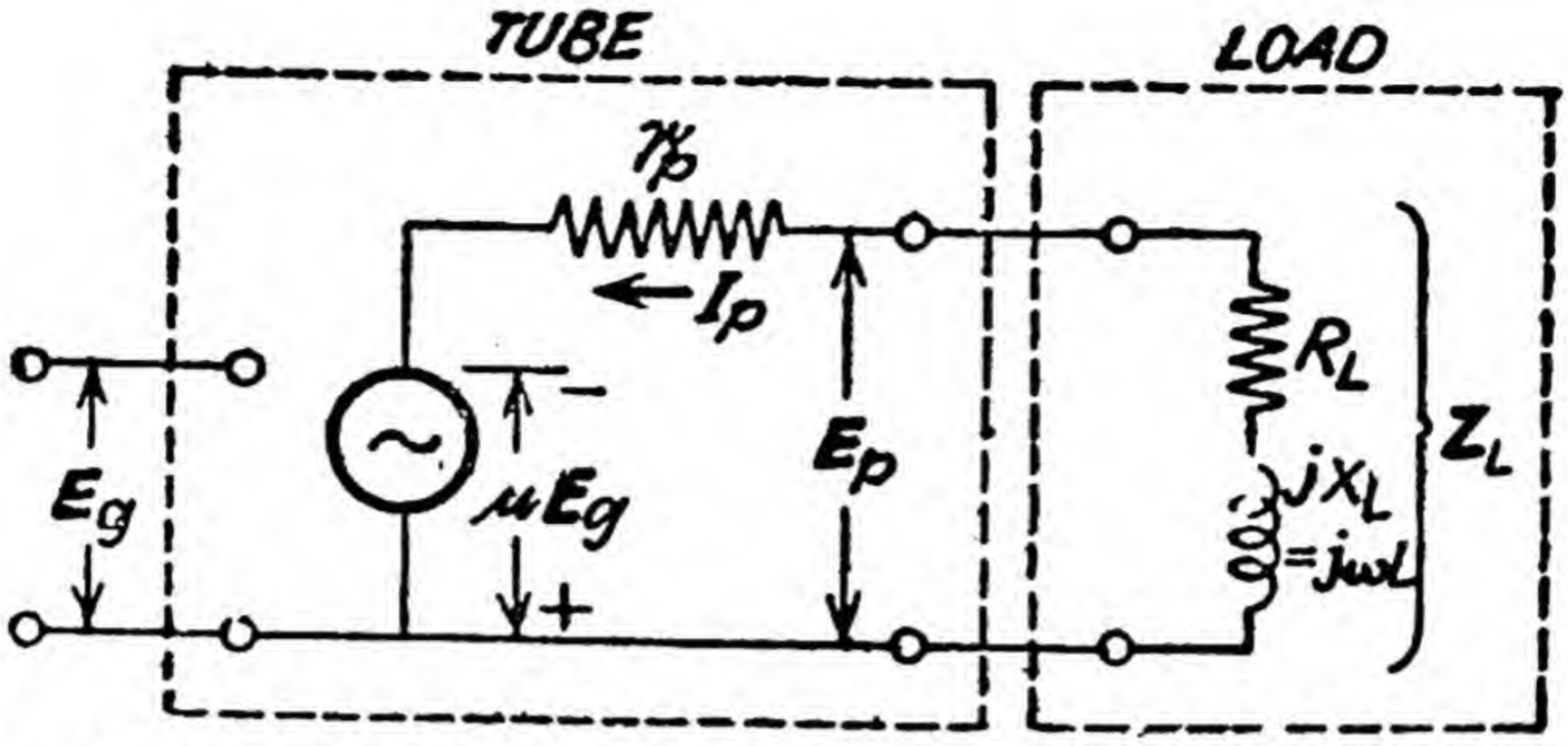


Fig. 13. A.C. equivalent circuit of a single stage amplifier using complex inductive load.

phase angle of plate current, $\phi = -\tan^{-1} \dfrac{\omega L}{r_p + R_L}$. ...(9)

$\therefore$ Phase angle α by which load voltage $\mathbf{I}_p\mathbf{Z}_L$ leads the grid voltage

$$= \theta + \phi$$

$$= \tan^{-1} \frac{\omega L}{R_L} - \tan^{-1} \frac{\omega L}{r_p + R_L}. \quad ...(10)$$

Now phase angle α by which plate voltage leads the grid voltage is given by

$$\alpha = 180° + \theta + \phi$$

$$= 180° + \tan^{-1} \frac{\omega L}{R_L} - \tan^{-1} \frac{\omega L}{r_p + R_L}. \quad ...(11)$$

Considering equations (8), (9), (10) and (11), we can draw the phasor diagram for a single amplifier using inductive load. This diagram is shown in figure (14).

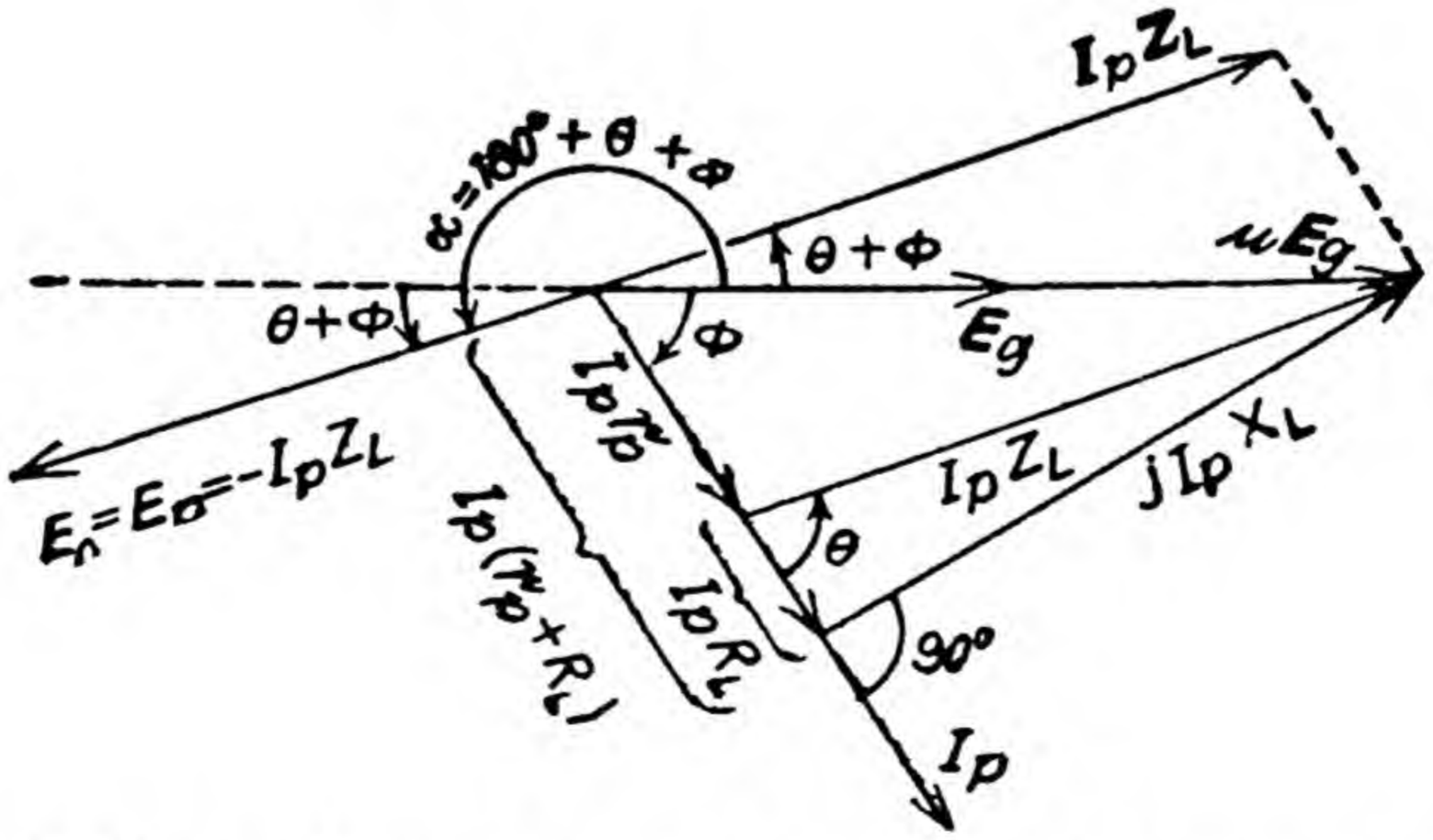


Fig. 14. Vector or phasor diagram for sinusoidal alternating components in single stage amplifier using inductive load.

We observe the following points :

(*i*) The a.c. plate current $\mathbf{I}_p$ lags behind the dividing voltage $\mu\mathbf{E}_g$ and the phase angle is ϕ.

(*ii*) The voltage drop $\mathbf{I}_p\ (r_p+R_L)$ across the series resistance (r_p+R_L) is in phase with $\mathbf{I}_p$. The voltage drop $(j\mathbf{I}_pX_L)$ across inductance L leads $\mathbf{I}_p$ by 90°. The sum of the two components, *i.e.*, $\mathbf{I}_p\ (r_p+R_L)$ and $j\mathbf{I}_pX_L$ is equal to the total driving voltage $\mu\mathbf{E}_g$.

(*iii*) The voltage drop $\mathbf{I}_pR_L$ and $j\mathbf{I}_pX_L$ added vectorially constitute the voltage drop $\mathbf{I}_p\mathbf{Z}_L$ which leads the current $\mathbf{I}_p$ by an angle θ. The phase angle of this voltage with respect to $\mathbf{E}_g$ is $(\theta+\phi)$.

(*iv*) The output voltage is opposite to $\mathbf{I}_p\mathbf{Z}_L$ and its phase with respect to $\mathbf{E}_g$ is α.

Now we consider the voltage gain for this case.

As $\mathbf{Z}_L=R_L+jX_L=R_L+j\omega L$, the complex a.c. plate current is given by

$$\mathbf{I}_p=\frac{\mu\mathbf{E}_g}{r_p+R_L+j\omega L}. \qquad \ldots(12)$$

The magnitude of the current is given by

$$I_p=\frac{\mu E_g}{\sqrt{\{(r_p+R_L)^2+\omega^2L^2\}}}. \qquad \ldots(13)$$

Now output voltage $\mathbf{E}_p$ is given by

$$\mathbf{E}_p=\mathbf{E}_0=-\mathbf{I}_p\,\mathbf{Z}_L$$

$$=-\mu\mathbf{E}_g\,\frac{R_L+j\omega L}{(r_p+R_L)+j\omega L}. \qquad \ldots(14)$$

The magnitude of the voltage is given by

$$E_0=\mu E_g\,\frac{\sqrt{(R_L^2+\omega^2L^2)}}{\sqrt{\{(r_p+R_L)^2+\omega^2L^2\}}}. \qquad \ldots(15)$$

$$\therefore \text{ Complex voltage gain } \mathbf{A}=\frac{\mathbf{E}_0}{\mathbf{E}_g}=-\mu\,\frac{R_L+j\omega L}{(r_p+R_L)+j\omega L}. \qquad \ldots(16)$$

The magnitude of complex voltage gain is given by

$$A=\mu\,\frac{\sqrt{(R_L^2+\omega^2L^2)}}{\sqrt{\{(r_p+R_L)^2+\omega^2L^2\}}}. \qquad \ldots(17)$$

Case (II), $\mathbf{R}_L$ is pure resistance :

In this case $\mathbf{Z}_L=R_L+j0$,

and $$\mathbf{A}=-\mu\,\frac{R_L}{r_p+R_L}=-\frac{\mu}{1+r_p/R_L}. \qquad \ldots(18)$$

The phase angle of output voltage *i.e.*, of the voltage gain is 180°. This shows that if all the reactive components are excluded from the amplifier circuit, the phase shift produced by one stage is always 180° irrespective of frequency.

Case (III). Z_L is pure inductance :

In this case $\mathbf{Z}_L=0+j\omega L$.

The complex a.c. component of plane current $\mathbf{I}_p$ is given by

$$\mathbf{I}_p=\frac{\mu\mathbf{E}_g}{r_p+j\omega L}. \qquad (19)$$

The magnitude of this current is given by

$$I_p=\frac{\mu E_g}{\sqrt{(r_p^2+\omega^2L^2)}}$$

The output voltage $\mathbf{E_0}$ is given by

$$\mathbf{E_0}=-\mathbf{I_p Z_L}=-\frac{\mu \mathbf{E_g}}{r_p+j\omega L}\times j\omega L=\frac{-\mu \mathbf{E_g}}{1+r_p/j\omega L}$$

$$=\frac{\mu E_g}{\sqrt{\{1+(r_p/\omega L)^2\}}}\Big/\underline{270^\circ-\tan^{-1}\frac{\omega L}{r_p}} \quad ...(20)$$

Its magnitude is given by $\dfrac{\mu E_g}{\sqrt{\{1+(r_p/\omega L)^2\}}}$

Voltage gain

$$\mathbf{A}=\frac{-\mu j\omega L}{r_p+j\omega L}=\frac{\mu}{\sqrt{\{1+(r_p/\omega L)^2\}}}\Big|\underline{270^\circ-\tan^{-1}\frac{\omega L}{r_p}}. \quad ...(21)$$

Thus $$A=\mu\frac{1}{\sqrt{\{1+(r_p/\omega L)^2\}}}, \quad ...(22)$$

and $$\theta_A=270^\circ-\tan^{-1}\frac{\omega L}{r_p}=-90^\circ-\tan^{-1}\frac{\omega L}{r_p}. \quad ...(23)$$

Figure (15) shows the manner in which the voltage amplification vary with Z_L/r_p for pure resistance, pure inductance and complex inductive load. It is clear from the figure that for the same value of Z_L/r_p, pure inductive load produces greater voltage gain. In other words, we can say that to produce the same voltage gain, the impedance of a pure inductive load required is much smaller than a pure resistive load. The another advantage with an inductive load is that the plate supply voltage E_{bb} may be smaller than for a resistive load because of the steady component of voltage drop in the load resistor.

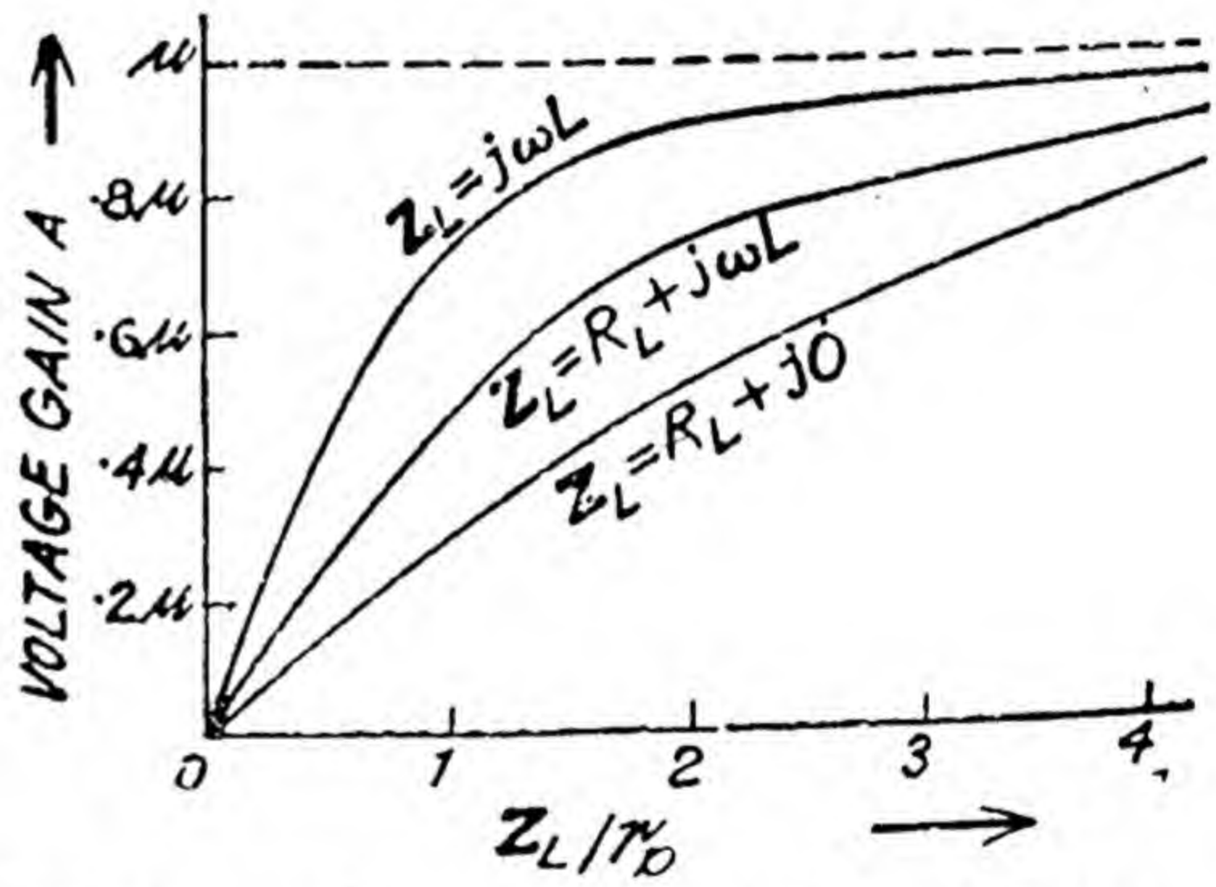


Fig. 15. Voltage gain of amplifier as a voltage with Z_L/r_p.

Curves of figure (16) show the form of the variation of phase angle of voltage gain with Z_L/r_p for pure resistive load, pure inductance load and for complex inductive load. In case of a pure resistive load, the

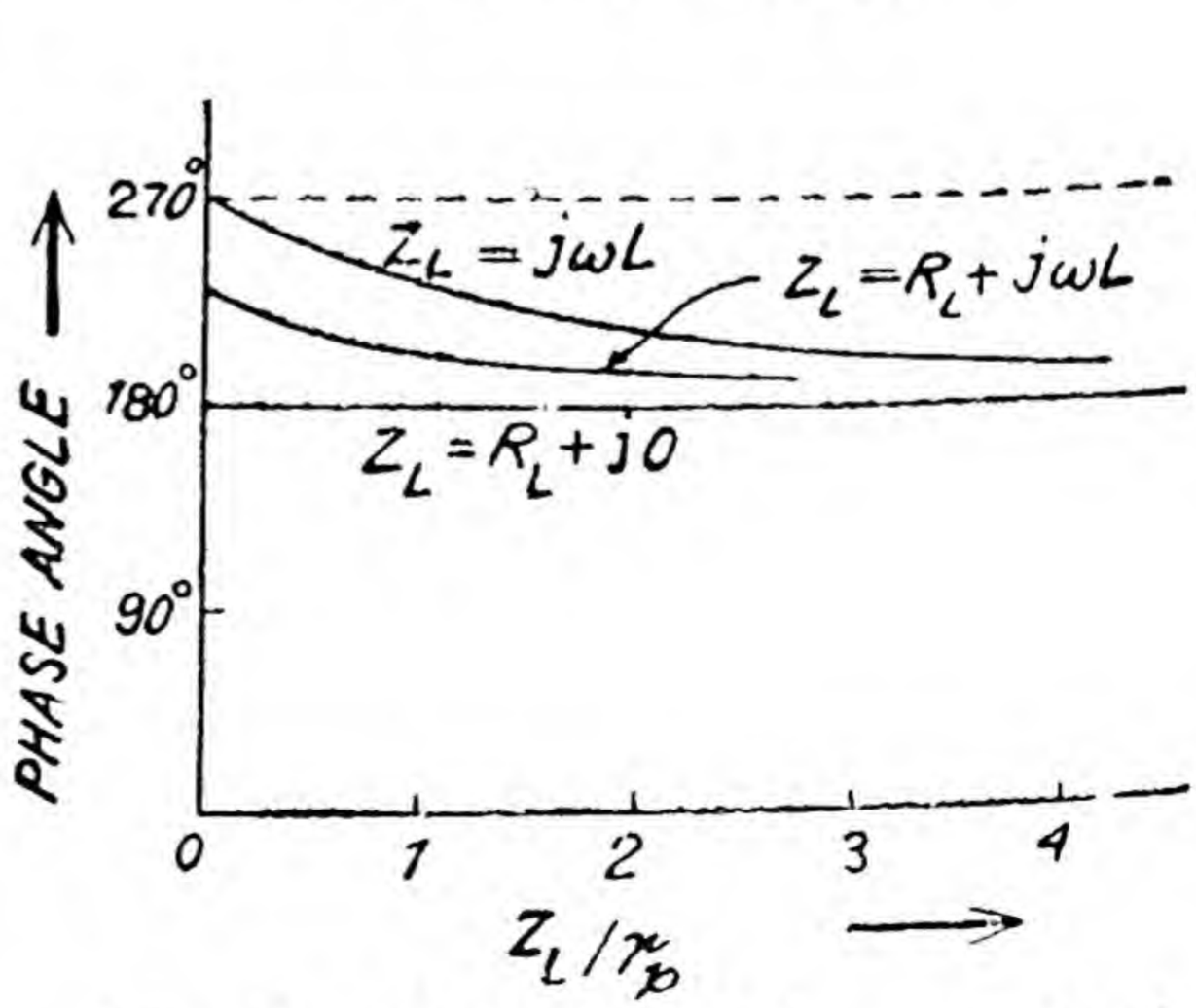


Fig. 16. Variation of phase angle of voltage with Z_L/r_p.

phase angle has a constant value 180°. For pure inductive load and for $Z_L/r_p=0$, the phase angle is 270°. As the value Z_L/r_p increases, the phase angle reduces and approaches to 180° for large value of Z_L/r_p. The curve for a complex inductive load lies in between the two curves.

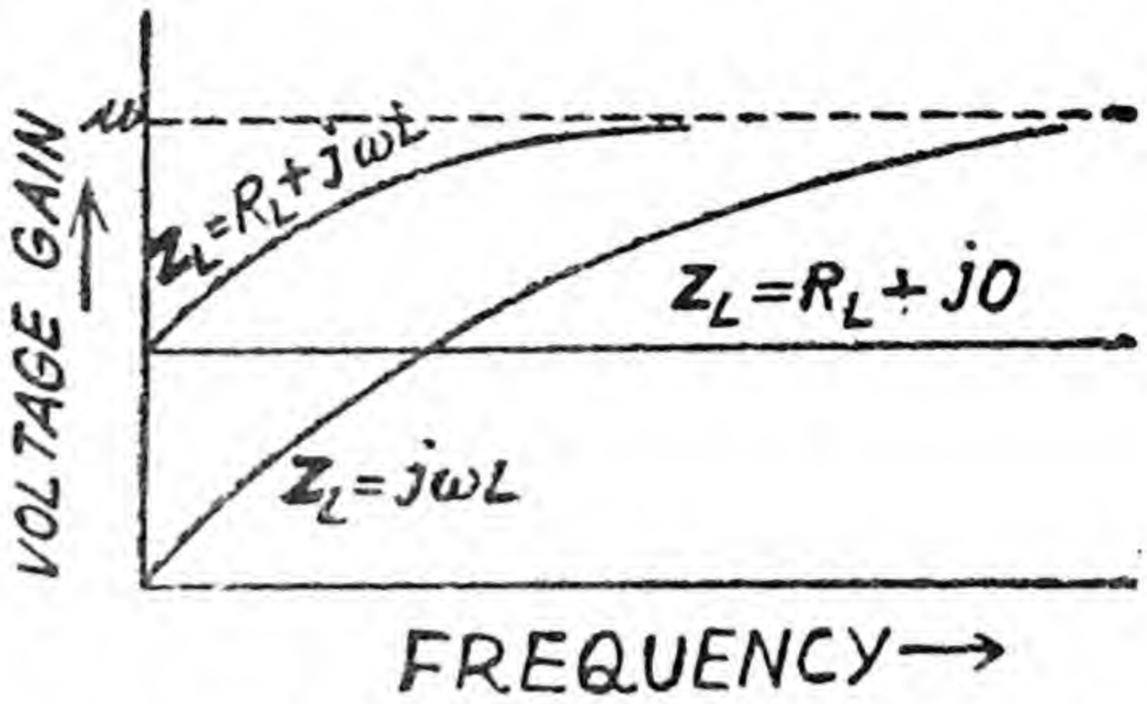


Fig. 17. Showing the variation of voltage gain with frequency.

The variation of voltage gain with frequency is also shown in figure (17). The voltage gain is independent of frequency in pure resistive load while increases in case of inductive and complex inductive loads.

9·9. CLASSIFICATION OF AMPLIFIERS :

Amplifiers are classified in ways descriptive of their character and properties. They are commonly classified in five ways :

(1) according to use,
(2) according to circuits,
(3) interms of their frequency range of operation,
(4) according to portion of the cycle during which plate current flows, *i.e.* upon their operating characteristics, and

(5) according to the range of frequencies amplified compared with central frequency.

(1) According to use to which they are subjected :

(*a*) Voltage amplifier,
(*b*) Power amplifier.

In voltage amplifier, the main aim is to get the maximum possible voltage gain. The tubes used have relatively high amplification factor μ. Such amplifiers may have high impedances load which may be tuned or untuned.

Tubes used as power amplifiers have low amplification factor and high transconductance. The power amplifiers may control a large plate current at a relatively high plate potential.

(2) According to circuit :

In this method of classification, amplifiers may be classified as (*a*) single stage amplifiers and (*b*) cascade amplifiers. Single stage amplifier uses only one amplifier and is suitable when small amplification is sufficient. The cascade amplifier consists of many amplifiers. The output of one amplifier is fed to the input of next amplifier. In this way considerable gain is possible. The cascade amplifiers are further classified according to the method of coupling as follows :

(*i*) Direct coupled amplifiers using d.c. coupling between the stages.

(*ii*) Resistance-capacitance coupled amplifiers using resistance as the load and a capacitor as the coupling element.

(*iii*) *L-C* coupled amplifiers, using inductor as the load and a capacitor as the coupling element.

(*iv*) Transformer-coupled ampliners using transformer as the coupling element.

(3) Interms of frequency range of operation :

(*a*) *Audio frequency amplifiers* : (30 c/s. to 1500 c/s.).

They are used in ratio receivers, public address systems and other amplifications involving audible frequencies. The frequency range of such units can be as narrow as 400 to 4000 cycles for amplifiers intended for speech only or 50 to 500 cycles for average quality music system as wide as 20 to 20,000 cycles for full high fidelity systems

(*b*) *Viedo frequency amplifiers* : (50 c/s to 6 mc/s).

These are used in Television receivers, radar receivers, and other devices producidg visible pictures. The frequency range of these units is much greater than audio amplifiers. For use in television, the typical frequency band needed is from approximately 50 cycles to ·5 megacycles. For use in some oscilloscopes or in radar, even higher upper frequency ranges are often needed.

(*c*) *Intermediate frequency amplifiers* : (150 kc/s. to many mc/s.)

They are used in resonant circuits and to amplify a relative narrow band of frequencies to each side of a mean or carrier frequency. In fact they are radio frequency amplifiers but the intermediate frequency is used, because somewhere in the complete unit there are one or more other amplifiers operating at some higher radio frequency. Broadcast band radio receivers usually employ an intermediate frequency of approximately 455 k.c. Ultra high frequency receivers use intermediate frequency values around 60 megacycles.

(*d*) *Radio frequency amplifiers* : (100 kc/s. to many thousand mc/s.)

These amplifiers, in general, are used to amplify a relatively narrow band of frequencies to each side of some mean value of a radio frequency. For example, in broadcast band radio receiver, a radio frequency amplifier may be set to operate at frequency of 1200 k.c. $\pm$ 5 k.c.

(4) According to tube operation : Another means of classifying amplifiers are by the duration of the conduction period in relation to one cycle of the input signal. In this respect, circuits are divided into four categories :

(a) Class A : Plate current flows for the full cycle of input. This is based on class *A* operation.

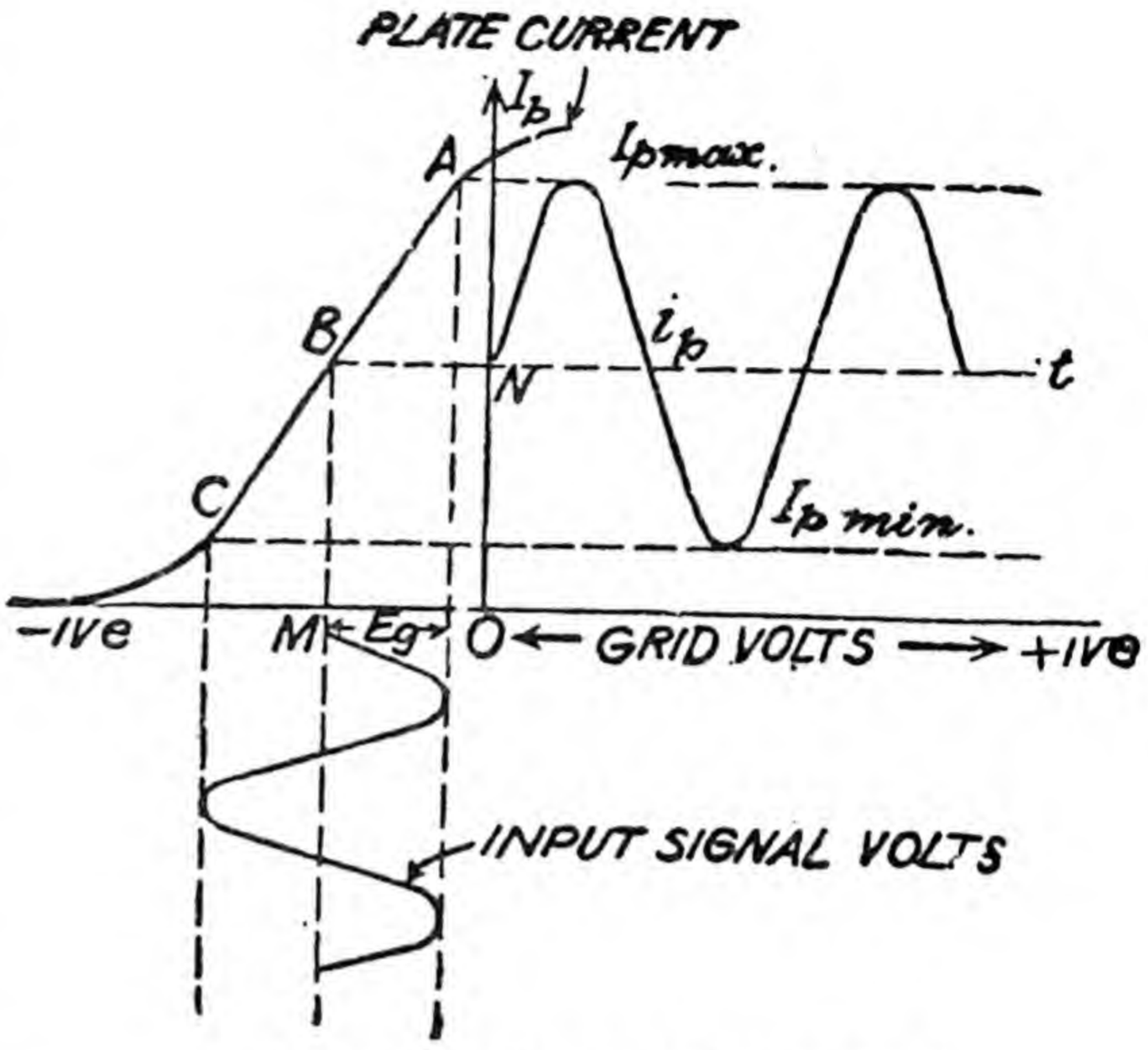


Fig. 18. Class *A* operation.

Class A operation : In this operation grid bias and the input alternating signal voltages are of such values that plate current flows in the output circuit at all times and the output wave shapes of the plate current are practically the same as the existing grid voltage. In order to realize the above condition, the operating point *B* is so chosen, as to allow the grid to be positive at the positive peaks of the signal and to be in the curved region of the characteristic at the negative peaks of the signal voltage.

(b) Class AB : Based on class *AB* operation in which current flows for less than one cycle, but for more than half cycle (fig. 19).

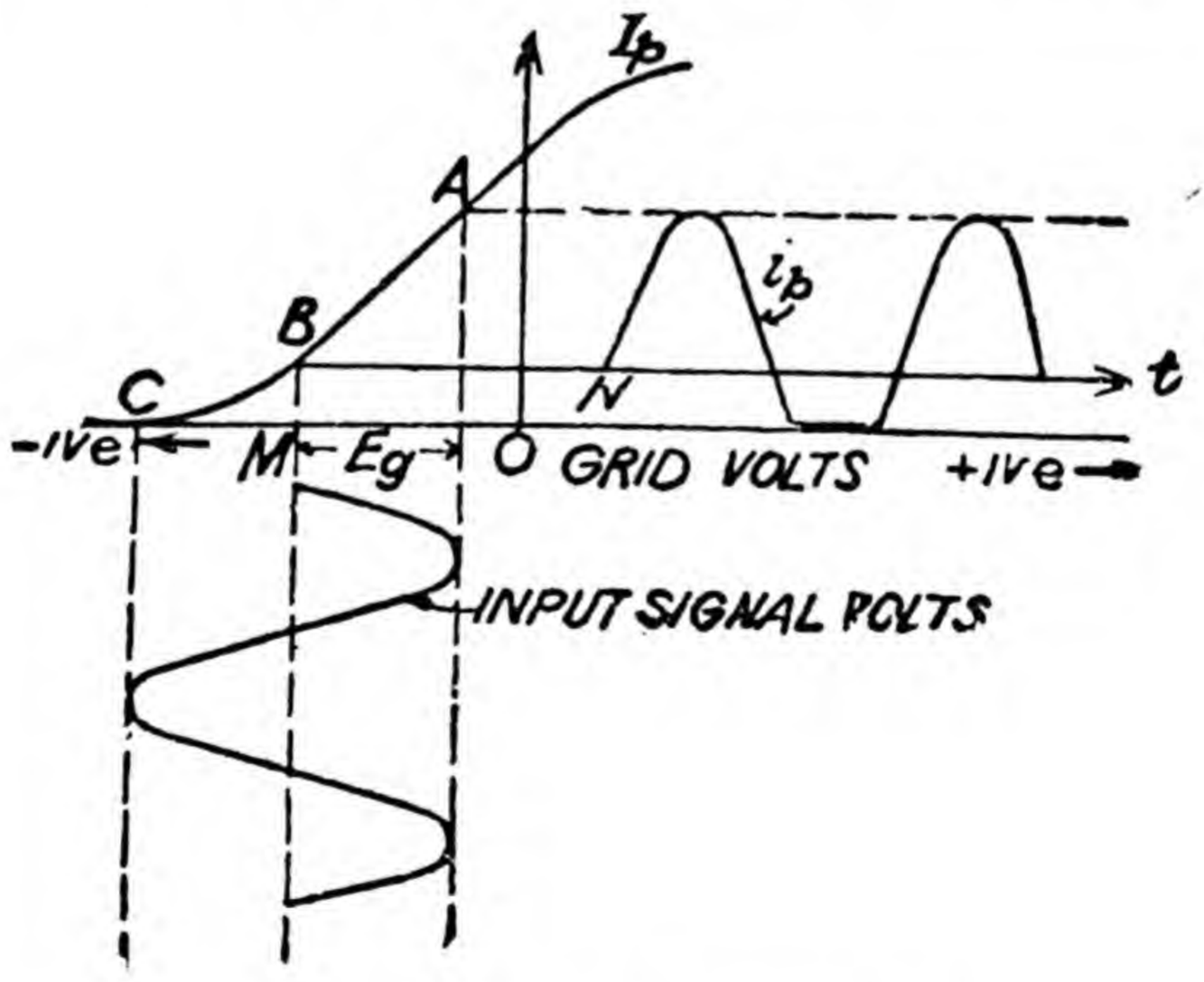


Fig. 19. Class *AB* operation.

Class AB Operation : Class *AB* operation is one in which the grid bias and the signal input voltage are of such values that plate current will flow for more than half, but less than the complete time of the input cycle. The operating point *B* is such that at negative peaks of the signal voltage, the grid reaches beyond the cut-off value. *i.e.* for a part of negative input signal, plate current becomes zero.

(c) Class B : Current flows for approximately one half cycle of the input signal and is based on class *B* operation.

Class B Operation : Class *B* operation is one in which the grid bias is made approximately equal to the cut-off value. In this case, the plate current will be approximately zero for zero-signal input and plate current will flow for one half of the input cycle, *i.e.* in this case the grid is biased just in the vicinity of cut-off value so that only positive half of the signal voltage is effective and other half is completely ineffective.

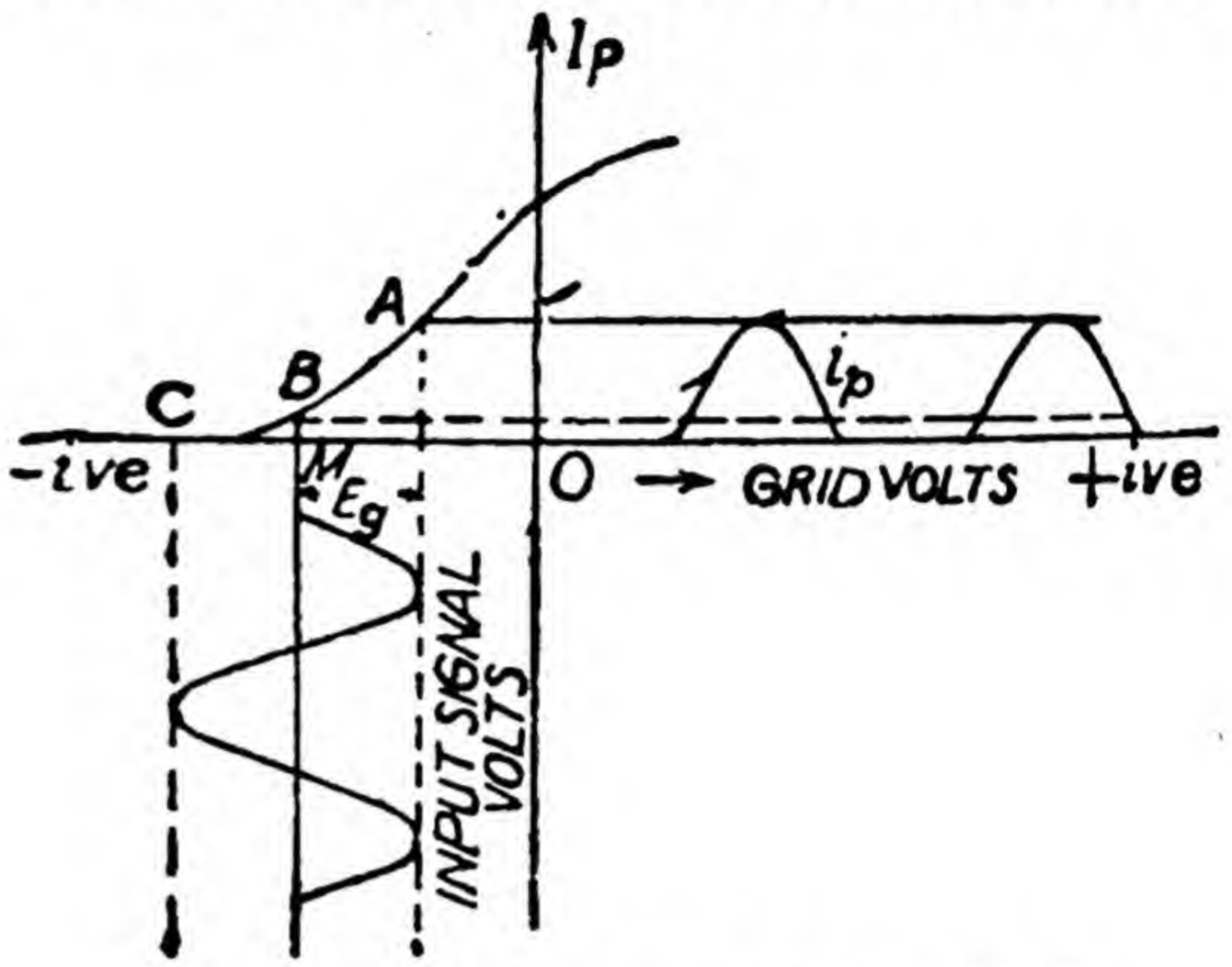


Fig. 20. Class B operation.

(d) Class C : Current flows for less than one half cycle of input (Fig. 21). The power amplifiers are based on class *C* operation.

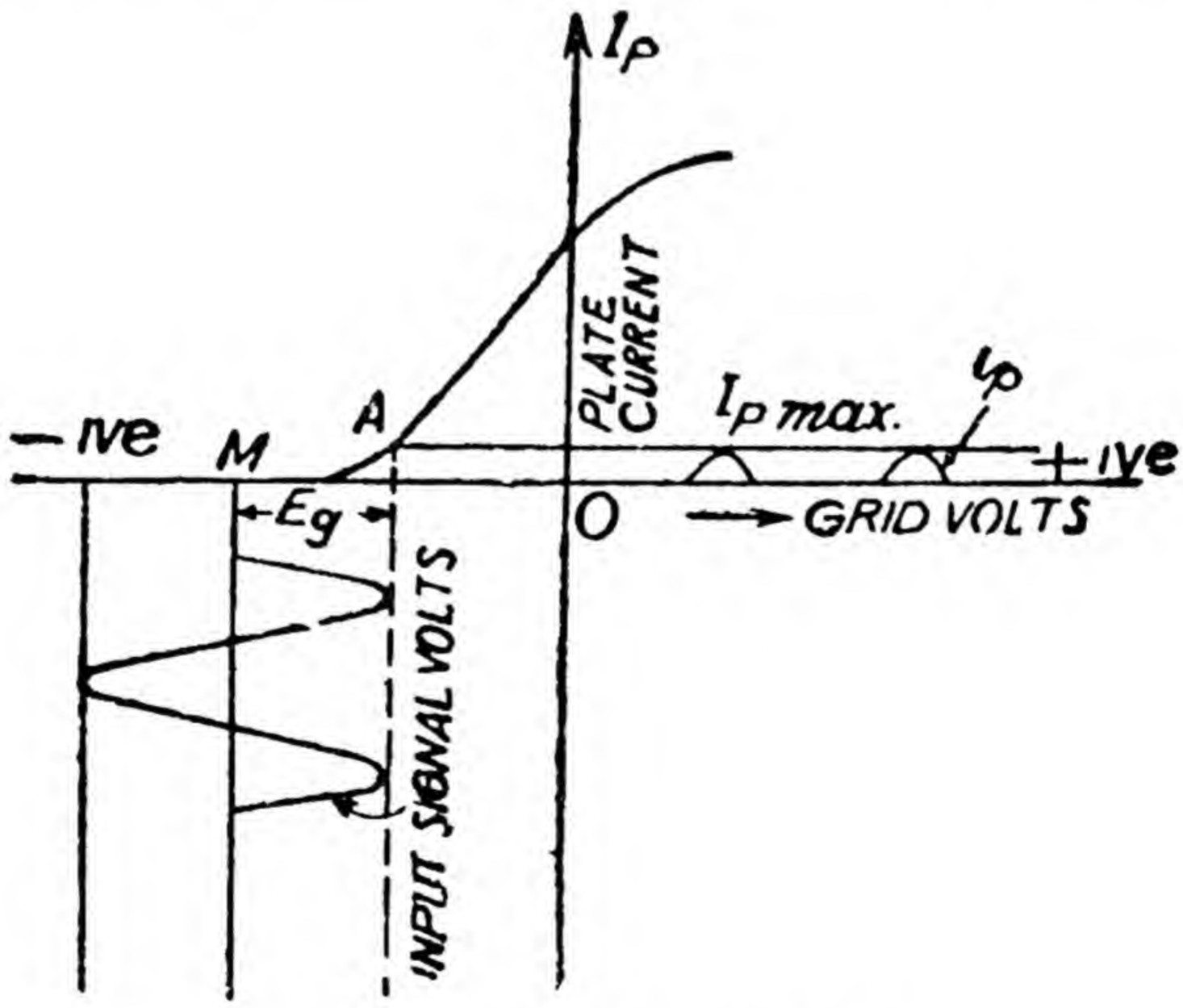


Fig. 21. Class *C* operation.

Class C operation : Class *C* operation is one in which the grid bias is considerably greater than cut-off value, so that the plate current is zero for zero signal input and plate current will flow for less than one-half the time of the input cycle. In this case, the operating point is chosen beyond the cut-off value ; hence we get only pulses of plate current corresponding to the positive peaks of the signal voltage.

(5) According to the range of frequencies amplified compared with central frequency :

These amplifiers are classified as (*i*) narrow band amplifiers and (*ii*) wide band amplifiers.

(*i*) A narrow band amplifier is one in which the band of frequencies amplified is small in comparison with the central frequency. Most of the radio frequency tuned amplifiers are narrow band amplifiers.

(*ii*) A wide band amplifier is one in which the band of frequencies amplified is large compared with the central frequency. RC coupled amplifiers come under this category.

Types of Amplifications :

(*i*) *Voltage amplification* : Voltage amplification is the ratio of the signal voltage available at the output terminals of an amplifier, transformer or other four terminal network, to the signal voltage impressed at the input terminals.

(*ii*) *Current amplification* : Current amplification is the ratio of the signal current produced in the output circuit of an amplifier to the signal current supplied to its input circuit.

(*iii*) *Power amplification* : Power amplification is the ratio of the power delivered to the output circuit of an amplifier to the power supplied to the input circuit.

In case of power amplifier, it is usually better to specify performance by its power sensitivity.

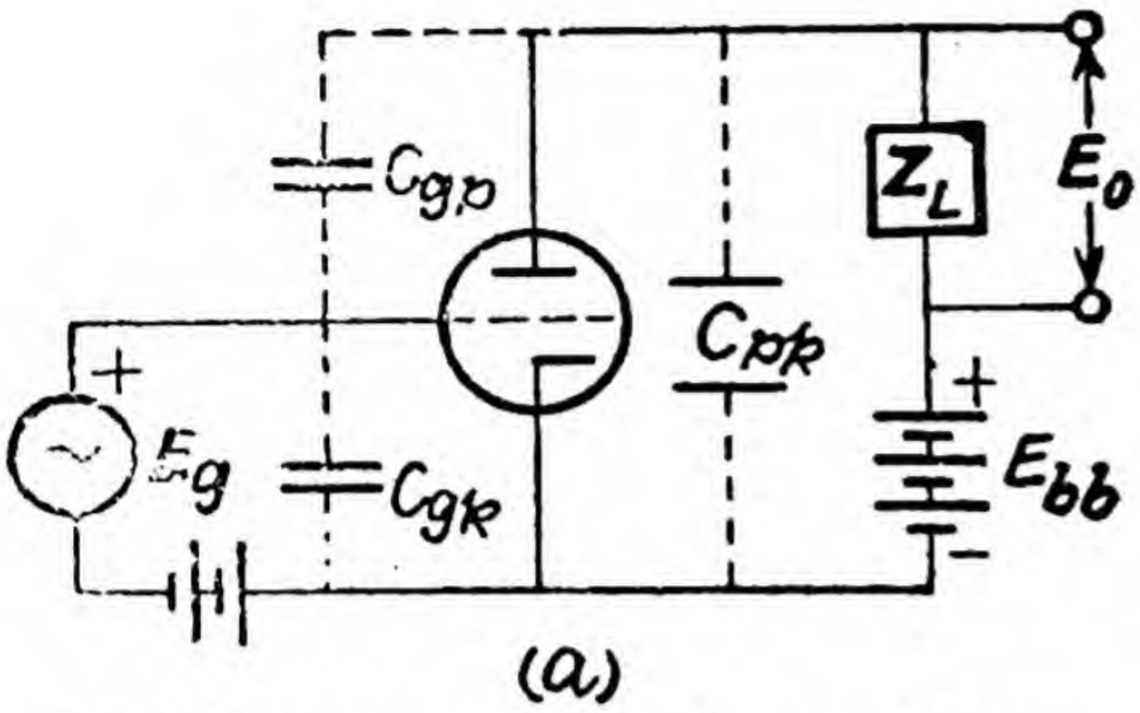
(*iv*) *Power sensitivity* : Power sensitivity is defined as the ratio of the signal frequency power, delivered by the output circuit of an amplifier, to the square of the effective value of the signal voltage impressed at the input terminals. This is measured in ohms.

9·10. INTER-ELECTRODE CAPACITANCE IN A TRIODE :

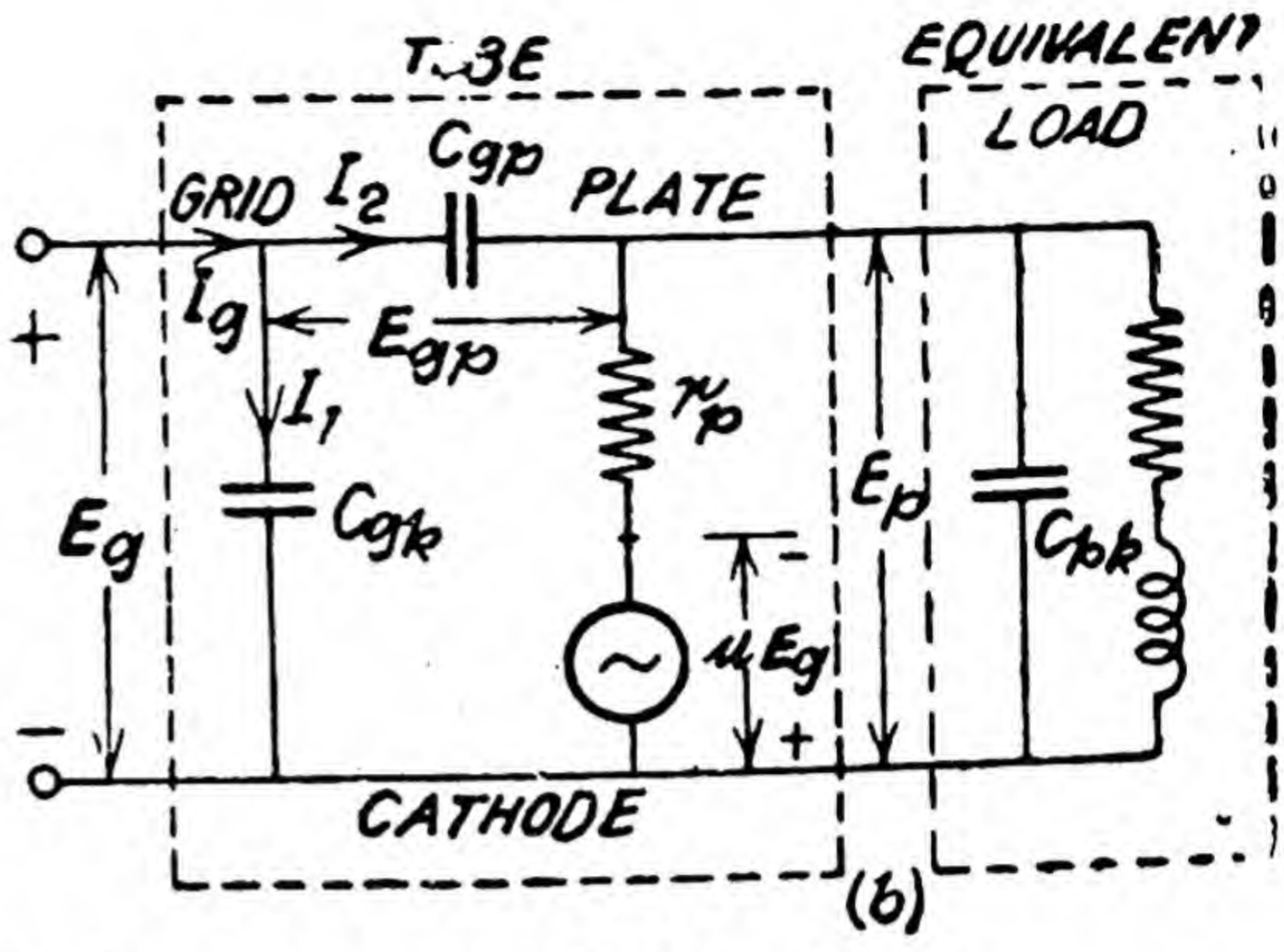
The grid of a vacuum tube amplifier is maintained at a negative potential with respect to cathode so that even for the positive cycle of grid voltage signal, the grid does not become positive and no grid current flows. This is only true for alternating current of very low frequencies. The alternating current of high frequencies can flow in the grid circuit due to interelectrode capacitances that exist between the different electrodes. These capacitances are important in the behaviour of the circuit.

Fig. 22 shows the schematic diagram of the triode amplifier and its a.c. equivalent circuit including the interelectrode capacitances.

In the circuit, C_{gp} denotes the capacitance between the grid and plate, C_{gk} is the capacitance between grid and cathode and C_{pk} is the plate-cathode capacitance. These capacitances are distributed throughout the length of the electrodes and exist inside the tube. For convenience these are taken as lump capacitances external to the tube.



(a) Vacuum tube triode amplifier with different inter-electrode capacitances.



(b) Equivalent a.c. circuit of triode amplifier with different inter-electrode capacitances.

Fig. 22.

The capacitance C_{pk} exists directly across the output terminals and hence acts if it were a part of load. The capacitance C_{gk} exists directly across the input terminals and hence acts as if it were a part of the grid signal. The capacitance C_{gp} is a mutual capacitance that interconnects the grid and plate circuit. Thus, it constitutes a feed back path from the plate to the grid, as well as a forward path from the grid to the plate. Since at low frequencies this capacitance is small and hence the approximation that the plate circuit is independent of grid circuit is valid but at higher frequencies this is not true. At higher frequencies, these capacitances seriously affect the operation. These capacitances have two effects: (i) modification of input admittance and (ii) modification of complex voltage gain.

9·11. INPUT ADMITTANCE OF A TRIODE AMPLIFIER :

We have seen that due to the presence of inter-electrode capacitance C_{gp} between grid and plate, the grid circuit is not isolated from plate circuit. When an alternating signal $\mathbf{E}_g$ is applied between grid and cathode an alternating current flows through the interelectrode capacitance between grid and cathode. The currents through C_{gk} and C_{gp} are supplied by signal source. This results in the input admittance of the amplifier. The input impedance $\mathbf{Z}_i$ of the triode amplifier is defined as the ratio of the input signal voltage $\mathbf{E}_g$ and $\mathbf{I}_g$ (current flowing through the grid). The reciprocal of the input impedance is known as input admittance $\mathbf{Y}_i$. Thus

Input impedance $$\mathbf{Z}_i = \frac{\mathbf{E}_g}{\mathbf{I}_g}.$$

and Input admittance $$\mathbf{Y}_i = \frac{\mathbf{I}_g}{\mathbf{E}_g}.$$

An expression for the input admittance of amplifier is derived below :

Refering to figure 22 (a) and 22 (b), let $\mathbf{I}_1$ and $\mathbf{I}_2$ be the effective values of current through C_{gk} and C_{gp}. Now

$$\mathbf{I}_g=\mathbf{I}_1+\mathbf{I}_2. \qquad \ldots(1)$$

The current $\mathbf{I}_2$ through the grid plate capacitance C_{gp} in figure 22 (b) is given by

$$\mathbf{I}_2=\mathbf{E}_{gp}\, j\omega C_{gp}, \qquad \ldots(2)$$

where $\mathbf{E}_{gp}$ is the complex effective value of the voltage rise from the plate to the grid. This voltage is given by

$$\mathbf{E}_{gp}=\mathbf{E}_g-\mathbf{E}_p. \qquad \ldots(3)$$

We know that complex voltage gain $\mathbf{A}=\dfrac{\mathbf{E}_p}{\mathbf{E}_g}$

or $$\mathbf{E}_p=\mathbf{A}\ \mathbf{E}_g. \qquad \ldots(4)$$

Substituting the value of $\mathbf{E}_p$ from equation (4) in equation (3), we get

$$\mathbf{E}_{gp}=\mathbf{E}_g-\mathbf{A}\ \mathbf{E}_g=\mathbf{E}_g\ (1-\mathbf{A}). \qquad \ldots(5)$$

Substituting the value of $\mathbf{E}_{gp}$ from equation (5) in equation (2)

$$\mathbf{I}_2=j\omega\ C_{gp}\ \mathbf{E}_g\ (1-\mathbf{A}). \qquad \ldots(6)$$

again $$\mathbf{I}_1=j\omega\ C_{gk}\ \mathbf{E}_g. \qquad \ldots(7)$$

From equations (6) and (7), the value of $\mathbf{I}_g$ is given by

$$\mathbf{I}_g=\mathbf{I}_1+\mathbf{I}_2=j\omega\ C_{gk}\ \mathbf{E}_g+j\omega\ C_{gp}\ \mathbf{E}_g\ (1-\mathbf{A})$$

$$=j\omega\ \mathbf{E}_g\ [C_{gk}+C_{gp}\ (1-\mathbf{A})]. \qquad \ldots(8)$$

Now the input admittance $\mathbf{Y}_i=\dfrac{\mathbf{I}_g}{\mathbf{E}_g}$

$$\mathbf{Y}_i=j\omega\ [C_{gk}+C_{gp}\ (1-\mathbf{A})] \qquad \ldots(9)$$

The voltage gain **A** may be put as $\mathbf{A}=A_r+jA_i$, where A_r and A_i are the real and imaginary components of complex voltage gain **A** respectively. Now

$$\mathbf{Y}_i=j\omega\ C_{gk}+j\omega\ C_{gp}\ (1-A_r)+\omega\ C_{gp}\ A_i \qquad \ldots(10)$$

$\therefore$ Input shunt conductance $=\omega\ C_{gp}\ A_i$ $\qquad \ldots(11)$

and input shunt susceptance $=\omega\ [C_{gk}+C_{gp}\ (1-A_r)]$ $\qquad \ldots(12)$

The capacitance between grid and plate (C_{gp}), thus results in a much larger capacitance $C_{gp}\ (1-A_r)$ between grid and cathode. This is known as "Miller effect".

In accordance with equation (10), the input a.c. equivalent circuit of a triode amplifier may be represented by a shunt resistance $R=1/\omega C_{gp}A_i$ in parallel with input capacitance $C_{gk}+C_{gp}(1-A_r)$ as shown in figure 23.

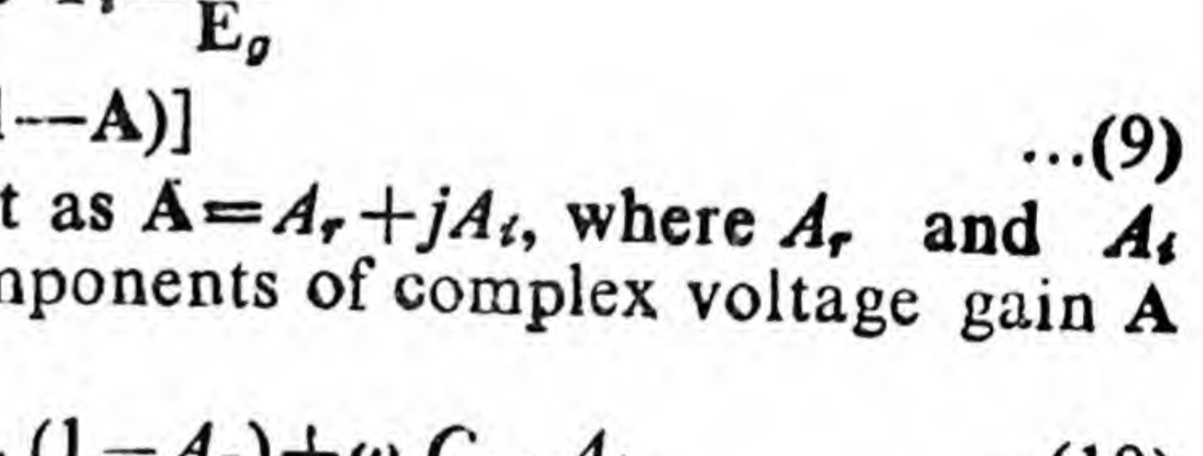


Fig. 23. A.C. equivalent input circuit of triode amplifier.

The value of complex voltage gain **A** and $\mathbf{Y}_i$ in case of pure resistive load

and complex load are calculated below :

Case I. Pure resistive load :

Neglecting the reactances of C_{gp} and C_{pk} when the frequency is low, the complex voltage gain A in pure resistive load R_L is given by

$$\mathbf{A} = -\mu \frac{R_L}{r_p + R_L} \qquad \ldots(13)$$

In this case, $A_r = -\mu \frac{R_L}{r_p + R_L}$ and $A_i = 0$.

$$\therefore \quad \mathbf{Y}_i = j\omega \left[C_{gk} + C_{gp} \left(1 + \frac{\mu R_L}{r_p + R_L} \right) \right] \qquad \ldots(14)$$

Therefore, the input capacity C_i of the triode amplifier is equal to

$$C_i = C_{gk} + C_{gp} \left(1 + \frac{\mu R_L}{r_p + R_L} \right) \qquad \ldots(15)$$

Case II. Complex load :

Now we shall find out the nature of the input impedance of a triode amplifier whose load is a resistance in series with an inductance. Let effective load impedance including the effect of C_{pk} be $\mathbf{Z}$, then

$$\mathbf{Z} = R + jX. \qquad \ldots(16)$$

$$\text{Now } \mathbf{A} = \frac{\mathbf{E}_0}{\mathbf{E}_v} = -\frac{\mu \mathbf{Z}}{r_p + \mathbf{Z}} = -\frac{\mu (R + jX)}{(r_p + R + jX)}$$

$$\text{or} \quad A_r + jA_i = -\frac{\mu (R + jX)}{(r_p + R) + jX}$$

$$= -\mu \frac{R(r_p + R) + X^2}{(r_p + R)^2 + X^2} - j \frac{\mu r_p X}{(r_p + R)^2 + X^2} \qquad \ldots(17)$$

Comparing the real and imaginary terms on both sides, we have

$$A_r = -\frac{\mu R (r_p + R) + X^2}{(r_p + R)^2 + X^2} \qquad \ldots(18)$$

and

$$A_i = -\frac{\mu r_p X}{(r_p + R)^2 + X^2}. \qquad \ldots(19)$$

Substituting values of A_r and A_i from equations (18) and (19) in equations (11) and (12), we have input shunt conductance

$$= -\mu\omega C_{gp} \frac{r_p X}{(r_p + R)^2 + X^2} \qquad \ldots(20)$$

and input shunt susceptance

$$= \omega \left[C_{gk} + C_{gp} \left(1 + \frac{\mu (r_p + R) R + X^2}{(r_p + R)^2 + X^2} \right) \right] \qquad \ldots(21)$$

$$\therefore \quad Y_i = j\omega \left[C_{gk} + C_{gp} \left\{ 1 + \frac{\mu R (r_p + R) + X^2}{(r_p + R)^2 + X^2} \right\} \right]$$

$$- \mu\omega C_{gp} \frac{r_p X}{(r_p + R)^2 + X^2} \qquad \ldots(22)$$

Thus Y_i contains a resistive component [last term of equation (22) whose value is negative if X is positive *i.e.*, if the anode load is inductive]. This negative resistance means that power is flowing back to the grid through the grid-plate capacity and if this power is more than which is required to supply the power dissipated in any positive resistance in the source of the grid voltage, oscillations will result. This is one of the chief difficulties in the use of triodes as amplifiers at radio frequencies. On the other hand, if the anode load is capacitive, the resistive component of the grid input impedance is always positive. A positive resistance means that the energy is transferred from the grid to plate circuit through C_{gp}.

In the above treatment, we have neglected the fact that the current through C_{gp} also flows through the plate load, thereby alternating the plate voltage slightly. With a more accurate treatment it can be shown that C_{gp} should be replaced by

$$C_{gp}\left\{1+\left(\frac{Zr_p}{Z+r_p}\right)j\omega C_{gp}\right\}^{-1}$$

9·12. METHODS OF BIASING THE CONTROL GRID :

Commonly three methods of biasing the control grid are used. They are : (*i*) fixed bias, (*ii*) cathode bias or self bias and (*iii*) grid leak bias. Here we shall discuss these methods.

(i) Fixed bias : It has already been mentioned that the grid should be maintained at a negative potential with respect to cathode. The simpler method is to connect a dry battery between grid and cathode. This is cumbersome and moreover it is difficult to maintain a constant voltage battery in the circuit. To overcome this trouble the grid bias voltage is obtained from the plate supply source with the help of potential divider. This provides a fixed bias.

(ii) Cathode bias or self bias : The economical method for providing the grid bias is to place a resistance R_k (cathode bias resistor) in between the cathode and ground. The advantage of this method is that no power supply is necessary for grid bias. The circuit arrangement of a triode amplifier using self bias is shown in figure 24.

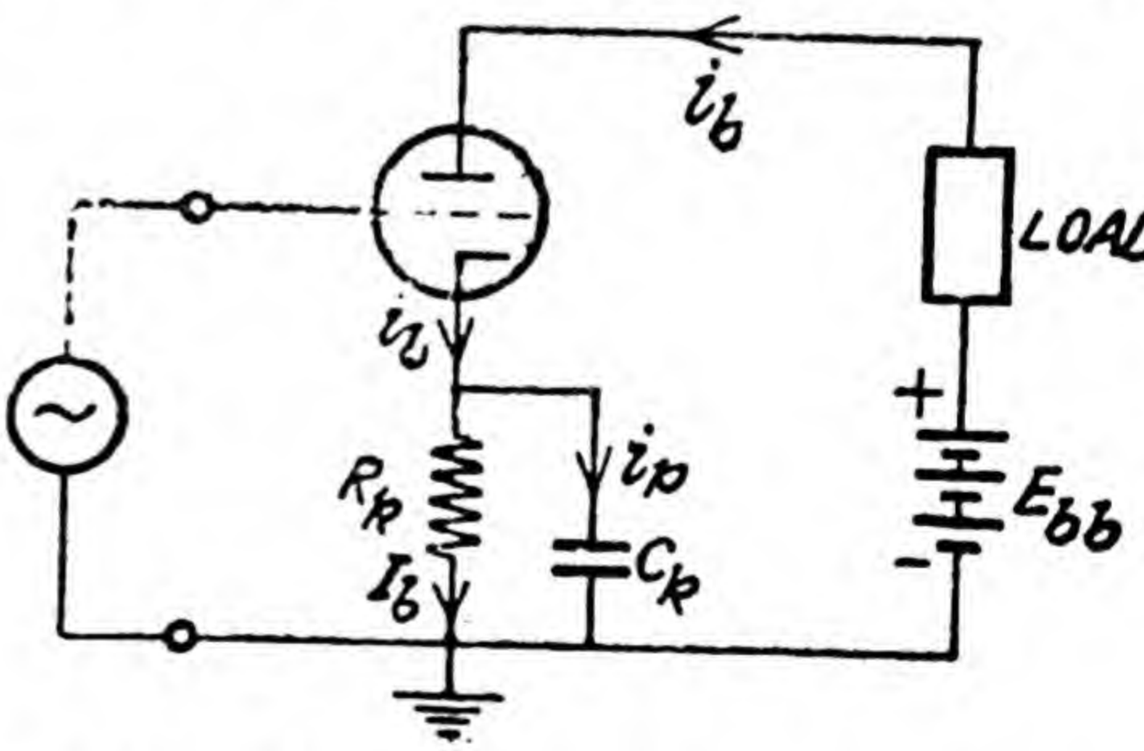


Fig. 24. Vacuum tube triode amplifier using self bias or cathode bias.

When no signal is applied, the plate currenr I_b flows through R_k and makes a voltage drop across it. The polarity of this voltage drop I_bR_k is such that the cathode end is positive and the ground end is negative. Thus the plate current flowing through R_k makes the grid negative with respect to cathode.

When a signal is applied and only R_k is introduced in cathode and ground, it is found that the gain of the amplifier is reduced. The reason is as follows :

When an a.c. signal is applied, the plate current i_b consists of two compenents : (*i*) the d.c. plate current I_b and (*ii*) a.c. component i_p. Now in this case, in addition to d.c. current I_b, the a.c. current i_p also flows through R_k. The a.c. current causes a voltage drop which is in opposite phase to the applied signal voltage. Thus the voltage across the grid cathode of the tube is the difference of the two voltages, *i.e.*., applied signal voltage and a.c. drop across R_k. This is a form of negative feed back and the gain of the amplifier is reduced.

The reduction of the gain can be prevented by connecting a capacitor C_k of high value and small reactance in parallel with R_k. When the resistance of the condenser is far smaller than R_k almost all a.c. component of current flows through C_k without making any voltage drop across R_k. The d.c. current I_b flowing through R_k provides the grid bais. *Thus R_k C_k combination in the cathode lead provides the grid bias.* High capacitance, low voltage electrolytic capacitors are used for this purpose. The use of this bais is restricted to class *A* and *AB* amplifiers.

(iii) Grid leak bais : Figure (25) shows a vacuum tube amplifier using grid leak bias. In grid leak bais resistor R_g is connected between grid and cathode and a capacitor C is connected in series with the grid as shown in figure 25.

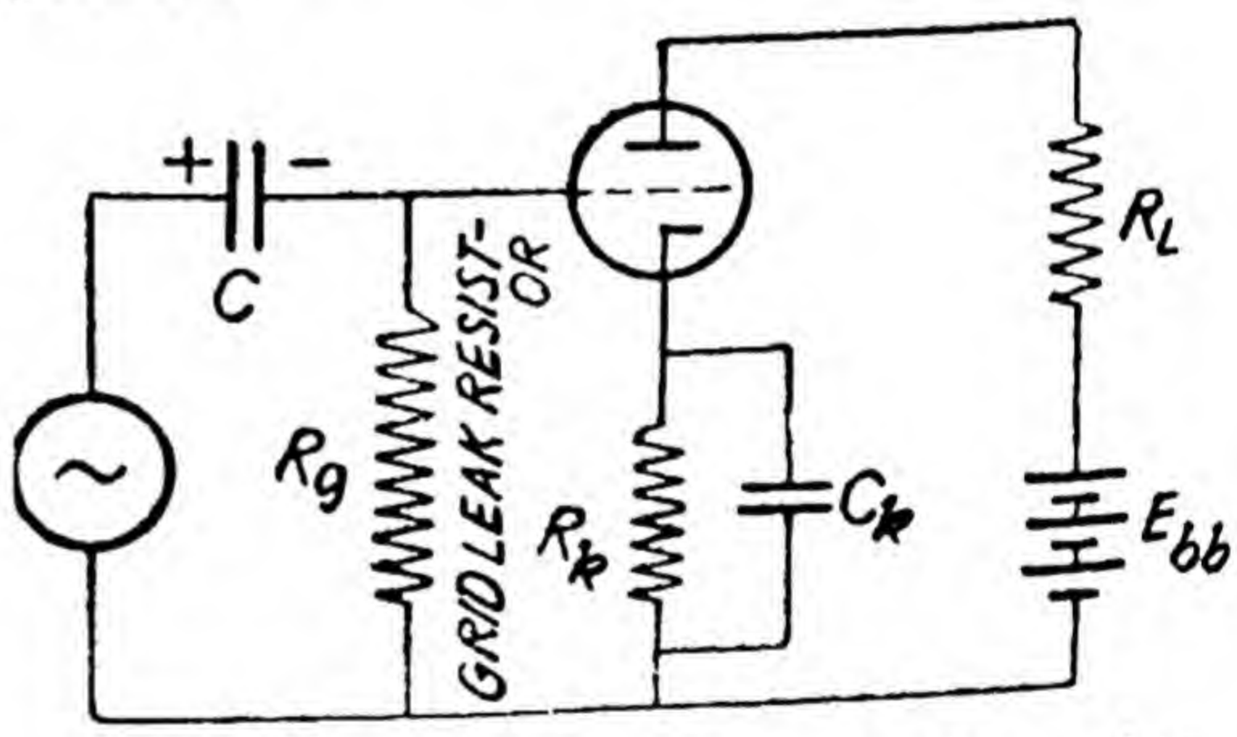


Fig. 25. Vacuum tube amplifier using grid leak bias.

When positive half cycle of the a.c. signal is applied, the total grid voltage reaches to zero potential or it may become positive. Due to this fact grid current flows and the condenser is charged with polarity shown in the fig. During the negative half cycle, the condenser discharges through grid leak resistor R_g because in this case no grid current flows. As R_g is of the order of $250K\Omega$, the time constant of the discharge of the condenser is very large. After few cycles of alternate charge and discharge, a constant d.c. voltage appears across the condenser which provides the necessary grid bias. The grid leak bias is generally used in conjuction with self bias. Usually the grid leak bias is used when the grid current flows *i.e.* for *AB*, *B* or *C* operations.

9·13. DISTORTION IN AMPLIFIERS :

The output voltage of an ideal class *A* amplifier should have the same waveform as the input signal. In order that the output

waveform shall be identical with that of the input, three conditions must be satisfied.

(*i*) The output must contain only those frequencies that are contained in the input.

(*ii*) The output must contain all frequencies contained in the input and the relative amplitude of various components must be the same in the input.

(*iii*) If any component of the output is shifted in phase relative to the corresponding component of the input, all components must be shifted by the same number of electrical degrees of the fundamental cycle of that component.

Any deviation present in the output waveform from input wave- is called distortion. Failure to satisfy the above conditions results form in three corresponding types of distortion :

1. Non-linear distortion or amplitude distortion.
2. Frequency distortion.
3. Phase distortion.

1. **Amplitude distortion :** Amplitude distortion is due to the production of new frequencies in the output which were not present in the input. From the dynamic (mutual) characteristic we note that any grid signal such as E_{g1} that falls within the straight line section of this dynamic characteristic (a_1 to b_1), will be amplified without distortion (Fig. 26 *a*). On the other hand, signals that fall outside the straight section, such as Eg_2, will have their peaks flattened. Since amplitude distortion is caused by operation in the non-linear areas, it is also called non-linear distortion.

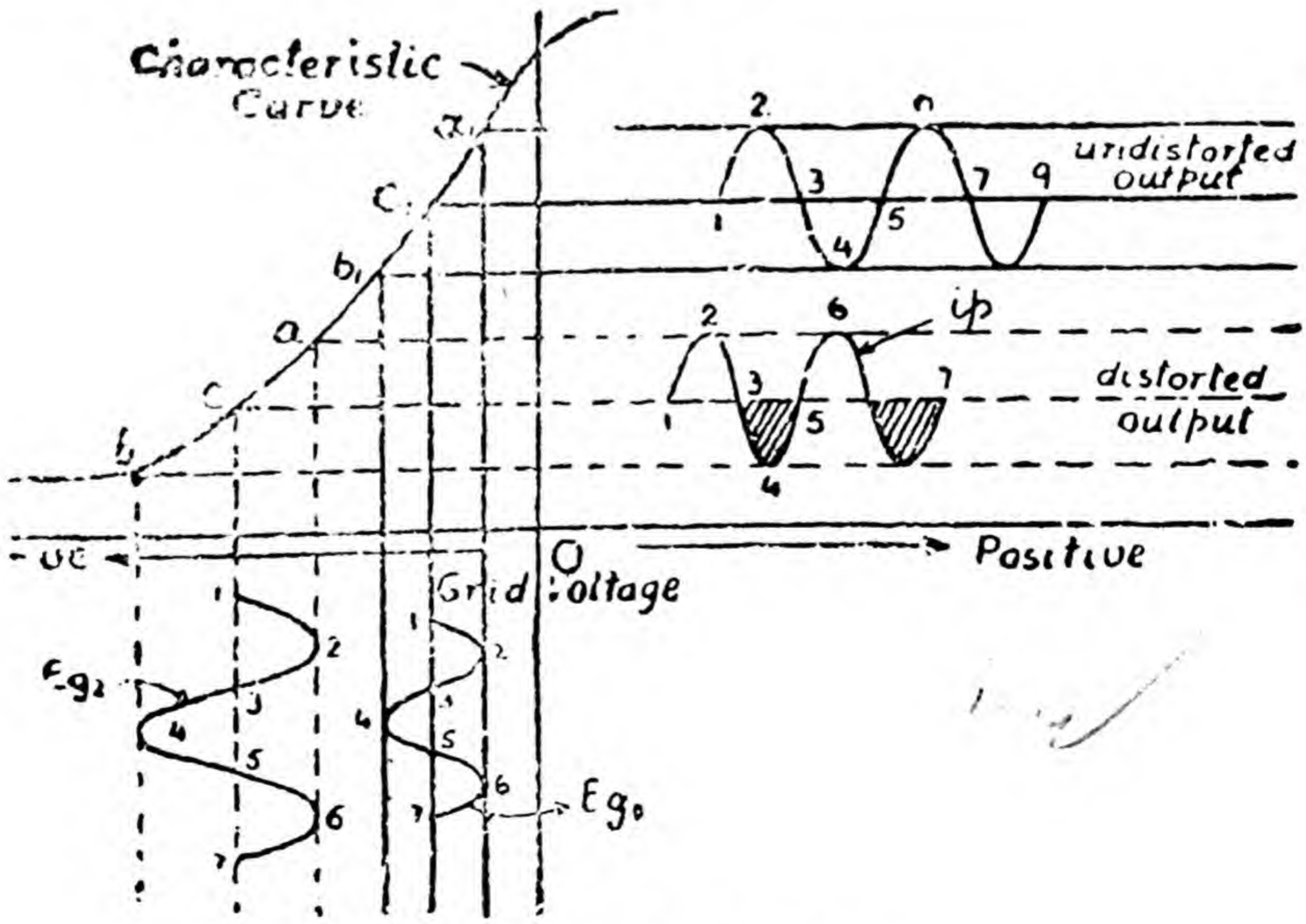


Fig 26 (a) Amplitude distortion.

Another type of amplitude distortion is caused by operating the tube with incorrect grid bias. When the tube is operated with a grid bias whose value is less than the maximum value of the input signal voltage, the grid voltage will be driven positive during the portion of each cycle in which the positive value of the input signal is greater than the grid bias. The manner in which the output signal is distorted for this type of operation is shown in fig. 26. (*b*).

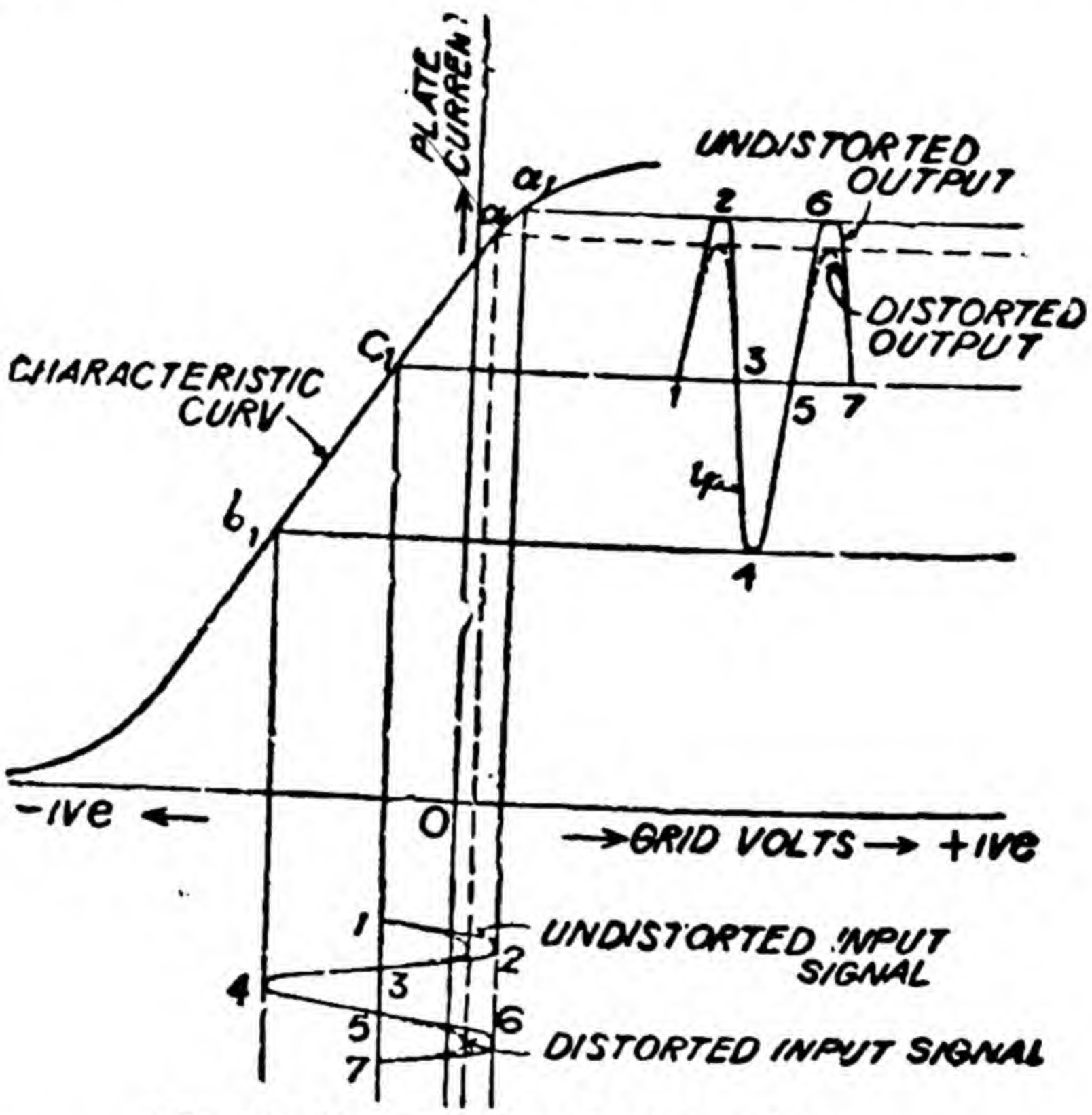


Fig. 26. (*b*) Amplitude distortion.

Distortion is also caused by applying too large an input signal. When the tube is operated with its correct value of grid bias but the applied signal is too large, either half or both halves of the input signal may be distorted.

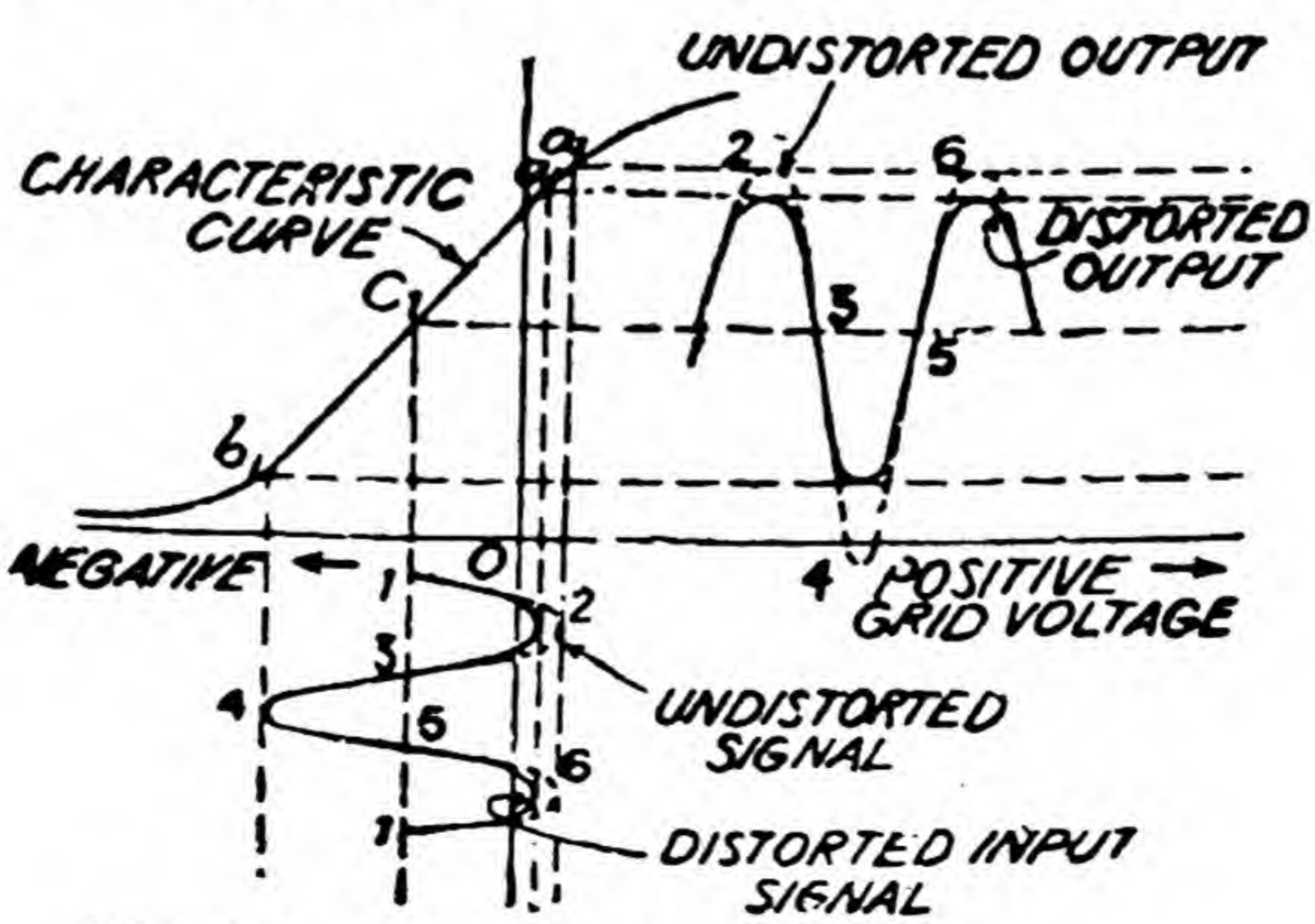


Fig. 27. Amplitude distortion.

When amplitude distortion occurs, the output waveshape is no longer a sine wave. The resulting complex wave consists of the original frequency and harmonics of the original frequency. The current and voltage relations may be expressed in general by the series

$$I_p = I_0 + aE_g + bE_g^2 + cE_g^3 + \ldots \qquad \ldots(1)$$

Assume that the input wave is a simple sine function of time, of the form

$$E_g = \hat{E}_g \sin \omega t.$$

The current is

$$I_p = I_0 + a\hat{E}_g \sin \omega t + b\hat{E}_g^2 \sin^2 \omega t + c\hat{E}_g^3 \sin^3 \omega t + \ldots$$

$$= I_0 + a\hat{E}_g \sin \omega t + \frac{b\hat{E}_g^2}{2} - \frac{b\hat{E}_g^2}{2} \cos 2\omega t + \ldots$$

$$= I_0 + \frac{b\hat{E}_g^2}{2} + a\hat{E}_g \sin \omega t - \frac{b\hat{E}_g^2}{2} \cos 2\omega t + \ldots$$

Thus the current contains

(*i*) $\frac{b\hat{E}_g^2}{2}$, *i.e.* rectified portion of the output current arising due to distortion.

(*ii*) $b\hat{E}_g \sin \omega t$, *i.e.*, output current with the frequency of the input voltage.

(*iii*) $-\frac{b\hat{E}_g^2}{2} \cos 2\omega t$, *i.e.*, second harmonic component of the plate current.

Percentage of second harmonic distortion

$$= \frac{\text{amplitude of second harmonic}}{\text{amplitude of fundamental}} \times 100\%.$$

$$= \frac{\frac{b\hat{E}_g^2}{2} \times 100}{a\hat{E}_g} = \frac{50b\hat{E}_g}{a}\%.$$

Knowing the values of a and b, percentage distortion can be calculated. If the maximum, mean and minimum values of the current is obtained from the dynamic characteristic, the values of a and b are calculated in the following manner :

From fig. 28, it is observed that

when $\omega t = 0$, $I_p = I_p$ (mean) $= I_0$.

$\omega t = \pi/2$, $I_p = I_p$ (max.) $= I_0 + a\hat{E}_g + b\hat{E}_g^2$.

$\omega t = 3\pi/2$, $I_p = I_p$ (min.) $= I_0 - a\hat{E}_g + b\hat{E}_g^2$.

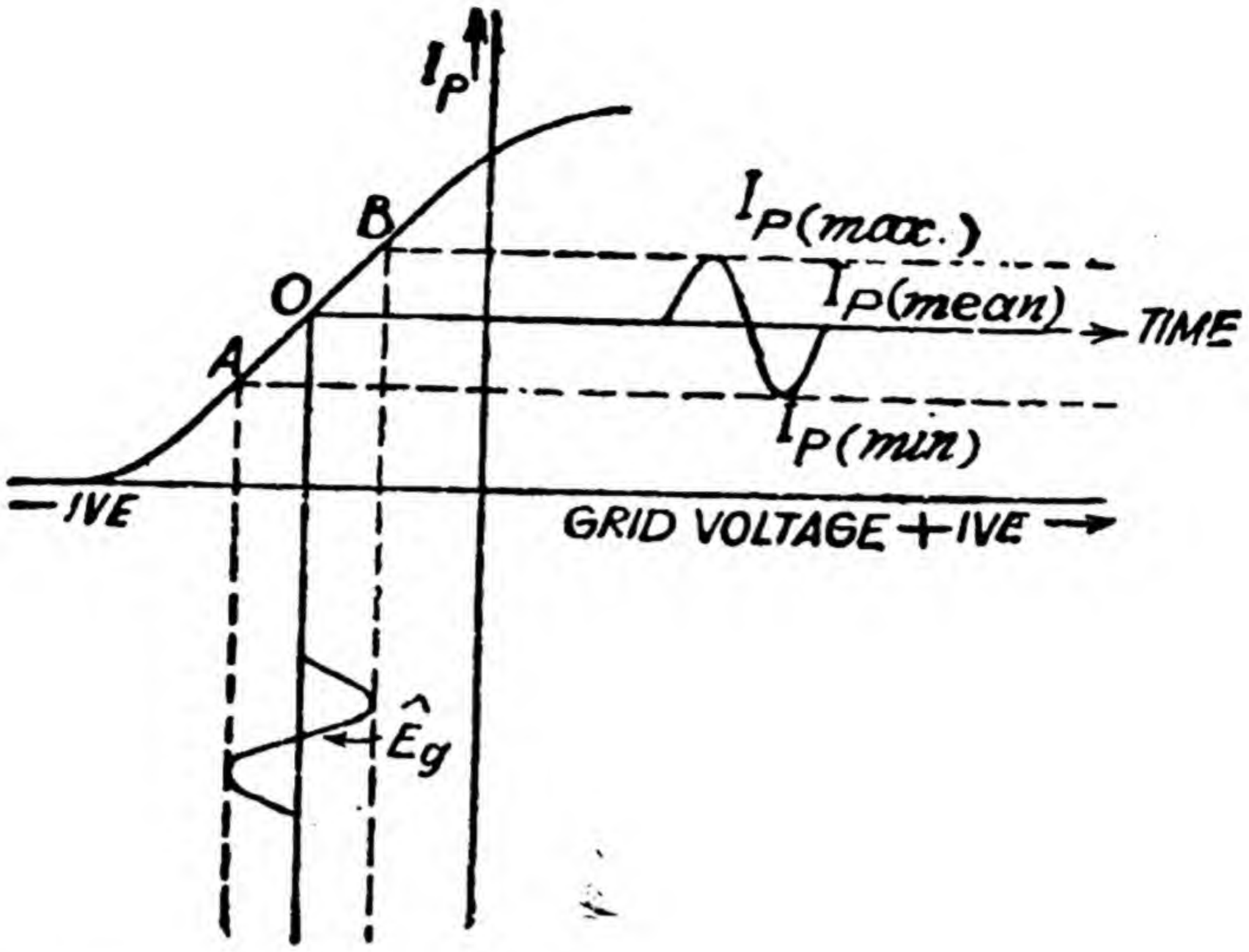


Fig. 28. Showing max., min and mean values of plate current.

From these equations, we have

$$I_p \text{ (mean)} - I_p \text{ (min.)} = a\hat{E}_g - b\hat{E}_g^2.$$

$$I_p \text{ (max.)} - I_p \text{ (min.)} = 2a\hat{E}_g.$$

$$I_p \text{ (max.)} + I_p \text{ (min.)} - 2I_p \text{ (mean)} = 2b\hat{E}_g^2,$$

$$\text{Thus } \frac{I_p \text{ (max.)} + I_p \text{ (min.)} - 2I_p \text{ (mean)}}{2\,[I_p \text{ (max.)} - I_p \text{ (min.)}]}$$

$$= \frac{2b\hat{E}_g^2}{2\ 2a\hat{E}_g} = \frac{b\hat{E}_g}{2a}$$

and the percentage of second harmonic distortion will be

$$= \frac{I_p \text{ (max.)} + I_p \text{ (min.)} - 2I_p \text{ (mean)}}{2\,[I_p \text{ (max)} - I_p \text{ (min)}]} \times 100\%.$$

This holds only when I_p is related to the applied grid voltage by equation (1). This equation is assumed to represent the dynamic characteristic.

2. Frequency distortion : Frequency distortion exists when all frequencies are not amplified by the same amount. This is caused by the circuit associated with the amplifier tube and sets a limit to the range of frequencies that will be satisfactorily amplified in any particular case. The tube may cause frequency distortion if the period of applied signal is of the same order of magnitude as the time of flight of the electron from cathode to the anode

of the tube. Referring to fig. 29, it can be seen that an amplifier having a horizontal gain-frequency plot as shown by the dashed line will be ideal or will have no frequency distortion.

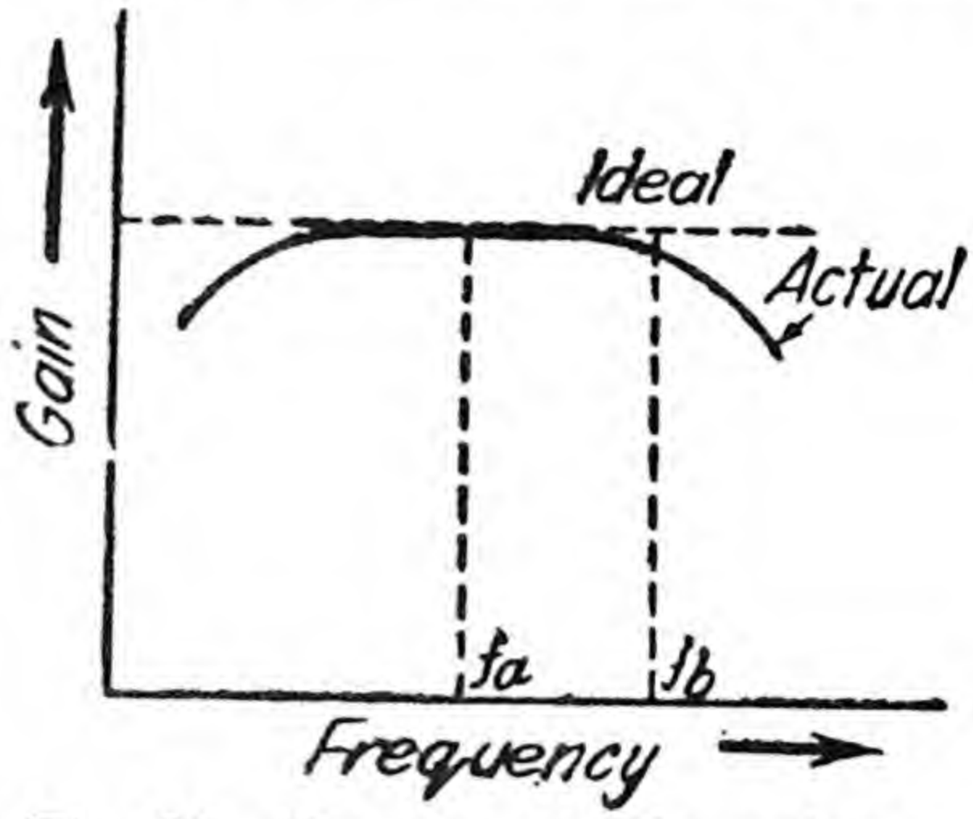


Fig 29. Frequency distortion.

An actual amplifier curve may drop at one or both ends of frequency axis, and the gain for frequencies f_a and f_b will differ. If a complex waveform having both frequencies present as

$$e_g = k_a \sin \omega_a t + k_b \sin \omega_b t$$

is applied, and if the gain for f_a is n and for f_b is $0{\cdot}6n$, then the out-put will be

$$e_0 = nk_a \sin \omega_a t + 0{\cdot}6nk_b \sin \omega_b t$$

Obviously, the output waveform will differ from the input and frequency distortion is present.

3. **Phase distortion or time delay distortion** : If a complex non-sinusoidal signal is to be applied to the input of an amplifier in order to amplify it satisfactorily, the relative phase of the various components must remain unaltered by the process of amplification. If this is not so phase distortion is said to be present in the amplifier. Distortion of this nature will occur in the amplifier if the plot of phase shift against freuqency is not a straight line which passes through zero on the phase shift axis or through a point which is an integral multiple of π radians.

Phase distortion is caused by the presence of reactive and resistive elements in the amplifier circuit. This distortion is not of importance in amplifier of the a.f. type since delayed distortion is not preceptible to the ear. It is very objectionable in systems that depend on waveshape for their operation, as for example, in television.

Suppose that the input signal to the amplifier is periodic and may be expressed analytically by

$$e_g = k_a \sin (\omega t + \theta_1) + k_b \sin (2\omega t + \theta_2) + \ldots$$

If the gain is n, constant in magnitude but produces a phase shift that is proportional to the frequency, the output will be of the form

$$e_0 = nk_a \sin (\omega t + \theta_1 + \psi) + nk_b \sin (2\omega t + \theta_2 + 2\psi) + \ldots$$

This output potential has the same waveshape as the input signal, but a time delay between two wave exists. By writing $\omega t' = \omega t + \psi$, we have

$$e_0 = nk_a \sin (\omega t' + \theta_1) + nk_b \sin (2\omega t' + \theta_2) + \ldots$$

This equation is referred to a new time scale t'. Thus delay in distortion arise due to the frequency character of the circuit associated with the vacuum tube.

A simple example illustrating time delay distortion is shown

in fig. 30 predicting :

(*a*) original wave,

(*b*) original wave after suffering time delay distortion of τ sec

(*c*) original wave after suffering time delay distortion.

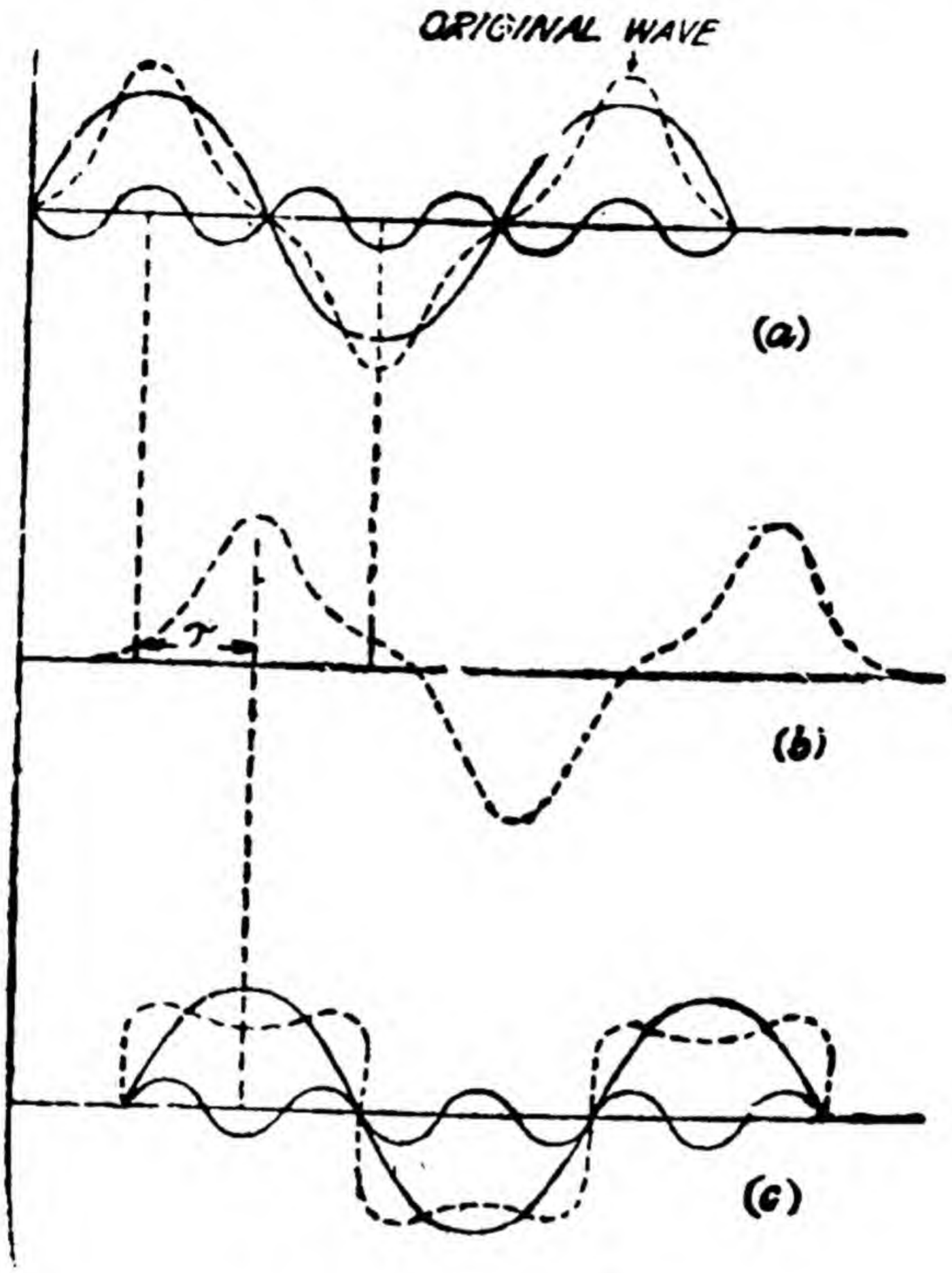


Fig. 30. Phase distortion.

Diagram shows the effect of time delay distortion on a complex wave consisting of a fundamental and third harmonic. In the wave suffering time delay distortion, the relative phase of third harmonic components differ by 18° from its relative phase in the original wave.

9·14. WAVEFORM DISTORTION DUE TO NONLINEARITY OF THE TUBE CHARACTERISTICS :

We have seen that the dynamic characteristic of a vacuum tube is not generally a straight line but it has a slight curvature. This is due to the fact that tube characteristics are not linear, parallel and equidistant for equal increments of the parameter. The effect of this nonlinearity is that the output current waveforms are not exactly same as input waveforms *i.e.*, the nonlinearity of the tube gives rise to distortion. Even for a pure resistive load, the dynamic transfer characteristic is slightly non-linear and hence

results in waveform distortion. It is observed that the amplitude of output voltage or current during positive half cycle is greater than during the negative half cycle. This type of distortion is known as amplitude distortion. The degree of amplitude distortion is measured by the presence of different harmonic components. Harmonic components are the frequency components whose frequencies are integral multiple of the original frequency.

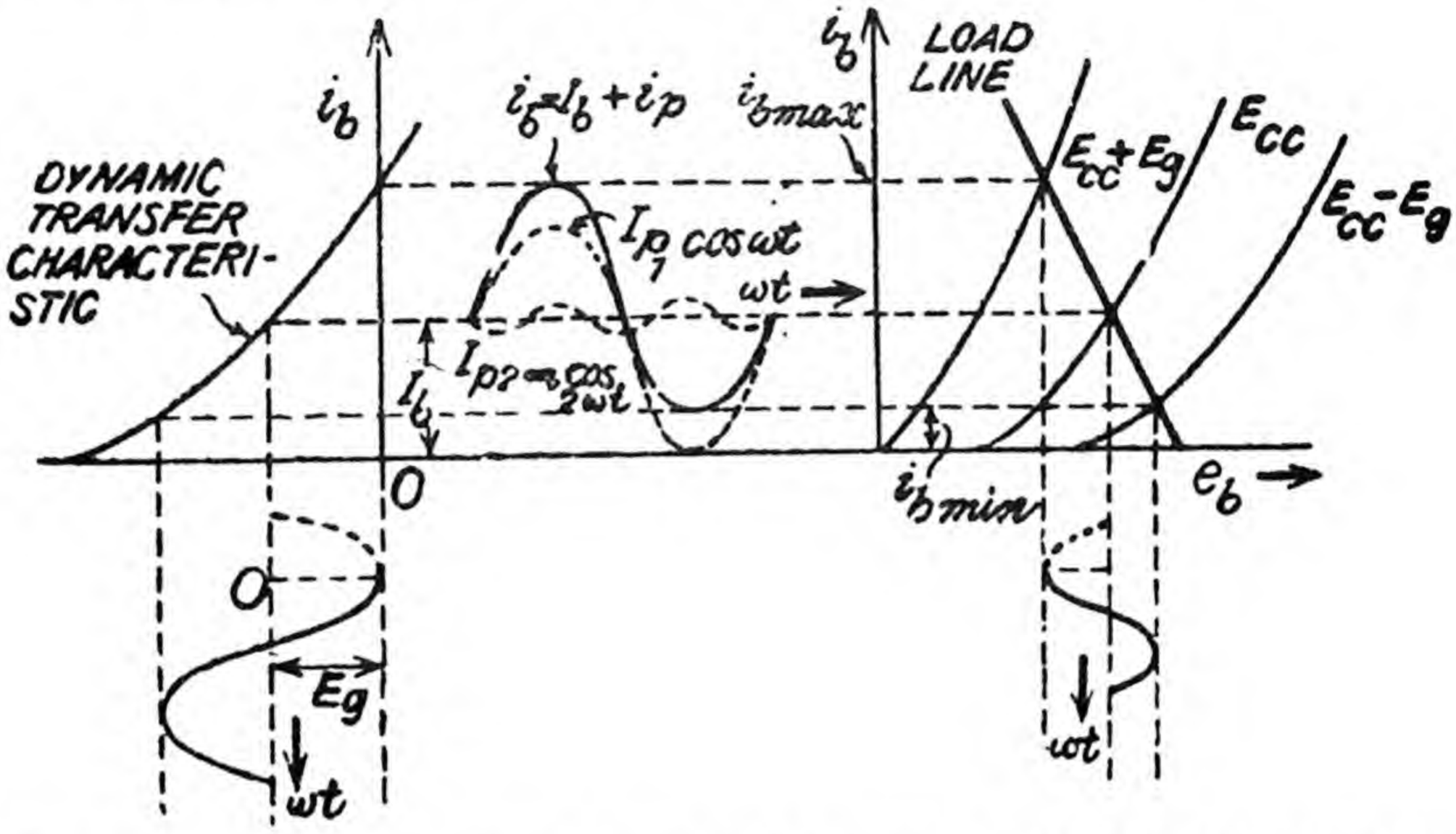


Fig. 31. **Waveform distortion due to non-linearity of triode characteristics.**

Figure (31) shows the type of waveform distortion due to non-linearity of tube characteristics. For simplicity, a pure resistance load is assumed, and grid signal voltage is assumed to be sinusoidal given by

$$e_g = E_g \cos \omega t. \qquad \ldots(1)$$

Let the dynamic characteristics be represented by a power series as

$$i_p = a_1 e_g + a_2 e_g^2 + a_3 e_g^3 + \ldots \qquad \ldots (2)$$

where a_1, a_2, a_3 are constants depending on the tube characteristics and the load.

Substituting the value of e_g from equation (1) in equation (2), we have

$$\begin{aligned} i_p &= a_1 E_g \cos \omega t + a_2 (E_g \cos \omega t)^2 + a_3 (E_g \cos \omega t)^3 + \ldots \\ &= a_1 E_g \cos \omega t + a_2 E_g^2 . \tfrac{1}{2} (1 + \cos 2\omega t) + \ldots \\ &= \tfrac{1}{2} a_2 E_g^2 + a_1 E_g \cos \omega t + \tfrac{1}{2} a_2 E_g^2 \cos 2\omega t + \ldots \end{aligned} \qquad \ldots(3)$$

In equation (3), the term $\frac{1}{2} a_2 E_g^2$ is the d.c. component, $a_1 E_1 \cos \omega t$ is the fundemental frequency term, $\frac{1}{2} a_2 E_g \cos 2\omega t$ is the second harmonic term and so on. Let I_{p0}, I_{p1m}, I_{p2m} represent the constant term amplitude of fundamental harmonic plate current, second harmonic plate current, etc., respectively, then

$$i_p = I_{p0} + I_{p1m} \cos \omega t + I_{p2m} \cos 2\omega t + \ldots \qquad \ldots(4)$$

We know that

$$i_b = I_b + i_p$$

$$\therefore \quad i_b = I_b + I_{p0} + I_{p1m} \cos \omega t + I_{p2m} \cos 2\omega t + \ldots \quad \ldots(5)$$

These four components are shown in the figure (31). These plate characterfstic with load line are also shown. It can be seen from the figure that i_b is maximum at $\omega t = 0$ and minimum at $\omega t = \pi$ radians. Equation (5) may now be written as

$$i_{b\ max.} = I_b + I_{p0} + I_{p1m} + I_{p2m} \quad \ldots(6)$$

$$i_{b\ min.} = I_b + I_{p0} - I_{p1m} + I_{p2m} \quad \ldots(7)$$

From equations (6) and (7), we have

$$I_{p1m} = \frac{i_{b\ max.} - i_{b\ min.}}{2}. \quad \ldots(8)$$

Equation (3) shows that

$$I_{p0} = I_{p1m}$$

$$\therefore \quad I_{p2m} = I_{p3} = \tfrac{1}{2}\left[\frac{i_{b\ max.} + i_{b\ min.}}{2} - I_b\right], \quad \ldots(9)$$

The value of amplitudes of fundamental and second harmonic components of the plate current are given by equations (8) anc equation (9). In order to find out the values of I_{p1m} and I_{p2m}, it is only necessary to know $i_{b\ max.}$, $i_{b\ min.}$ and I_b. These quantities may be found from the plate characteristics graphically. For this purpose we have to draw the load line for a given value of plate supply voltage E_{bb} and load resistance R_L. The value of $i_{b\ max.}$, $i_{b\ min.}$ and I_b are obtained bv noting the plate current i_b for E_{cc}, $(E_{cc} + E_g)$ and $(E_{cc} - E_g)$ as iddicated in figure (31).

$$\frac{I_{p2m}}{I_{p1m}} \times 100 = \frac{\tfrac{1}{2}(i_{b\ max.} + i_{b\ min.}) - I_b}{(i_{b\ max.} - i_{b\ min})} \times 100. \quad \ldots(10)$$

NUMERICAL PROBLEMS

1. *Figure* (32) *shows the circuit of an amplifier. In the amplifier,* $E_{bb} = 300$ *volts,* $R_L = 48$ *KΩ,* $R_k = 2K\Omega$, $R_g = 100$ *KΩ and* $R = 100$ *KΩ. The capacitance* C_c *ond* C_k *are so large that these may be taken as the effective short circuit so far as the a. c. signal frequency is concerned. Draw the d.c. load line and the grid bias line to get the quiscent point. Draw also the a c. load line. The d.c. load line is to be drawn on the characteristics of the tube shown in figure* (33).

The total d.c. resistance in the plate cathode circuit is given by

$$R_{d.c.} = R_L + R_k$$
$$= 48 + 2$$
$$= 50\ K\Omega. \quad \ldots(1)$$

The loop equation for plate cathode circuit is

$$e_b = E_{bb} - i_b R_{d.c.}$$
$$= 300 - i_b\, 50 \quad \ldots(2)$$

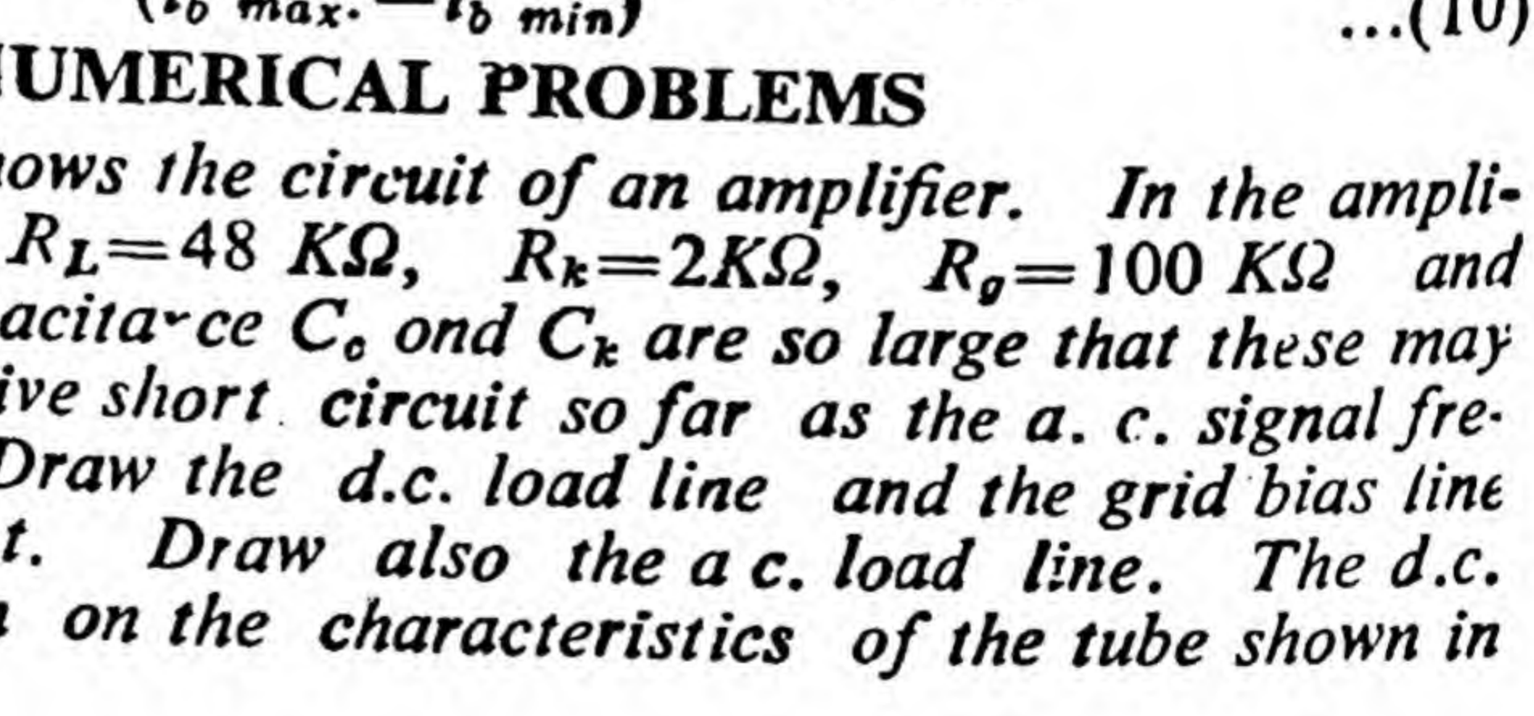


Fig. 32. Circuit of an amplifier.

From equation (2),

$e_b = 0$ when $i_b = 6$ m.a.

$e_b = 300$ when $i_b = 0$

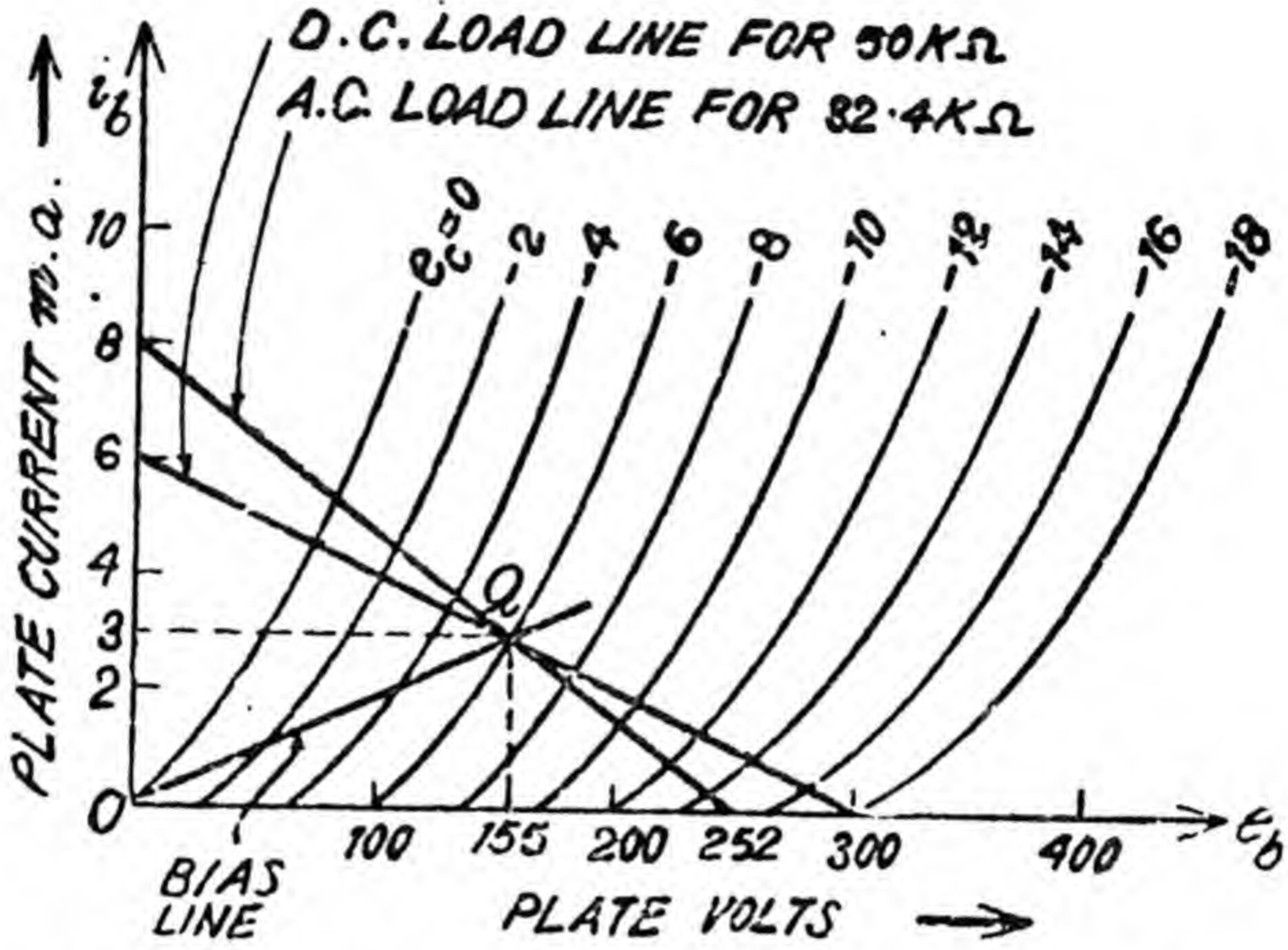


Fig. 33. Plate characteristics of the tube with load lines.

Using these values, a line is drawn on the plate characteristics. This line is the desired load line for 50 $K\Omega$ as shown in figure (33).

Now we draw the grid bias line to determine the operating point Q. Since the grid is maintained negative, the grid current is zero and hence the voltage developed across R is zero. The loop equation is

$$0 = e_c + i_b R_k$$

or $e_c = -i_b$ $\qquad R_k = -i_b.2.$

when $e_c = 0$, $\qquad i_b = 0$

$e_c = -6V$ $\qquad i_b = 3$ m.a.

$e_d = -4V$ $\qquad i_b = 2$ m.a.

The above points are connected on a line which gives the grid bias line. The bias line cuts the load line at a point Q known as the quiescent point or operating point. From the figure we have

$$E_{b0} = 155\ V. I_{b0} = 3 \text{ m.a. and } E_{c0} = -6V.$$

When C_c and C_k are large to be an effective short circuit, the a.c. passes through R_L and R_g. The total a.c. resistance in the plate cathode circuit would be a parallel combination of R_L and R_g. Thus

$$R_{a.c.} = \frac{R_L R_g}{R_L + R_g} = \frac{48 \times 100}{48 + 100} = 32{\cdot}4\ K\Omega.$$

The a.c. load line is a line passing through Q and having a slope given by $\theta = \tan^{-1}(-1/R_{a.c.})$. In order to draw the a.c. load line, we require a second point which is obtained as follows :

When $i_b = I_{b_1} = 3$ m,a., then the voltage drop across $R_{a.c.}$ is $R_{a.c.} \times i_b = 32{\cdot}4 \times 3 = 97{\cdot}2$ volts. The intercept of the a.c. load line on e_b axis is $E_{bn} + I_{bn} R_{a.c.} = 155 + 97{\cdot}2 = 252{\cdot}2$ *V*. Now a line is drawn connecting $e_b = 252{\cdot}2$ *V* to the quiescent point *Q*. This line is a.c. load line.

2. *A linear A_1 amplifier uses a triode has voltage gain -10 and -12 with load resistance of 8 and 12 KΩ respectively. Calculate the amplification factor and a.c. plate resistance of the tube.*

For resistive load, the voltage gain is given by

$$A = -\frac{\mu R_L}{r_p + R_L}.$$

In the first case $$-10 = -\frac{\mu \times 8}{(r_p + 8)}. \quad \text{.. (1)}$$

In the second case $$-12 = -\frac{\mu \times 12}{(r_p + 12)} \quad \text{...(2)}$$

By multiplying equation (1) by 3 and equation (2) by 2 and rearranging, we get

$$30\,(r_p + 8) = 24\,\mu, \quad \text{...(3)}$$

$$24\,(r_p + 12) = 24\,\mu. \quad \text{...(4)}$$

From equations (3) and (4), we have

$$30\,r_p + 240 = 24\,r_p + 288,$$

$$6r_p = 288 - 240 = 48$$

$$r_p = 8\ K\Omega.$$

Substituting the value of r_p in equation (1) and solving, we get

$$\mu = 20.$$

3. *An amplifier on using a valve V_1 having amplification factor of 20 and plate resistance of 25 kilo-ohms has the same gain as on using another valve V_2 having amplification factor of 10 and plate resistance 6 kilo-ohms. Find the value of load resistance if such a condition is to be obtained.*

We know that $$A = -\frac{\mu R_L}{r_p + R_L}.$$

For valve V_1, $$A = \frac{-20 \times R_L}{25 \times 10^3 + R_L}$$

For valve V_2, $$A = \frac{-10 R_L}{6 \times 10^3 + R_L}$$

Hence $$\frac{20\ R_L}{25 \times 10^3 + R_L} = \frac{10\ R_L}{6 \times 10^3 + R_L}$$

or $$2 \times 6 \times 10^3 + 2R_L = 25 \times 10^3 + R_L$$

or $$R_L = 13 \times 10^3 = 13\ K\Omega.$$

4. *A linear A_1 amplifier uses a triode having amplification factar μ of 10 and plate resistance r_p of 6000Ω. The load in the plate circuit consists of a resistance of 10kΩ, in series with an inductor of 1·5 henry. If the input is a sinusoidal voltage of effective value 2 volts and frequency 1000 Hz, find the amplitudes and phase angles of the alternating plate current, plate voltage and complex voltage gain.*

The effective value of a.c. plate current is given by

$$\mathbf{I}_p=\frac{\mu \mathbf{E}_g}{r_p+Z_L}=\frac{\mu \mathbf{E}_g}{(r_p+R_L)+j\omega L}.$$

The magnitude of the a.c. plate current is given by

$$I_p=\frac{\mu E_g}{\sqrt{\{(r_p+R_L)^2+\omega^2L^2\}}}$$

$$=\frac{10\times 2}{\sqrt{\{6\times 10^3+10\times 10^3)^2+(2\pi\times 1000+1\cdot 5)^2\}}} \text{ amp.}$$

$$=1\cdot 07 \text{ milliamperes.}$$

The phase angle is

$$=-\tan^{-1}\frac{\omega L}{r_p+R_L}$$

$$=-\tan^{-1}\frac{2\pi\times 1000\times 1\cdot 5}{(6000\times 10{,}000)}=-30^\circ\ 30'.$$

R.M.S. value of the plate voltage is given by

$$\mathbf{E}_p=-\mathbf{I}_p\mathbf{Z}_L=-\frac{\mu \mathbf{E}_g.\mathbf{Z}_L}{r_p+Z_L}.$$

Magnitude

$$=\frac{\mu \mathbf{E}_g\sqrt{(R_L^2+\omega^2L^2)}}{\sqrt{\{(r_p+R_L)^2+\omega^2L^2\}}}$$

$$=I_p\sqrt{(R_L^2+\omega^2L^2)}$$

$$=(1\cdot 07\times\ 0^{-3})\sqrt{\{(10\times 10^{3\ 2}+(2\pi\times 1000\times 1\cdot 5)^2\}} \text{ volts}$$

$$=14\cdot 8 \text{ volts.}$$

$$\text{Phase angle of } \mathbf{E}_p=180^\circ+\tan^{-1}\frac{\omega L}{R_L}-\tan^{-1}\frac{\omega L}{r_p+R_L}$$

$$=180^\circ+\tan^{-1}\frac{2\pi\times 1000\times 1\cdot 5}{10\times 10^3}-30^\circ\ 30'$$

$$=180^\circ+43^\circ\ 10'-30^\circ\ 30'=192^\circ\ 40'.$$

The complex voltage gain is given by

$$\mathbf{A}=\frac{\mathbf{E}_p}{\mathbf{E}_g}=-\frac{\mu Z_L}{r_p+Z_L}$$

$$=\frac{-\mu\ (R_L+j\omega L)}{(r_p+R_L)+j\omega L}$$

The magnitude is

$$=\frac{\mathbf{E}_p}{\mathbf{E}_g}=\frac{14\cdot 8}{2}=7\cdot 4$$

5. *The triode circuit shown in figure 34 (a) called a cascade amplifier is used in television receiver circuit as a radio frequency amplifier. Draw its equivalent circui and determine the voltage gain.*

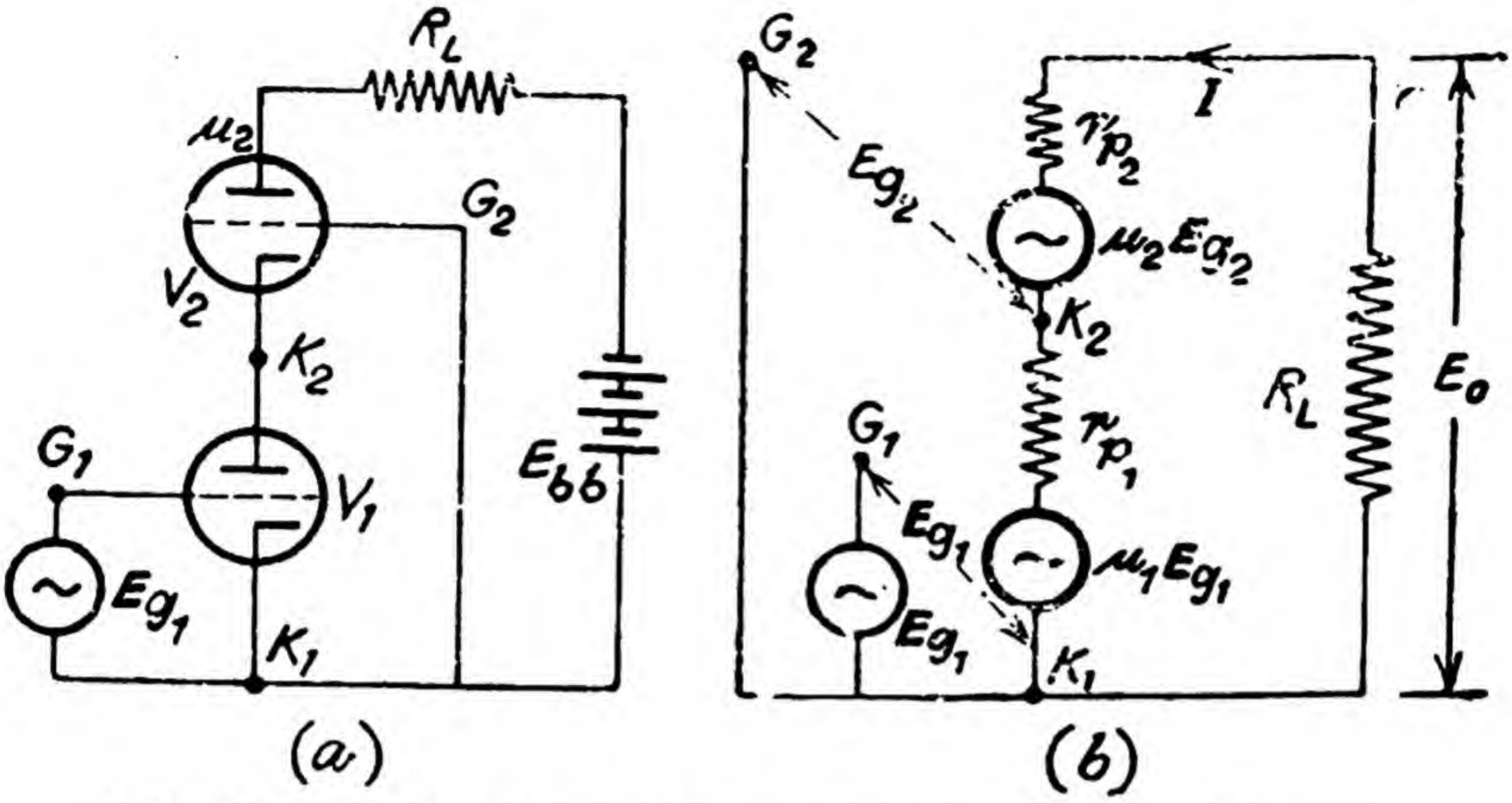


Fig. 34. **A cascade amplifier and its equivalent circuit.**

It can be noted that

$$0 = E_{g2} \quad I\, r_{p1} - \mu_1 E_{g1}$$

or

$$E_{g2} = \mu_1 E_{g1} + I r_{p1} \qquad \ldots(1)$$

Let I be the loop current, than loop equation may be written as

$$\mu_1 E_{g1} + \mu_2 E_{g2} = -I\,(r_{p1} + r_{p2} + R_L). \qquad \ldots(2)$$

Substituting the value of E_{g2} from equation (1) in equation (2), we get

$$\mu_1 E_{g1} + \mu_2\,(\mu_1 E_{g1} + I r_{p1}) = -I\,(r_{p1} + r_{p2} + R_L)$$

or

$$\mu_1\,(1 + \mu_2)\,E_{g1} = -I\,(r_{p1} + \mu_2 r_{p1} + r_{p2} + R_L)$$

or

$$I = \frac{-\mu_1\,(1 + \mu_2)\,E_{g1}}{r_{p1}\,(1 + \mu_2) + r_{p2} + R_L}.$$

The output voltage E_0 is given by

$$E_0 = IR_L = \frac{-\mu_1\,(1 + \mu_2)\,E_{g1}\,R_L}{r_{p1}\,(1 + \mu_2) + r_{p2} + R_L}$$

$$\therefore \quad A = \frac{E_0}{E_{g1}} = \frac{-\mu_1\,(1 + \mu_2)\,R_L}{r_{p1}\,(1 + \mu_2) + r_{p2} + R_L}.$$

6. *A linear A_1 amplifier uses a triode with load resistance of* 24,000 *ohms and a cathode bias circuit consisting of a parallel combination of R_k and C_k. The value of R_k is so chosen that with a plate supply voltage of* 352 *volts, the zero signal plate current is* 10 *m A. and zero signal grid bias is* −12 *volts. The amplification factor of the tube is* 20 *and dynamic plate resistance is* 12,000 *ohms. Condenser C_k is* 10 *micro-farads. Calculate (i) value of R_k (ii) zero signal plate voltage, and (iii) magnitude of voltage gain at frequency* 1000 *c/s.*

(*i*) We know that grid bias $= -R_k I_b$.

$$\therefore \quad R_k I_b = -12$$

or
$$R_k = \frac{12}{10 \times 10^{-3}} = 1200 \text{ ohms.}$$

(*ii*) Zero signal plate voltage E_b is given by

$$E_b = I_{bb} - I_b (R_L + R_k)$$
$$= 352 - 10 \times 10^{-3} (24{,}000 + 1200)$$
$$= 110 \text{ volts.}$$

(*iii*) The complex voltage gain A is given

$$\mathbf{A} = \frac{-\mu R_L}{r_p + R_L + (\mu + 1 \quad Z_k}$$

$$= \frac{-\mu R_L / Z_k}{\frac{r_\rho + R_L}{Z_k} + (\mu + 1)}$$

$\mathbf{Z}_k$ is given by $\quad \frac{1}{\mathbf{Z}_k} = \frac{1}{R_k} + j\omega C_k = \frac{1 + j\omega C_k R_k}{R_k}$

At frequency of 1000 c/s

$$\frac{1}{Z_k} = \frac{1 + j.2\pi \times 1000 \times 10^6 \times 1200}{1200}$$

$$= \frac{1 + j \,.\, 75{\cdot}36}{1200}.$$

$$\therefore \quad \mathbf{A} = \frac{-20 \times 24000 \times (1 + j\ 75{\cdot}36)/1200}{\frac{36.0\ \ 0\ (1 + j\ 75{\cdot}36)}{1200} + (20 + 1)}$$

$$\mathbf{A} = \frac{400\ (1 + j\ 75\ 36)}{30\ (1 + j\ 75{\cdot}36) + 21} = \frac{400 + j30144}{51 + j\ 2261}$$

Magnitude

$$\mathbf{A} = \frac{\sqrt{\{(400)^2 + (30144)^2\}}}{\sqrt{\{(51^{\ 2} + (2.61)^2\}}} = 21.$$

EXERCISES AND PROBLEMS

1. What is the purpose of a voltage amplifier ?
2. Explain in detail why a fixed negative bias is necessary to make most tubes operate properly as amplifiers.
3. Draw a circuit diagram of basic voltage amplifier and show the different voltages at different places.
4. How a vacuum tube amplifier works as a linear device ?
5. Explain the working of vacuum tube amplifier under zero signal (quiescent operation) condition and with a c. signal
6. How will you draw the dynamic characteristic of a valve ?
7. Explain the equivalent circuit representation (voltage and current) for triode amplifier.

8. Explain, giving basic circuit diagram, the working of a triode as an amplifier.
   If a resistance R_L is inserted in the anode circuit as load resistance, prove that the voltage amplification obtained is $\mu R_L/(R_L+r_p)$ where μ is amplification factor and r_p is the plate resistance of the triode.
9. Give a brief account of the classification of amplifiers.
10. What do you mean by interelectrode capacitances in a triode.
11. Calculate the input admittance of a triode amplifier in case of a complex load.
12. Describe the different methods used for biasing a control grid.
13. What do you mean by distortion in amplifiers. Give a brief discussion of (i) amplitude distortion, (ii) frequency distortion and (iii) phase distortion ?
14. How the nonlinearity of tube characteristics affects the waveform. Give the expression for the percentage of second harmonic in a vacuum triode.
15. A voltage amplifiers uses a triode having amplification factor 20 and plate resistance 40,000 ohms. Another tube with plate resistance of 60,000 ohms and amplification factor 15 is replaced for the previous tube. The amplification gain remains unaltered. Determine the load resistance to satisfy the condition.
16. The gain of a triode amplifier with resistive load of 5K ohm is 8, while with resistive load of 10K ohm is 10. Determine the amplification factor and plate resistance of the triode.
17. The a.c. plate resistance of a tride is 6000 ohms whose amplification factor is 8. In the plate circuit the load resistance is 12,000 ohms. Find the value of a.c. voltage to be applied between grid and cathode to get a.c. power output of 2 watts across the load.
18. A type 6J5 triode is operated as a single amplifier under specific conditions at 22,000 cps. The important factors are $\mu=20$, $r_p=7{,}700$ ohms, $C_{gp}=3{\cdot}4\mu\mu f$, $C_{gk}=3{\cdot}4\ \mu\mu f$, and $C_{pk}=3{\cdot}6\ \mu\mu f$. Calculate the input capacitance and the input resistance of the tube when the load is a resistor $R=20K$ ohm.
19. A class A, amplifier uses a high μ triode having $\mu=80$. $r_p=50K$ ohm, $C_{gp}=C_{gk}=2\mu\mu f$. A signal of frequency $(80{,}000/2\pi)$ Hz is applied at the input. Calculate the input shunt conductance, input shunt capacitance and input shunt admittance of the amplifier if the load including C_{pk} in shunt is (a) resistance of $200K$ ohms. (b) a pure inductive reactance of $200K$ ohm and (c) complex inductive impedance of 10^5+j10^6 ohm.
20. A linear A_1 amplifier uses a triode having $\mu=20$ and $r_p=10{,}000$ ohms. Load resistance is 14000 ohms. Calculate bias impedance consisting of parallel combination of resistance $R_k=500$ ohms and condenser C_k of 16 microfarad. If the zero signal bias developed across R_k is —10 volts and zero signal plate voltage is 120 volts, find the zero signal plate current and plate supply voltage. Also find the magnitude and phase angle of complex voltage gain at a frequency 1000 Hz. Find also the magnitude of voltage gain which occurs in the absence of cathode impedance.

10 UNTUNED POTENTIAL AMPLIFIERS

We have seen that when a sinusoidal voltage is applied between the grid and cathode of a vacuum tube amplifier, the output voltage at the plate circuit is several times larger than the input voltage. The waveform and frequency of the output voltage is the same as that of the input. Thus the input voltage is said to be amplified several times by the amplifier. When the grid of the vacuum tube is maintained at a negative potential even for the positive half cycle of the applied a.c signal, no grid current flows *i.e.*, no a.c. power is consumed in the grid circuit. The a c. signal at the grid without any loss can control an appreciably large current flowing through a large impedance in the plater circuit. Thus the amplitude of the output voltage is far larger than that of the input voltage. This gives a voltage gain of the amplifier. It can also be observed that the power gain will be still larger. A transformer may be used to step up the voltage but the power across the secondary terminals is always less than across the primary. Sometimes the transformers are used as loads in vacuum tube amplifiers.

The voltage gain of a single stage is less than the amplification factor of the tube used. Sometimes, it is required to have a voltage gain far exceeding the amplification factor of the tube. In such cases, a number of amplifiers are connected in such a way that the output of one is connected to the input of the next. Such an amplifier is known as cascade amplifier. Depending upon the network used for the coupling between output stage and input stage, the cascade amplifiers may classified as :

(1) resistance capacitance coupled amplifiers,
(2) inductance capacitance coupled amplifiers,
(3) transformer coupled amplifiers, and
(4) direct coupled amplifiers.

10·1. RESISTANCE CAPACITANCE COUPLED AMPLIFIER :

One of the most common voltage amplifier circuits. used when a sensibly constant amplification over a wide range of frequencies is desired, is the resistance capacitance coupled circuit. If the output voltage of the first tube is coupled to the grid of the next tube

through a resistance and capacitance, the amplifier is known as the resistance capacitance coupled amplifier. A two stage voltage amplifier employing triodes is shown in figure 1.

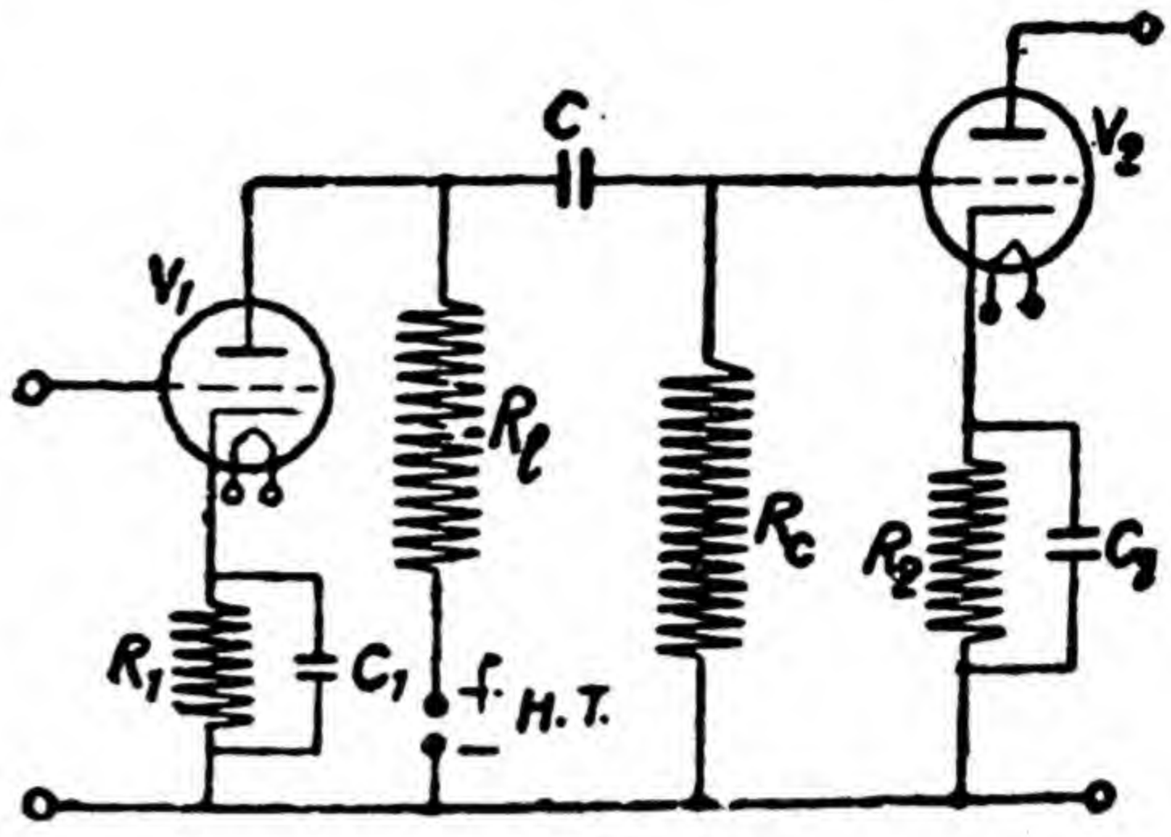


Fig. 1. R.C. Coupled Amplifier.

Function of circuit components :

Cathode bias circuit : Here the first stage of the triode circuit is the same as the basic amplifier circuit except that bias battery has been replaced by a self biased resistor R_1 in the cathode circuit. The resistor R_1 is shunted by the capacitor C_1. Components R_1 and C_1 provide the bias for the tube. The combination is known as self-bias or, more specially, a cathode bias circuit. In any voltage amplifier, freedom from amplitude distortion is obtained by operating the tube with the grid at some negative value compared to cathode which is done with the help of resistor R_1. Because of the voltage drop in resistor R_1 cathode is positive with respect to ground ; using positive cathode as reference, the grid is in turn negative compared to cathode.

Why is the resistor shunted by the capacitor C_1 ? When a.c. signal are applied at the input of this amplifier, the instantaneous grid voltage will change, and the plate current will vary accordingly. But if the plate current changes, the bias (drop across R_1) would also change, and this effect is generally not desired. Capacitor C_1 serves as a by-pass capacitor to prevent these bias variations.

Plate circuit resistance (R_l) : In order to achieve amplification, a load impedance must be connected between the plate of the tube and the plate power supply. Any variation in the plate current of V_1 and the coupling resistor R_1 will produce corresponding variations in the voltage across R_l because plate current flowing through this resistor will produce a voltage drop.

Effect of C and R_e Coupling : The blocking capacitor C prevents the high voltage that is applied to the plate of valve V_1 from reaching the grid of valve V_2. A high positive voltage at the grid of a tube would cause a very high current to flow in the grid and plate circuits of the tube ; this would overload and quickly damage the tube. If this capacitor were used without resistor R_e, tube V_2 would be blocked because the negative charges (the electrons that may be collected by the grid from the electron stream within the tube) on the grid side of the capacitor would increase the bias sufficiently to cause the tube to operate beyond cut-off. To prevent this blocking action a high resistance R_e called the grid leak is connected between the grid and cathode to provide a direct path for the accumulated electrons to leak off.

Therefore we conclude :

(*i*) R_o acts as a shunt across the load resistance, thus reducing effective load impedance, but this can be neglected if R_o is much greater than R_l.

(*ii*) More important effect is that the circuit combination R_o, C acts as a potential divider across the amplified voltage.

Ordinary range of capacitor C varies from ·001 to 0·1 μf for conventional a.f. stages. The bypass capacitor C_1 has a large capacity, of the order of 25 or 50 micro farad so that it may effectively bypass a.c. component of plate current. The value of R_1 may vary from 500 ohms to 5000 ohms. The typical value of R_o ranges from $50K\Omega$ to $20M\Omega$. The load resistor R_l is determined principally by the gain and the frequency band-width that are desired.

To analyse R-C coupled amplifier, we shall study it in the following frequency ranges :

(A) **Low frequency range :** ·At low frequencies the effect of inter-electrode capacitance can be neglected. The equivalent circuit is shown in figure 2.

According to Thevenin's theorem, any two terminal linear network containing linear impedances and generators can be replaced with an equivalent circuit consisting of a voltage source E' in series with an impedance Z'. The value of E' is the open circuit voltage between the terminals of the network and Z' is the impedance measured between the terminals with all other generators being removed (but not their internal impedances).

The Thevenin's equivalent of figure 2 is shown in figure 3.

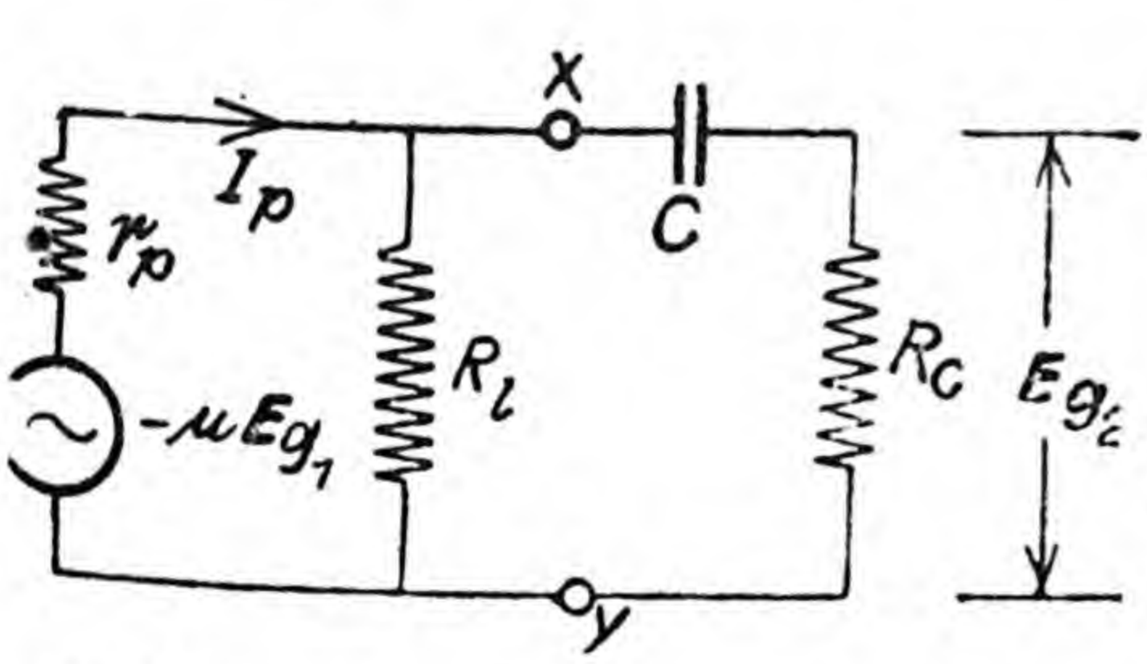


Fig. 2. a.c. equivalent circuit of R.C. coupled amplifier at low frequency range.

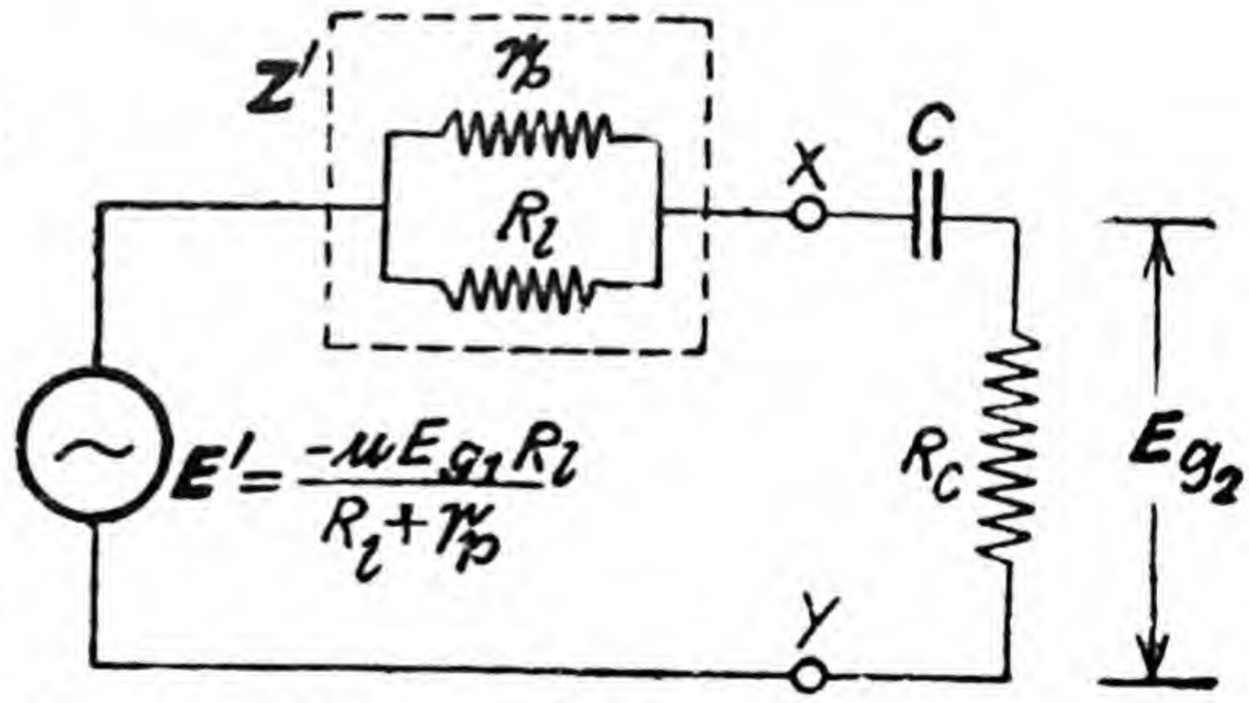


Fig. 3. Thevenin's equivalent of figure 2.

In figure 3, E′ is the open circuit voltage across the resistance R_l in figure 2.

Voltage across R_l = resistance × current

$$= R_l \times \mathbf{I}_p$$

$$= R_l \times \left(\frac{-\mu E_{g1}}{R_l + r_p} \right)$$

$$= -\frac{\mu E g_1 R_l}{R_l + r_p}.$$

because the voltage in the circuit is $-\mu E g_1$ and total resistance is $(R_l + r_p)$.

The impedance Z' between the two terminals X and Y is the impedance of the parallel combination of R_l and r_p as shown in figure 3.

$$\therefore \quad \frac{1}{Z'} = \frac{1}{r_p} + \frac{1}{R_l} = \frac{r_p + R_l}{r_p R_l}$$

or
$$Z' = \frac{r_p R_l}{R_l + r_p}.$$

The total impedance of Thevenin's equivalent circuit of figure 3 is given by

$$\frac{1}{j\omega C} + \frac{r_p R_l}{r_p + R_l} + R_c.$$

Now the current **I** in Thevenin's equivalent circuit is given by

$$\mathbf{I} = \frac{\dfrac{-\mu E g_1 R_l}{R_l + r_p}}{\dfrac{1}{j\omega C} + \dfrac{r_p R_l}{r_p + R_l} + R_o} \qquad \ldots(1)$$

The voltage $E g_2$ across R_c will be

$$E g_2 = \frac{\dfrac{-\mu E g_1 R_l}{R_l + r_p}}{\dfrac{1}{j\omega C} + \dfrac{r_p R_l}{R_p + R_l} + R_o} \times R_o \qquad \ldots(2)$$

The complex voltage gain in low frequency range is given by

$$A_l = \frac{E g_2}{E g_1} = \frac{\dfrac{-\mu R_l}{R_l + r_p}}{\dfrac{1}{j\omega C} + \dfrac{r_p R_l}{r_n + R_l} + R_o} \times R_o$$

$$= \frac{\dfrac{-\mu R_l R_o}{R_l + r_p}}{\dfrac{(r_p + R_l) + j\omega C (r_p R_l) + R_c\, j\omega C.(r_p + R_l)}{j\omega C (r_p + R_l)}}$$

$$= \frac{-\mu R_l R_o}{\dfrac{1}{j\omega C}(r_p + R_l) + r_p R_l + R_o r_p + R_c R_l} \qquad \ldots(3)$$

This expressiou shows that voltage amplification depends upon frequency *i.e.* increases with frequency.

(B) Intermediate frequency range : For intermediate frequen-

cies at which $R_c \gg \frac{1}{\omega C}$, the equivalent circuit may be simplified by the omission of C. Under such conditions the load for the valve becomes parallel combination of R_l and R_c.

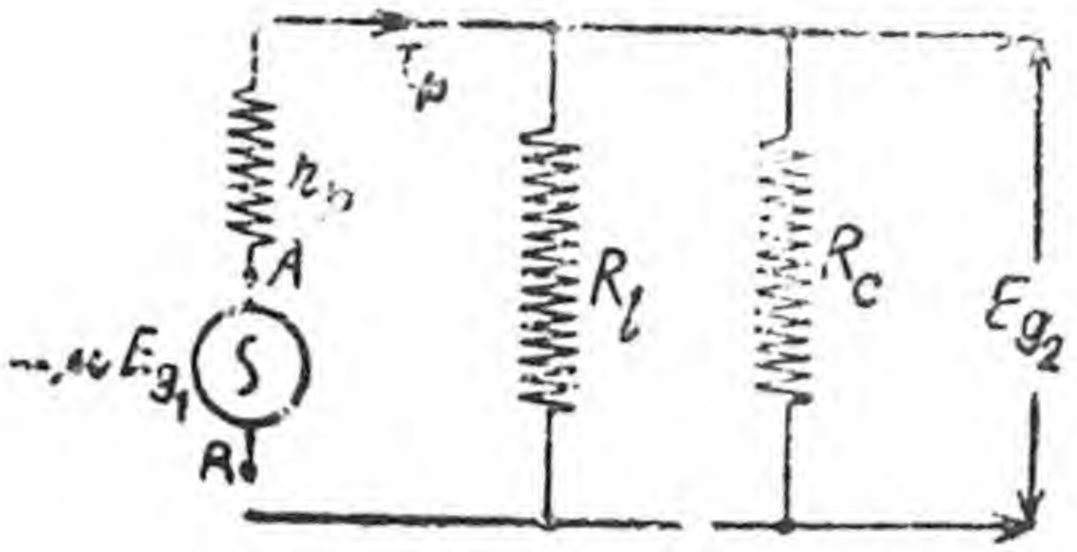


Fig. 4. Equivalent circuit at mid-frequency range.

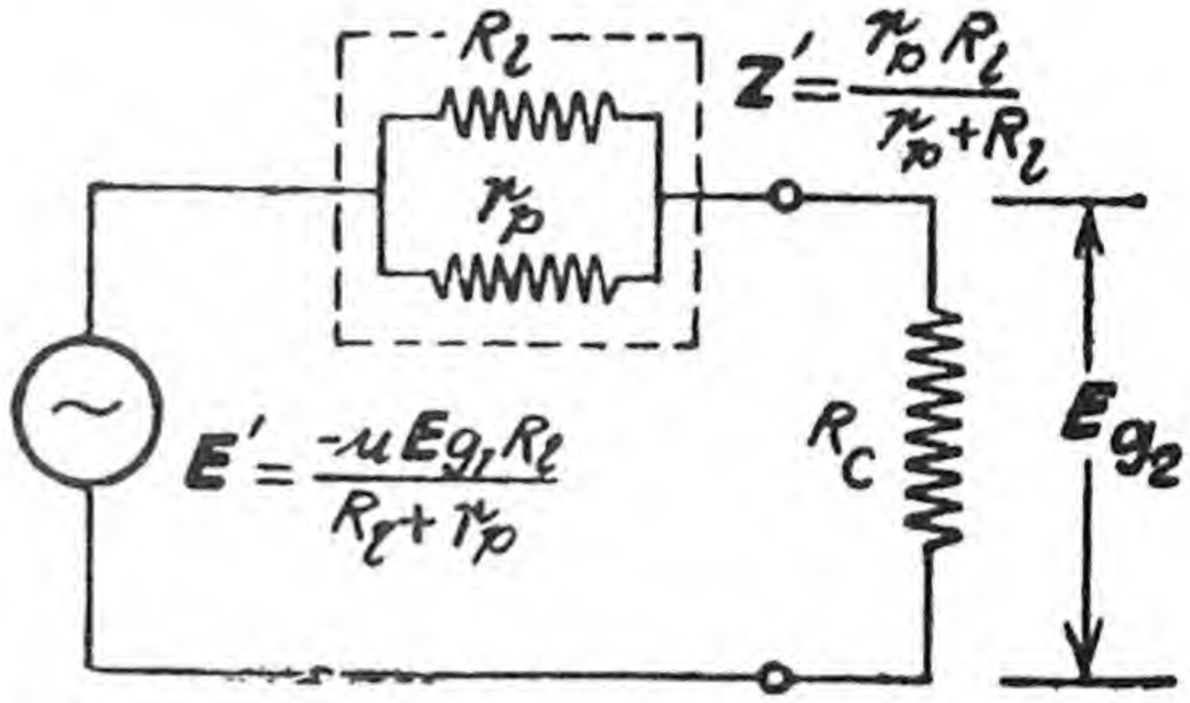


Fig. 5. Thevenin's equivalent circuit at mid frequency range.

The equivalent a.c. circuit and Thevenin's equivalent circuit are shown in figures 4 and 5 respectively. The voltage amplification in this case can be calculated as follows :

The total impedance of the Thevenin's equivalent circuit is

$$Z = R_c + \frac{r_p R_l}{r_p + R_l}$$

$$= \frac{R_c(r_p + R_l) + r_p R_l}{(r_p + R_l)}$$

The current **I** in the circuit is

$$\mathbf{I} = \frac{\frac{-\mu \mathbf{E}_{g_1} R_l}{r_p + R_l}}{\frac{R_c(r_p + R_l) + r_p R_l}{(r_p + R_l)}} = \frac{-\mu \mathbf{E}_{g_1} R_l}{R_c(r_p + R_l) + r_p R_l} \quad \ldots(4)$$

Voltage $\mathbf{E}_{g_2}$ developed across R_c is given by

$$\mathbf{E}_{g_2} = \text{current } \mathbf{I} \times R_c = -\frac{\mu \mathbf{E}_{g_1} R_l R_c}{R_c(r_p + R_l) + r_p R_l} \quad \ldots(5)$$

Voltage amplification $\mathbf{A}_m = \frac{\mathbf{E}_{g_2}}{\mathbf{E}_{g_1}} = \frac{-\mu R_l R_c}{R_c(r_p + R_l) + r_p R_l}$ ...(6)

From equation (6) we infer that when load is resistive, the amplification of the stage will be independent of frequency for frequencies beyond the value $R_c \gg \frac{1}{\omega C}$.

(C) High frequency range : At low frequencies, the effect of interelectrode capacitance can be neglected, but at very high frequencies the interelectrode capacitance becomes important and therefore its effect should be taken into account. The effect is to reduce the voltage amplification.

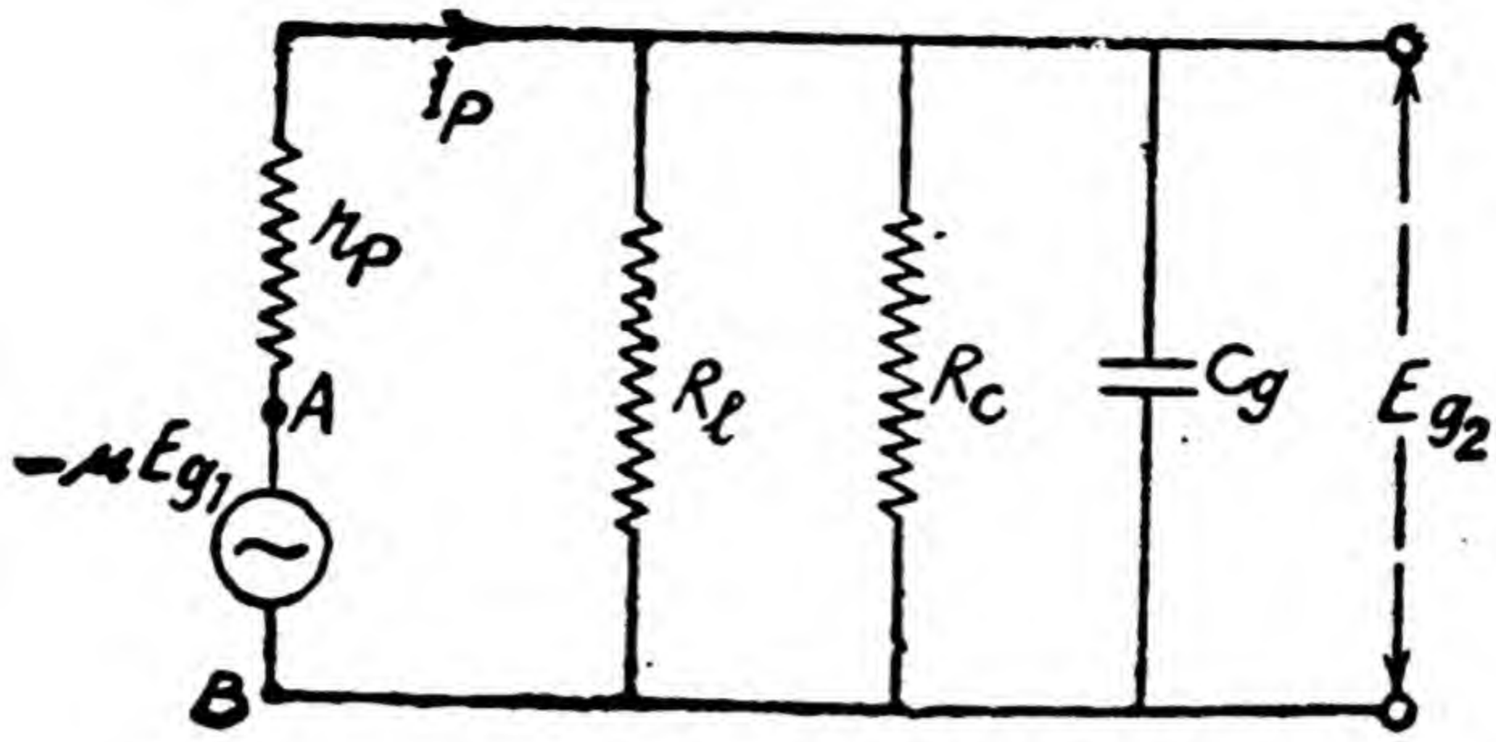


Fig. 6. Equivalent circuit.

The equivalent circuit for this frequency range is shown in figure 6. In the figure C_g denotes the effect of interelectrode capacitance at high frequencies. The Thevenin's equivalent circuit is shown in figure 7.

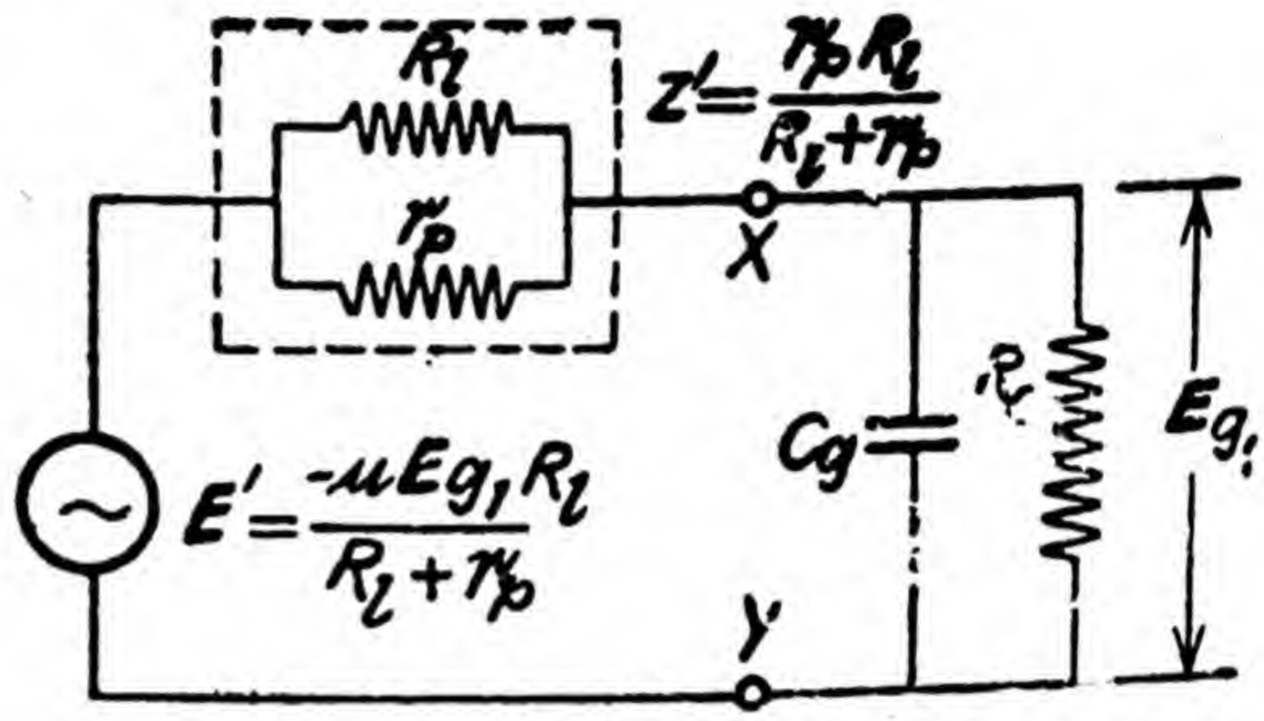


Fig. 7. Thevenin's equivalent circuit at high frequency range.

The total resistance of the circuit in this frequency range is given by

$$\frac{r_p R_l}{R_l + R_p} + \frac{R_c}{1 + j\omega C_g R_c},$$

where $\frac{R_c}{1 + j\omega C_g R_c}$ is impedance of the parallel combination of C_g and R_c.

$$\therefore \quad \text{Total resistance} = \frac{r_p R_l\,(1 + j\omega\, C_g R_c) + R_c\,(r_p + R_l)}{(r_p + R_l)\,.\,1 + j\omega\, C_g R_c)}$$

The current I in the circuit is

$$\mathbf{I} = \frac{-\dfrac{\mu \mathbf{E}_{g_1} R_l}{(R_l + r_p)}}{\dfrac{r_p R_l\,(1 + j\omega\, C_g R_c) + R_c\,(r_p + R_l)}{(r_p + R_l)\,(1 + j\omega\, C_g R_c)}} \qquad \ldots(7)$$

The voltage E_{g_2} across XY is given by

$$\mathbf{E}_{g_2} = \frac{-\dfrac{\mu \mathbf{E}_{g_1} R_l}{R_l + r_p} \times \dfrac{R_c}{1 + R_c . j\omega C_g}}{\dfrac{r_p R_l\,(1 + j\omega C_g R_c) + R_c\,(r_p + R_l)}{(r_p + R_l)\,(1 + j\omega\, C_g R_c)}} \qquad \ldots(8)$$

Complex voltage amplification $\mathbf{A}_h$ is given by

$$\mathbf{A}_h = \frac{\mathbf{E}_{g_2}}{\mathbf{E}_{g_1}} = \frac{-\mu R_l R_c}{r_p R_l + j\omega C_g R_c r_p R_l + R_c r_p + R_c R_l}$$

$$= \frac{-\mu R_l R_c}{r_p R_l + R_c r_p + R_c R_l\,(1 + j\omega C_g r_p)} \qquad \ldots(9)$$

Advantages : (*i*) RC coupled amplifier has the advantage of uniform response over wide range of frequencies,

(*ii*) cheap and easy to design, and

(*iii*) no distortion in the output.

Disadvantages : (*i*) These are not efficient, *i.e* higher voltage is drawn from sup ly source to give sufficient voltage to plate because most of the voltage drops across the load.

(*ii*) It is not suited to higher frequency voltage amplification as interelectrode capacitance becomes active at higher frequencies.

Frequency response curve : From the construction of equivalent circuits it is clear that the expressions for the voltage amplification in the intermediate frequency range does not depend upon frequency, or we say impedance is pure resistance which remains constant at all frequencies so that voltage amplification in this region is constant.

At low frequencies, capacitive resistance in the series circuit is high so that voltage reaching the second stage is less because it is equal to the voltage drop across R_l minus voltage drop across capacitive reactance. Thus the low frequency voltage that would otherwise be developed across the grid leak is reduced. This causes a decrease in voltage amplification at low frequencies as is indicated by the bend in the curve of figure 8.

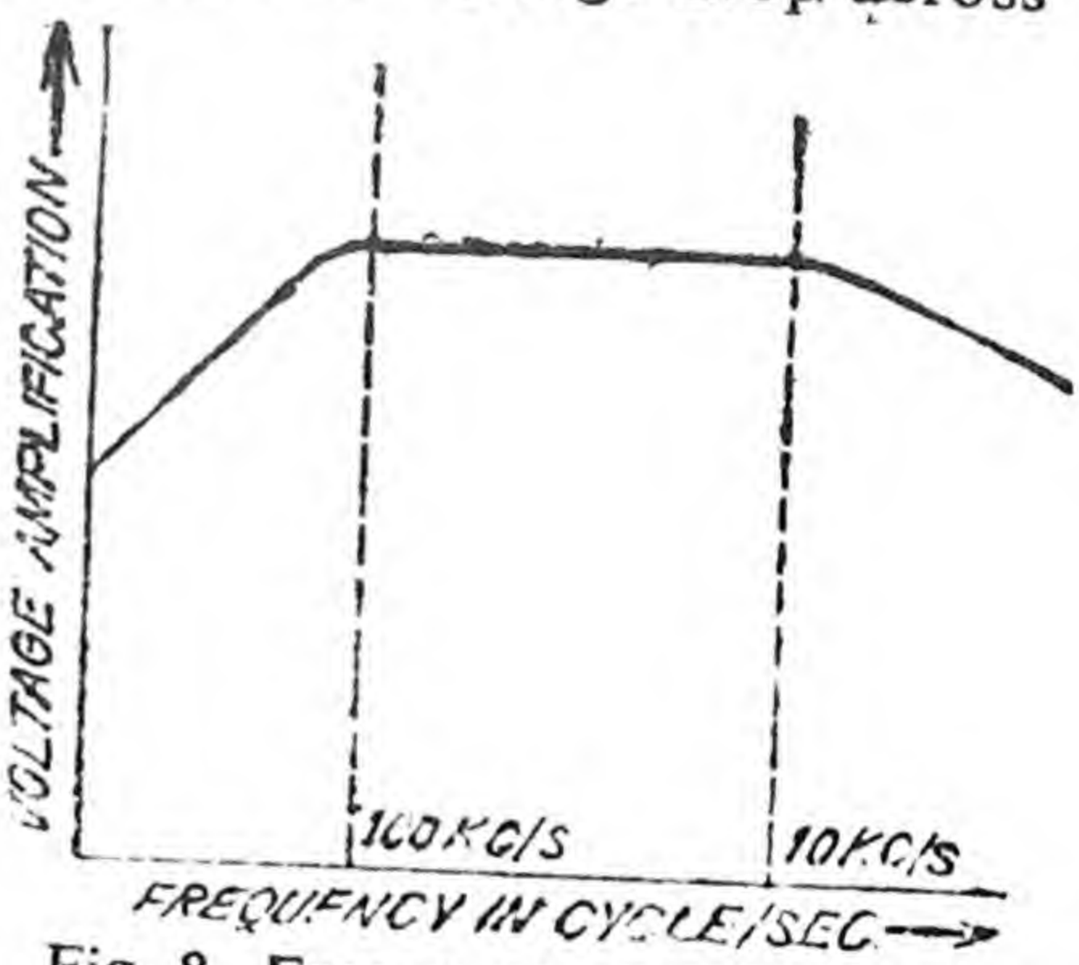


Fig. 8. Frequency response curve.

The reduction in amplification at high frequencies occurs due to the various capacitances which shunt the coupling and grid leak resistances. At sufficiently high frequencies, the effective load impedance is decreased enough because of low values of shunt resistances. Thus the output voltage of the amplifier is considerably reduced. The situation is shown by the bend in the curve of figure 8.

The distinguishing feature of this type of amplifier is that the voltage amplification remains substantially constant over a wide range of frequencies.

10·2. STUDY OF THE FREQUENCY AND PHASE RESPONSE OF RC-COUPLED AMPLIFIER :

The voltage amplifications of *RC* coupled amplifier at low, intermediate and high frequencies are given by

$$\mathbf{A}_l = -\frac{\mu R_l R_c}{r_p R_l + r_p R_c + R_l R_c + \dfrac{(r_p + R_l)}{j\omega C}}$$

$$= -\frac{\mu R_l R_o}{r_p R_l + r_p R_o + R_l R_o - \dfrac{j(r_p + R_l)}{\omega C}}$$

$$\mathbf{A}_m = -\frac{\mu R_l R_o}{r_p R_l + r_p R_o + R_l R_o}.$$

and $$\mathbf{A}_h = -\frac{\mu R_l R_o}{r_p R_l + r_p R_o + R_l R_o\,(1 + j\omega C_g r_p)}.$$

The intermediate frequency response of the amplifier is constant in magnitude and hence the low and high frequency performance can be generalized by using the mid frequency performance as standard of comparison.

$$\therefore \frac{\mathbf{A}_l}{\mathbf{A}_m} = \frac{-\dfrac{\mu R_l R_o}{r_p R_l + r_p R_o + R_l R_o - j\,(r_p + R_l)/\omega C}}{-\dfrac{\mu R_l R_o}{r_p R_o + r_p R_l + R_l R_o}}$$

$$= \frac{1}{1 - \dfrac{j}{\omega C}\left[\dfrac{r_n + R_l}{r_p R_l + r_p R_o + R_l R_o}\right]}. \qquad \ldots(1)$$

The second term in the denominator is an inverse function of frequency and at a particular frequency f_1 this term is equal to unity. The frequency can be calculated as follows :

$$\frac{1}{\omega C}\left[\frac{r_p + R_l}{r_p R_l + r_p R_o + R_l R_o}\right] = 1$$

or $$\frac{1}{2\pi f_0 C}\left[\frac{r_p + R_l}{r_p R_l + r_p R_o + R_l R_o}\right] = 1 \qquad \ldots(2)$$

$$\therefore \quad f_1 = \frac{1}{2\pi C}\left[\frac{1}{\dfrac{r_p R_l}{r_p + R_l} + R_o}\right].$$

Inserting (2) in equatian (1), we have

$$\frac{\mathbf{A}_l}{\mathbf{A}_m} = \frac{1}{1 - j f_1/f}. \qquad \ldots(3)$$

The equation gives the magnitude and phase of amplification at low frequencies with respect to values at mid frequencies. Since the phase angle at intermediate frequencies is 180°, as indicated by the negative sign, hence all angles for A_l lie in the third quadrant

Now $$\left|\frac{A_l}{A_m}\right| = \sqrt{\left(\frac{1}{1 + (f_1/f)^2}\right)}. \qquad \ldots(4)$$

The frequency f_1 is known as lower cut-off frequency. The

phase angle is given by

$$\theta_l = \tan^{-1}\frac{f_1}{f}$$ in addition to the inherent 180° shift of A_m,

i.e.

$$\theta_l = 180° + \tan^{-1}\frac{f_1}{f}. \qquad \ldots(5)$$

A dimensionless plot of equations (4) and (5) is shown in figure 9.

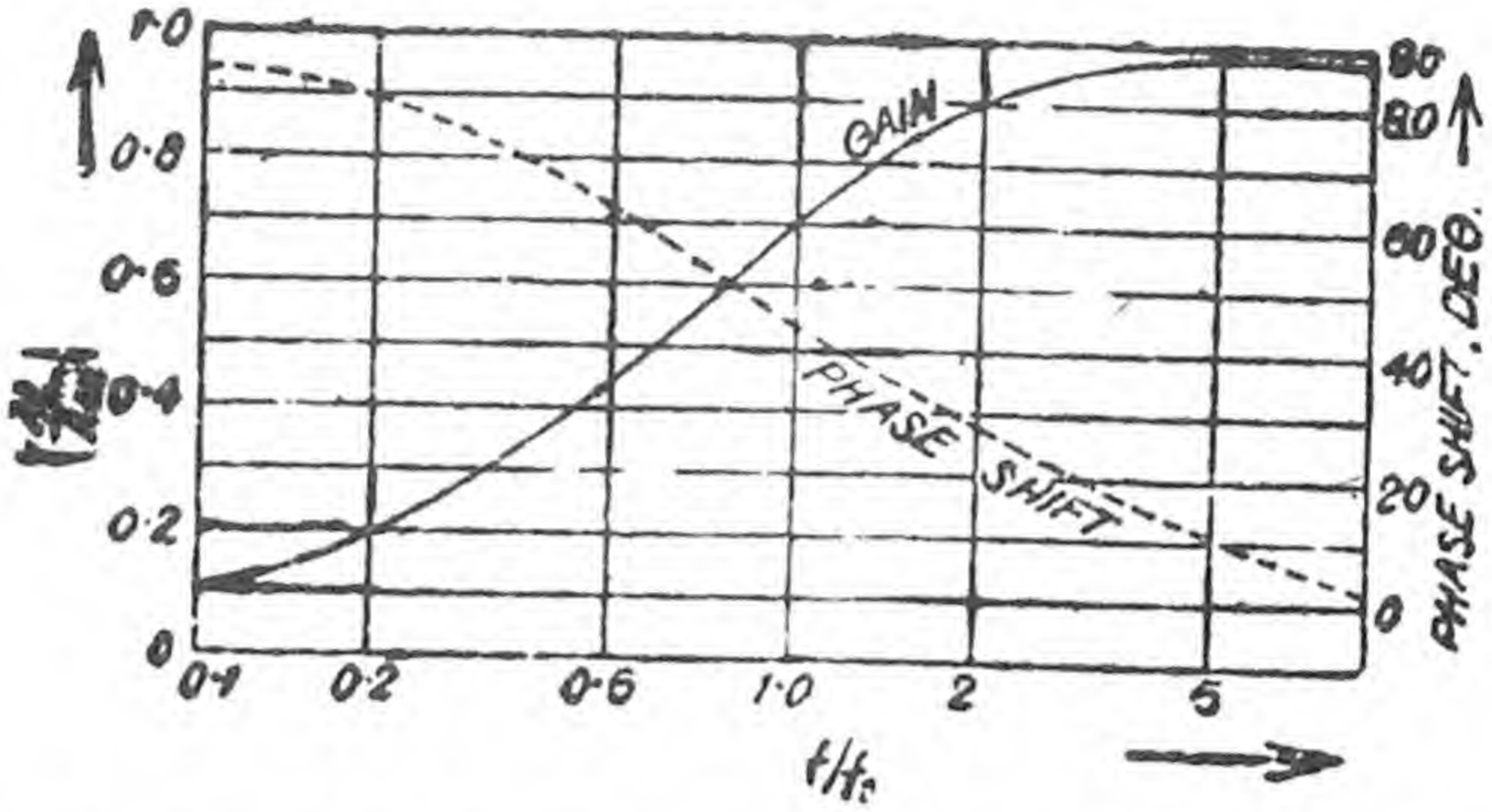


Fig. 9. Gain and phase shift curves for RC amplifier in low frequency range.

It is observed from the graph that as the ratio f/f_1 decreases, the gain asymptotically approaches zero while the phase shift approaches 90°.

Similarly, the high frequency gain ratio is

$$\frac{\mathbf{A}_h}{\mathbf{A}_m} = \frac{-\dfrac{\mu R_l R_o}{r_p R_l + r_p R_l + R_l R_c\,(1 + j\omega C_g r_p)}}{-\dfrac{\mu R_l R_o}{r_p R_l + r_p R_c + R_l R_o}}$$

$$= \frac{r_p R_l + r_p R_c + R_l R_o}{r_m R_o + r_p R_l + R_l R_c + R_l R_c\, j\omega C_g r_p}$$

$$= \frac{1}{1 + \dfrac{j\omega C_g}{\dfrac{1}{r_.} + \dfrac{1}{R_l} + \dfrac{1}{R_o}}} \qquad \ldots(6)$$

The frequency f_2 for which the second term in the denominator becomes unity is given by

$$f_2 = \frac{1}{2\pi C_g}\left(\frac{1}{r_p} + \frac{1}{R_l} + \frac{1}{R_o}\right). \qquad \ldots(7)$$

The frequency f_2 is commonly known as upper cut-off frequency. Substituting equation (7) in equation (6), we have

$$\frac{\mathbf{A}_h}{\mathbf{A}_m} = \frac{1}{1 + j.\, f/f_2}, \qquad \ldots(8)$$

from which $\left|\frac{A_h}{A_m}\right| = \sqrt{\left(\frac{1}{1+(f/f_2)^2}\right)}$ ...(9)

and $\theta_h = -\tan^{-1} f/f_2$ in addition to 180° shift of A_m gain

$\therefore$ $\theta_h = 180° - \tan^{-1}(f/f_2)$. ...(10)

A dimensionless plot of equations (9) and (10) is shown in figure 10.

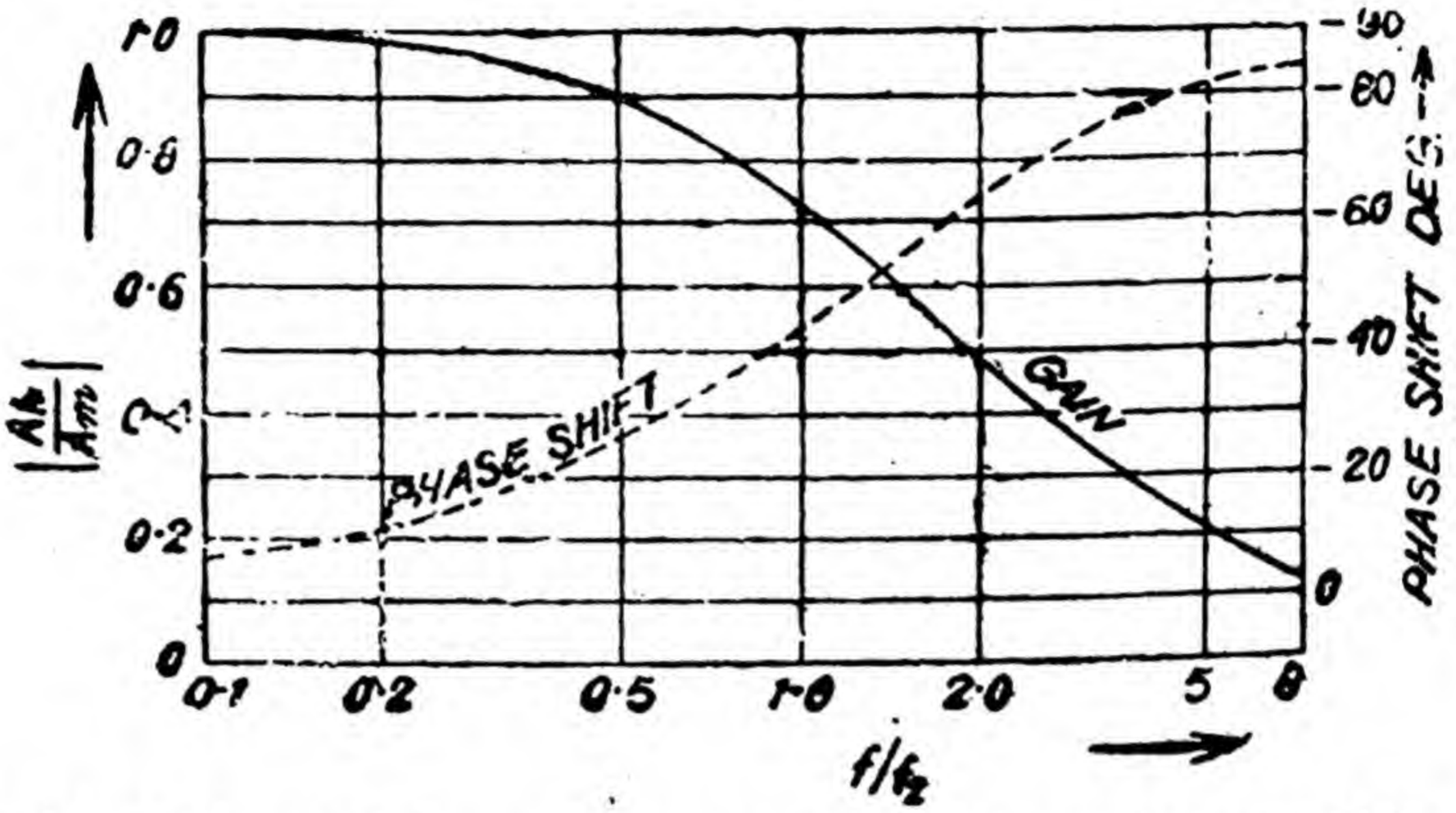


Fig. 10. Gain and phase shift curves for RC amplifier in high frequency range.

It is obvious from the figure 10 that at infinite frequency the gain would be zero while the phase shift approaches −90°.

10·3. GAIN BANDWIDTH PRODUCT OF RC COUPLED AMPLIFIER :

The width between the upper cut off frequency f_2 and the lower cut off frequency f_1 expressed in cycles is defined as the bandwidth of the amplifier. Thus

bandwidth $B = (f_2 - f_1)$. ...(1)

Generally f_1 is very small and hence f_2 alone may be regarded as bandwidth. According to equation (7),

$$B = f_2 = \frac{1}{2\pi C_g}\left(\frac{1}{r_p} + \frac{1}{R_l} + \frac{1}{R_c}\right). \quad ...(2)$$

The bandwidth may be increasd by decreasing the factor $\left(\frac{1}{r_p} + \frac{1}{R_l} + \frac{1}{R_c}\right)$ but at the same time the mid frequency gain also decreases and hence the product of mid frequency gain and band width is regarded as the index of the quality.

$$\therefore \quad A_m \times \text{bandwidth} = \frac{\mu R_l R_c}{r_p R_l + r_p R_c + R_l R_c} \times \frac{1}{2\pi C_g}\left(\frac{1}{r_p} + \frac{1}{R_l} + \frac{1}{R_c}\right)$$

$$= \frac{\mu}{r_p 2\pi C_p} = \frac{g_m}{2\pi C_g} \quad ...(3)$$

Equation (3) shows that the gain bandwidth product is independent of the circuit parameters R_l, R_c, C. If for the given tube and circuit parameters the bandwidth is increased, the gain will be

proportionately reduced. It is again clear from equation (3) that for large gain bandwidth product g_m should be large while inter-electrode capacitances should be small.

13·4. EFFECT OF CASCADING ON THE BANDWIDTH :

In order to increase the gain (which can be obtained by a single stage) several stages are used in series and the process is known as cascading. Let us consider the case of n such stages. If the voltage gain per stage is A, then the gain in cascading will be A^n. If A_m denotes the mid frequency gain, then overall gain of the amplifier will be $(A_m)^n$.

In case of low frequency range, the relative gain or the single stage is

$$\left(\frac{A_l}{A_m}\right)=\frac{1}{\sqrt{\left\{1+\left(\frac{f_1}{f}\right)^2\right\}}},$$

and hence the relative gain using n stages in cascade is given by

$$\left(\frac{A_l}{A_m}\right)^n=\frac{1}{\left[1+\left(\frac{f_1}{f}\right)\right]^{n/2}} \qquad ...(1)$$

Let f_{1n} be the lower cut off frequency of the cascade amplifier, which is defind as the frequency at which the relative gain is $\frac{1}{\sqrt{2}}$, *i.e.*

$$\left(\frac{A_l}{A_m}\right)^n=\frac{1}{\sqrt{2}}$$

so that
$$\frac{1}{\sqrt{2}}=\frac{1}{\left[1+\left(\frac{f_1}{f_{1n}}\right)^2\right]^{n/2}}$$

or
$$\left[1+\left(\frac{f_1}{f_{1n}}\right)^2\right]^n=2$$

or
$$1+\left(\frac{f_1}{f_{1n}}\right)^2=2^{1/n}$$

or
$$\left(\frac{f_1}{f_{1n}}\right)^2=2^{1/n}-1$$

or
$$\left(\frac{f_1}{f_{1n}}\right)=\sqrt{(2^{1/n}-1)}$$

or
$$f_{1n}=\frac{f_1}{\sqrt{(2^{1/n}-1)}}. \qquad ...(2)$$

In the high frequency range, the relative voltage gain of n stage cascade amplifier is given by

$$\left(\frac{A_h}{A_m}\right)^n=\frac{1}{\left[1+\left(\frac{f}{f_2}\right)^2\right]^{n/2}}$$

If f_{2n} is the upper cut off frequency, then

$$\frac{1}{\sqrt{2}}=\frac{1}{\left[1+\left(\frac{f_{2n}}{f_2}\right)^2\right]^{n/2}},$$

or
$$\left[1+\left(\frac{f_{2n}}{f_2}\right)^2\right]^{n/2}=\sqrt{2}$$

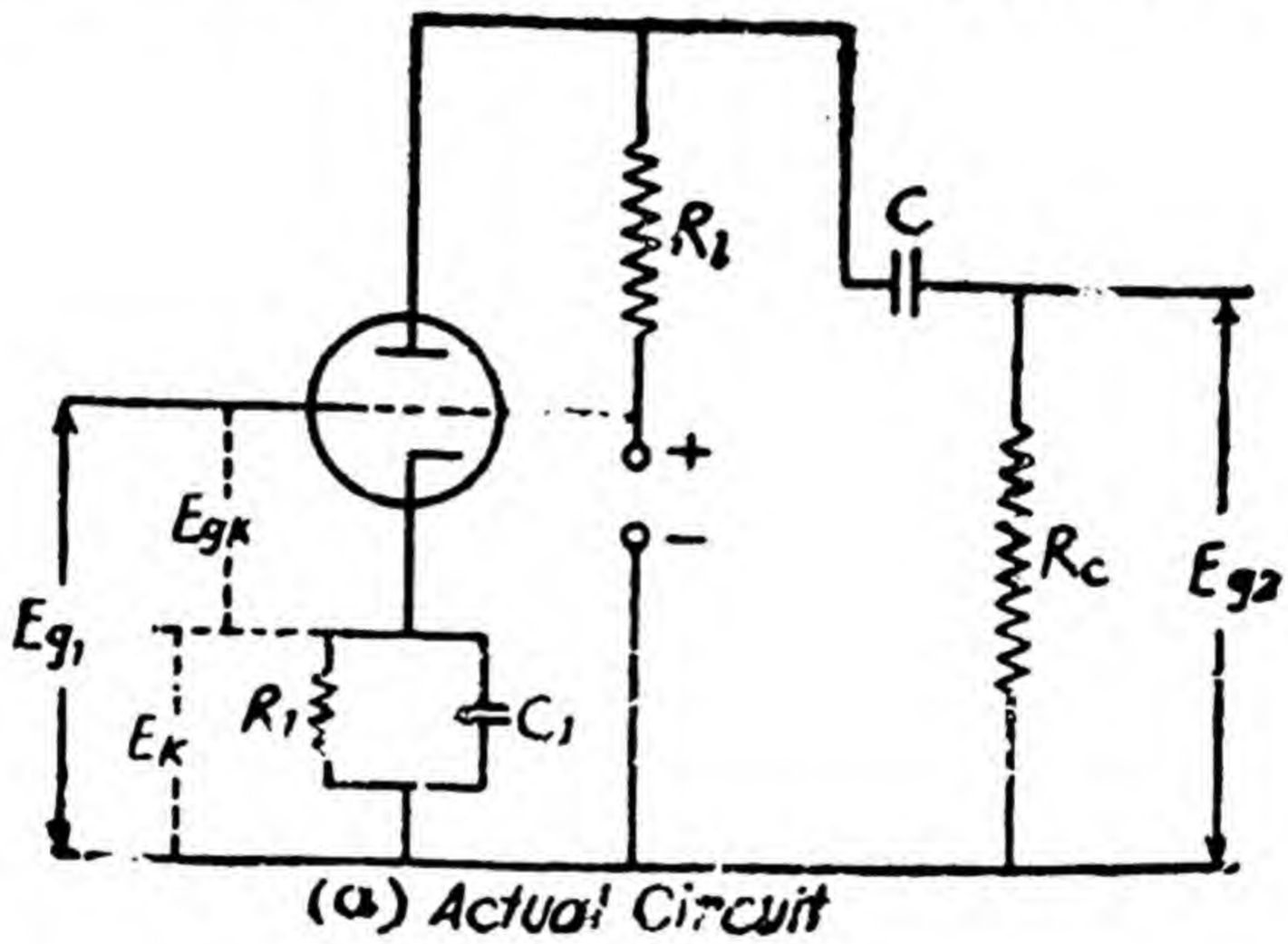
or
$$1+\left(\frac{f_{2n}}{f_2}\right)^2=(2)^{1/n}.$$

$$f_{2n}=f_2(2^{1/n}-1)^{1/2}. \qquad ...(3)$$

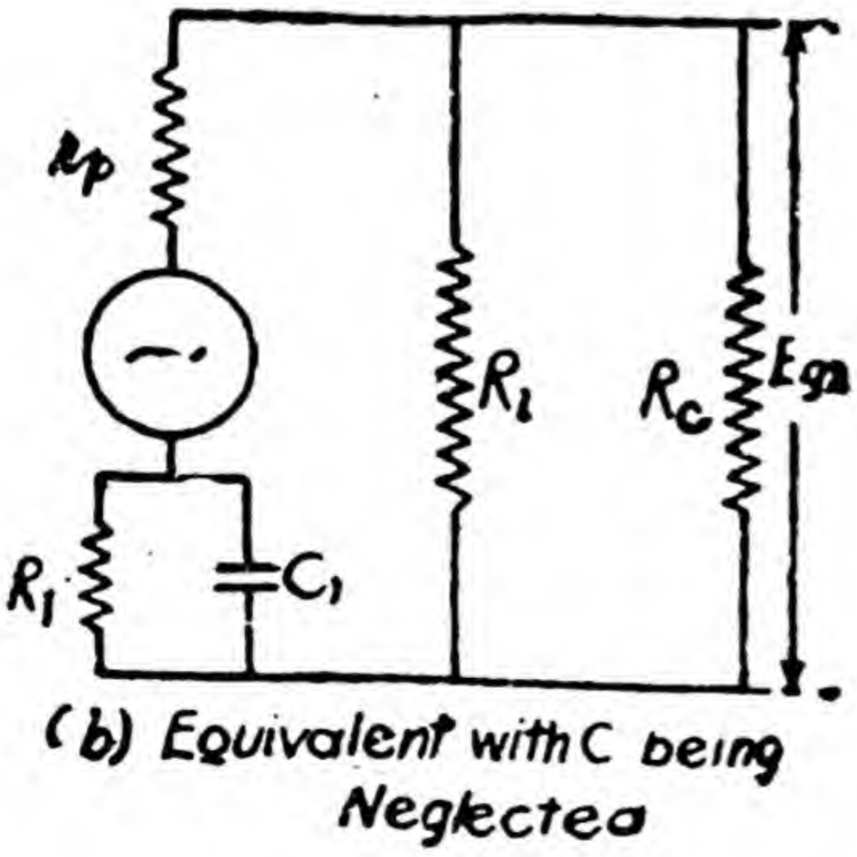
Since $f_2 >> f_1$, the effective bandwidth is given by f_{2n}. It is clear from equation (3) that the bandwidth is increased by a factor $\sqrt{(2^{1/n}-1)}$.

10·5. EFFECT OF CATHODE RESISTANCE AND CONDENSER ON THE FREQUENCY RESPONSE OF AN AMPLIFIER :

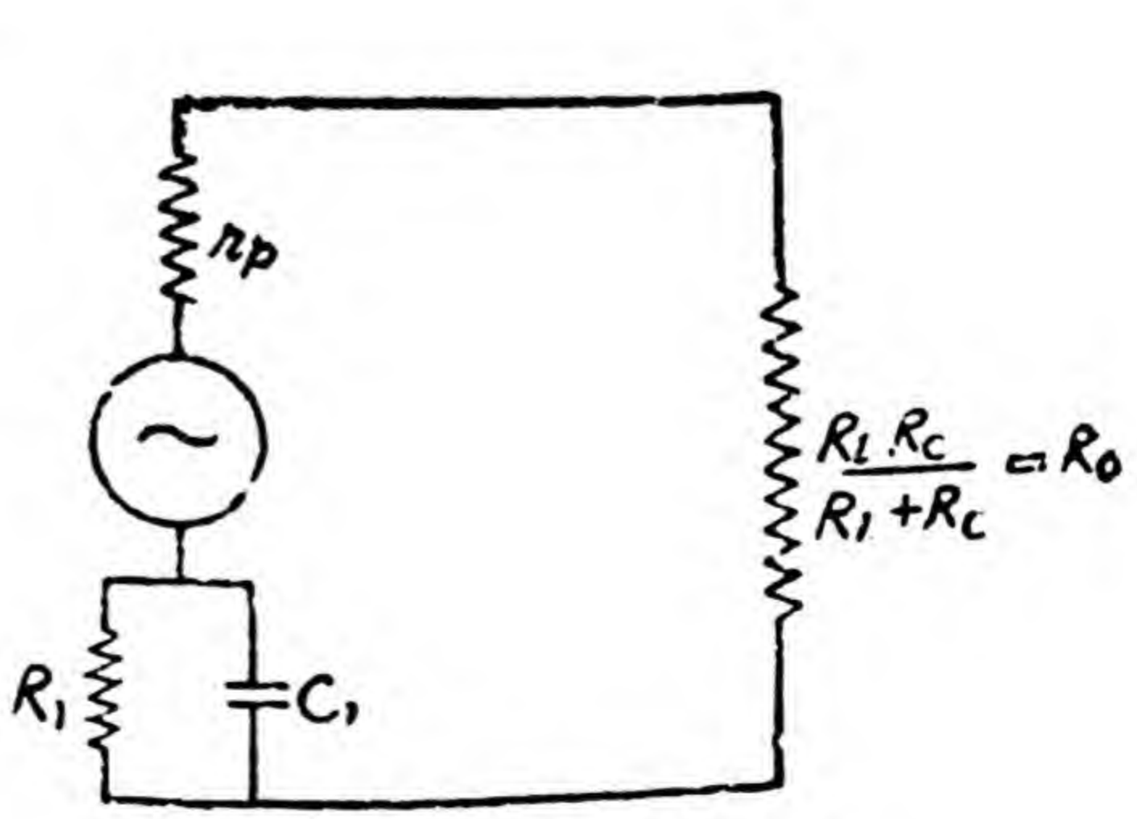
At low frequencies the effect of the condenser C can be neglected, otherwise the circuit becomes complicated. Figure 11 shows the actual circuit of RC amplifier and its equivalent.



(a) Actual Circuit



(b) Equivalent with C being Neglected



(c) simplified form

Fig. 11.

The impedance of R_1 and C_1 is given by

$$\frac{1}{\mathbf{Z}_k}=\frac{1}{R_1}+j\omega C_1$$

$$=\frac{1+j\omega C_1 R_1}{R_1}$$

$$\mathbf{Z}_k=\frac{R_1}{1+j\omega C_1 R_1},$$

$$\mathbf{E}_{g_2}=\frac{\mu \mathbf{E}_{gk} R_0}{r_p+R_0+Z_k}. \qquad \ldots(1)$$

$$\mathbf{E}_{gk}=\mathbf{E}_{g_1}-\mathbf{E}_k$$

$$=\mathbf{E}_{g_1}-\mathbf{I}_0\mathbf{Z}_k,$$

where $\mathbf{I}_0$ is cathode circuit current.

$$\therefore \quad \mathbf{I}_0=\frac{\mu \mathbf{E}_{gk}}{r_p+R_0+Z_k},$$

$$\therefore \quad \mathbf{E}_{gk}=\mathbf{E}_{g_1}-\frac{\mu \mathbf{E}_{gk} Z_k}{r_p+R_0+Z_k}$$

or

$$\mathbf{E}_{gk}\left(1+\frac{\mu Z_k}{r_p+R_0+Z_k}\right)=\mathbf{E}_{g_1}$$

or

$$\mathbf{E}_{gk}=\frac{\mathbf{E}_{g_1}}{1+\dfrac{\mu Z_k}{r_p+R_0+Z_k}}$$

$$\mathbf{E}_{gk}=\frac{\mathbf{E}_{g_1}(r_p+R_0+Z_k)}{r_p+R_0+Z_k+\mu Z_k}. \qquad \ldots(2)$$

From (2) substituting the value of $\mathbf{E}_{gk}$ in (1),

$$\mathbf{E}_{g_2}=\frac{\mu R_0}{(r_p+R_0+Z_k)}\cdot\frac{\mathbf{E}_{g_1}(r_p+R_0+Z_k)}{(r_p+R_0+Z_k+\mu Z_k)},$$

Gain at low frequency $A_l=\dfrac{\mathbf{E}_{g_2}}{\mathbf{E}_{g_1}}=\dfrac{\mu R_0}{r_p+Z_k(1+\mu)+R_0}.$

$$\mathbf{A}_l=\frac{\mu R_0}{r_p+Z_k(1+\mu)+R_0}. \qquad \ldots(3)$$

At medium frequency Z_k can be neglected ; hence

$$A_m=\frac{\mu R_0}{r_p+R_0}.$$

Thus gain at low frequency is decreased due to the factor $\mathbf{Z}_k(1+\mu)$ in the denominator.

10·6. LOW FREQUENCY COMPENSATION OF RC COUPLED AMPLIFIER :

It has been observed that the low frequency response of RC amplifiers falls off. Technically, it is possible to extend the low frequency responnse to as low a frequency as desired. The loss of the gain at low frequencies can be compensated by the addition of a low frequency compensating filter in series with the plate load resistor as shown in fig. 12. The filter consists of a resistance R_f and a con-

denser C_f. The filter provides a plate load impedance that rises with decrease in frequency and can compensate for the reduction in gain caused by capacitor C.

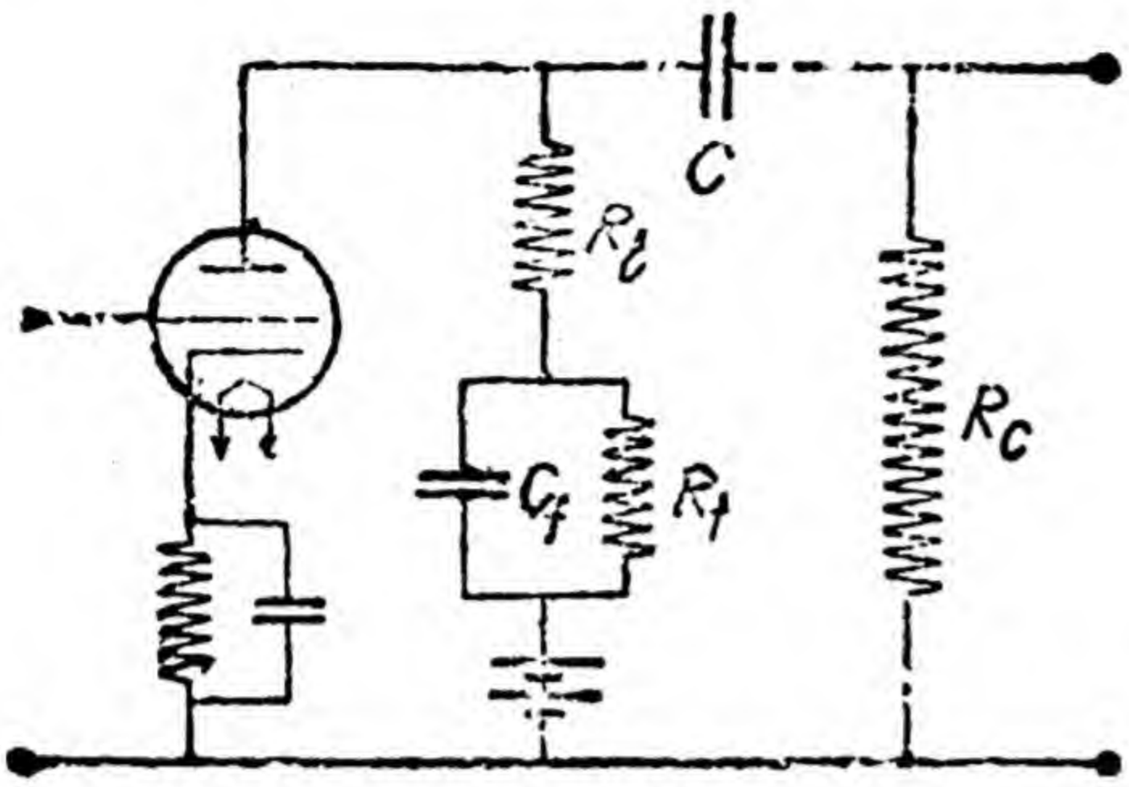
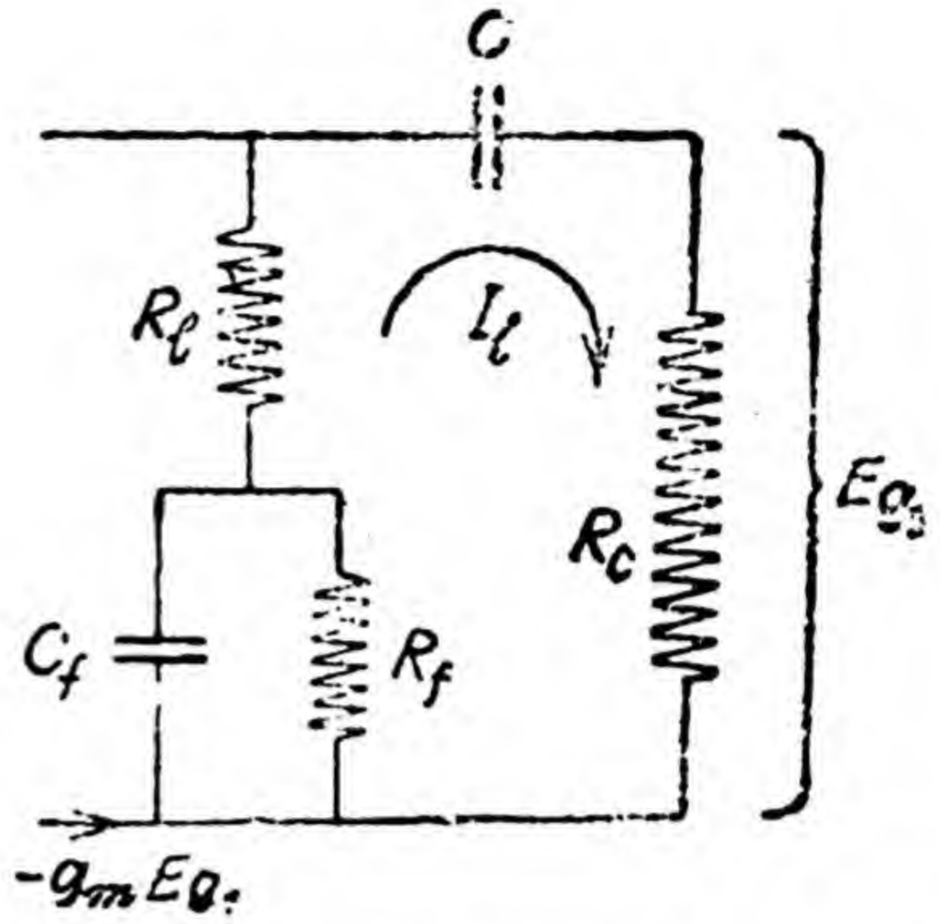


Fig. 12. (*a*) Low frequency compensation of *R-C* amplifier.



12. (*b*) Equivalent circuit of frequency compensation. Since r_p is very large, it may be emitted from circuit.

From the equivalent circuit, the current $\mathbf{I}_l$ is given by

$$\mathbf{I}_l = \frac{-g_m \mathbf{E}_{g_1} \left\{ \dfrac{\left(\dfrac{R_f}{j\omega C_f}\right)}{R_c + \left(R_f + \dfrac{1}{j\omega C_f}\right)} \right\}}{R_c + \dfrac{1}{j\omega C} + R_l + \dfrac{R_f / j\omega C_f}{R_f + \dfrac{1}{j\omega C_f}}} \qquad \ldots(1)$$

Since $\left(R_c + \dfrac{1}{\omega C}\right) > > \left(R_l + \dfrac{R_f/j\omega C_f}{R_f + 1/j\omega C_f}\right)$.

We can ignore the latter. Therefore equation (1) can be written as :

$$I_l = -g_m \mathbf{E}_{g_1} \left(R_l - \frac{C_f}{R_f - j/\omega C_f}\right) \cdot \frac{1}{(R_c - j/\omega C)} \cdot \qquad \ldots(2)$$

The output voltage $\mathbf{E}_{g_2}$ is given by

$$\mathbf{E}_{g_2} = \mathbf{I}_l R_c = -g_m \mathbf{E}_{g_1} \left(R_l - \frac{jR_f/\omega C}{R_f - j/\omega C_f}\right)\left(\frac{R_c}{R_c - j/\omega C}\right) \qquad \ldots(3)$$

$$\therefore \qquad \mathbf{A}_l = \frac{\mathbf{E}_{g_2}}{\mathbf{E}_{g_1}} = -g_m \left(R_l + \frac{jR_f/\omega C_f}{R_f - j/\omega C_f}\right) \frac{R_c}{(R_c - j/\omega C)} \cdot \qquad \ldots(4)$$

Since $R_f > > \dfrac{1}{\omega C_f}$, we can ignore the latter in comparison.

Therefore

$$\mathbf{A}_l = -g_m \left(R_l - \frac{jR_f/\omega C_f}{R_f}\right)\left(\frac{R_c}{R_c - j/\omega C}\right)$$

$$\mathbf{A}_l = -g_m \left(\frac{\omega C_f R_l - j}{C_f}\right)\left(\frac{CR_c}{\omega C R_c - j}\right) \qquad \ldots(5)$$

If $C_f R_l = C R_c$, the gain at low frequencies is independent of frequency. In other words time constant of the coupling circuit is equal to the time constant $C_f R_l$, A_l is independent of frequency.

10·7. HIGH FREQUENCY COMPENSATION OF RC COUPLED AMPLIFIER

There are two methods for the high frequency compensation of the RC amplifier :

1. Shunt compensation
2. Series compensation

1. Shunt compensation : In this method an inductance L is placed in shunt with C_g in the plate load so as to counteract the effect of C_g in reducing the load impedance at high frequencies. This method is known as shunt compensation because the increased reactance of shunt coil L at high frequencies serves to peak the response due to the total shunt capacitance. The value of the inductance is so chosen that it resonates with the total effective capacitance of the output of one tube and input of the following tube in the neighbourhood of the falling frequency. The circuit arrangement is shown in fig. 13.

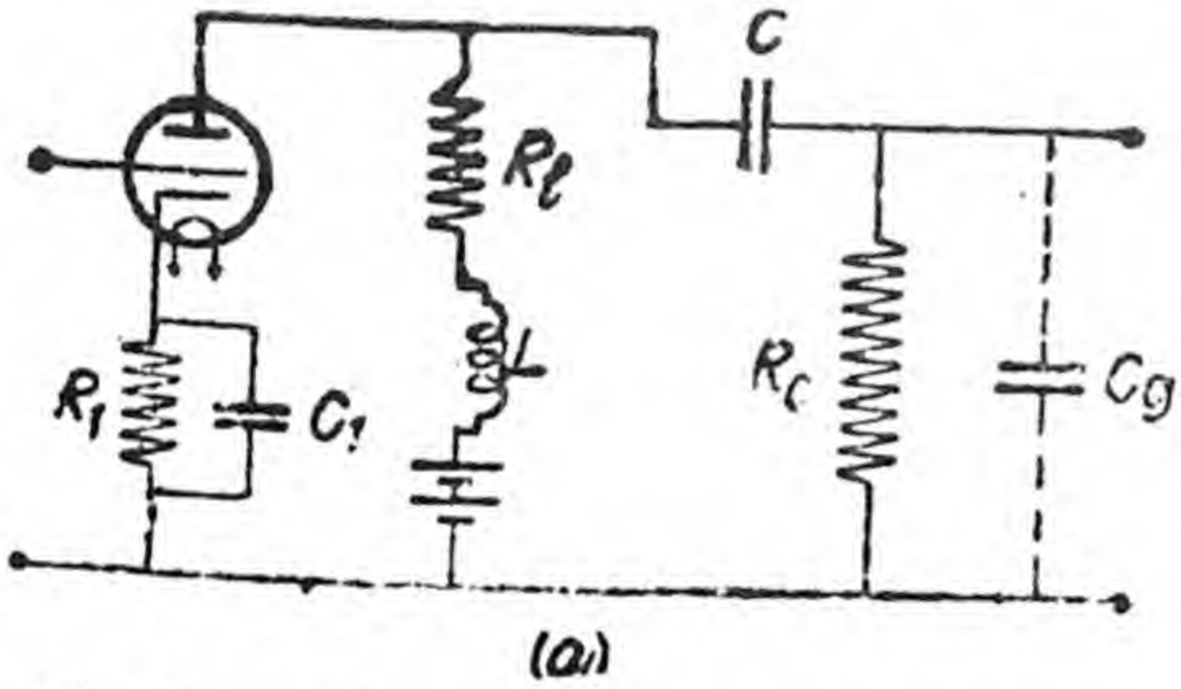
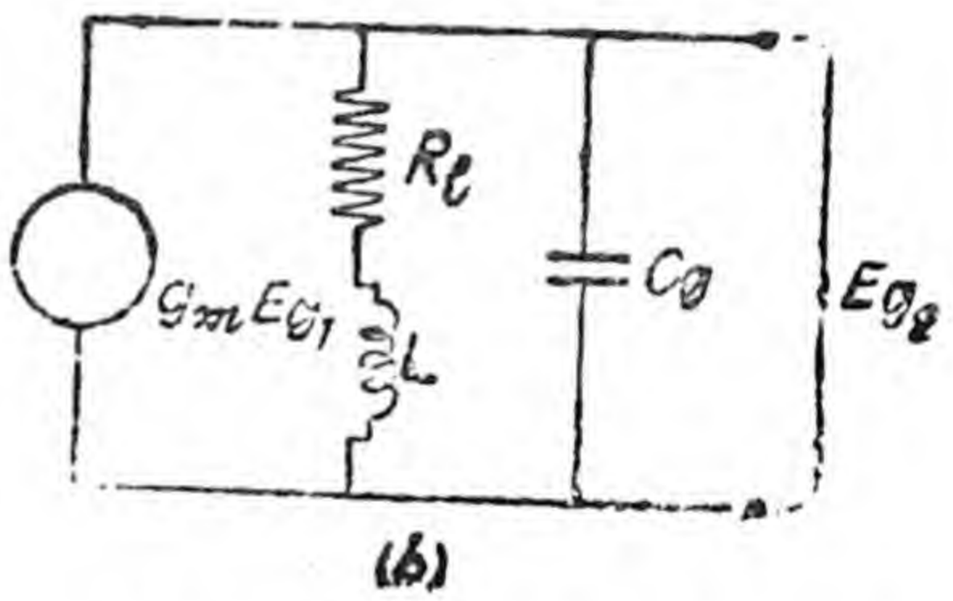


Fig 13. Shunt compensated circuit



Equivalent circuit of 13 (*a*)

Considering the equivalent circuit of figure 13.

$$A_h = -g_m \frac{(R_l + j\omega L)\ (-j/\omega C_g)}{(R_l + j\omega L - j/\omega C_g)}$$

$$= -g_m R_l \frac{(\omega L/R_l - j)}{\omega C_g R_l + j\,(\omega^2 LC_g - 1)} \quad \text{.. (1)}$$

The magnitude of the high frequency gain A_h is

$$|A_h| = -g_m R_l \sqrt{\left(\frac{\omega^2 L^2/R_l^2 + 1}{1 + \omega^2(C_g^2 R_l^2 - 2LC_g) + \omega^4 L^2 C_g^2}\right)}. \quad \text{...(2)}$$

In case of uncompensated amplifier $L = 0$ and the upper cut off frequency f_2 is given by equation 7 (article 10·2) as

$$f_2 = \frac{1}{2\pi C_g R_l}. \quad \text{...(3)}$$

Substituting this value of f_2 in equation (2), we have

$$|A_h| = -g_m R_l \times \sqrt{\left(\frac{1+(L^2/C_0^2R_l^4)\,(f/f_2)^2}{1+(1-2L/C_0R_l^2)\,(f/f_2)^2+L^2/C_0^2R_l^4)(f/f_2)^4}\right)} \quad ...(4)$$

Since mid frequency gain is $A_m = -g_m R_l$, therefore from equation (4), we have

$$\frac{|A_h|}{|A_m|} = \sqrt{\left(\frac{1+(L^2/C_0^2R_l^4)\,(f/f_2)^2}{1+(1-2L/C_0R_l^2)\,(f/f_2)^2+(L^2/C_0^2R_l^4)(f/f_2)^4}\right)} \quad ...(5)$$

It is convenient to define a quantity Q as

$$Q = \frac{L}{C_0 R_l^2}.$$

Introducing Q in equation (5), we have

$$\frac{|A_h|}{|A_m|} = \sqrt{\left(\frac{1+Q^2\,(f/f_2)^2}{[1+Q.(f/f_2)]^2+(f/f_2)^2}\right)}. \quad ...(6)$$

A graph plotted between relative gain $|A_h| / |A_m|$ as a function of (f/f_2) for different values of Q is shown in fig. 14.

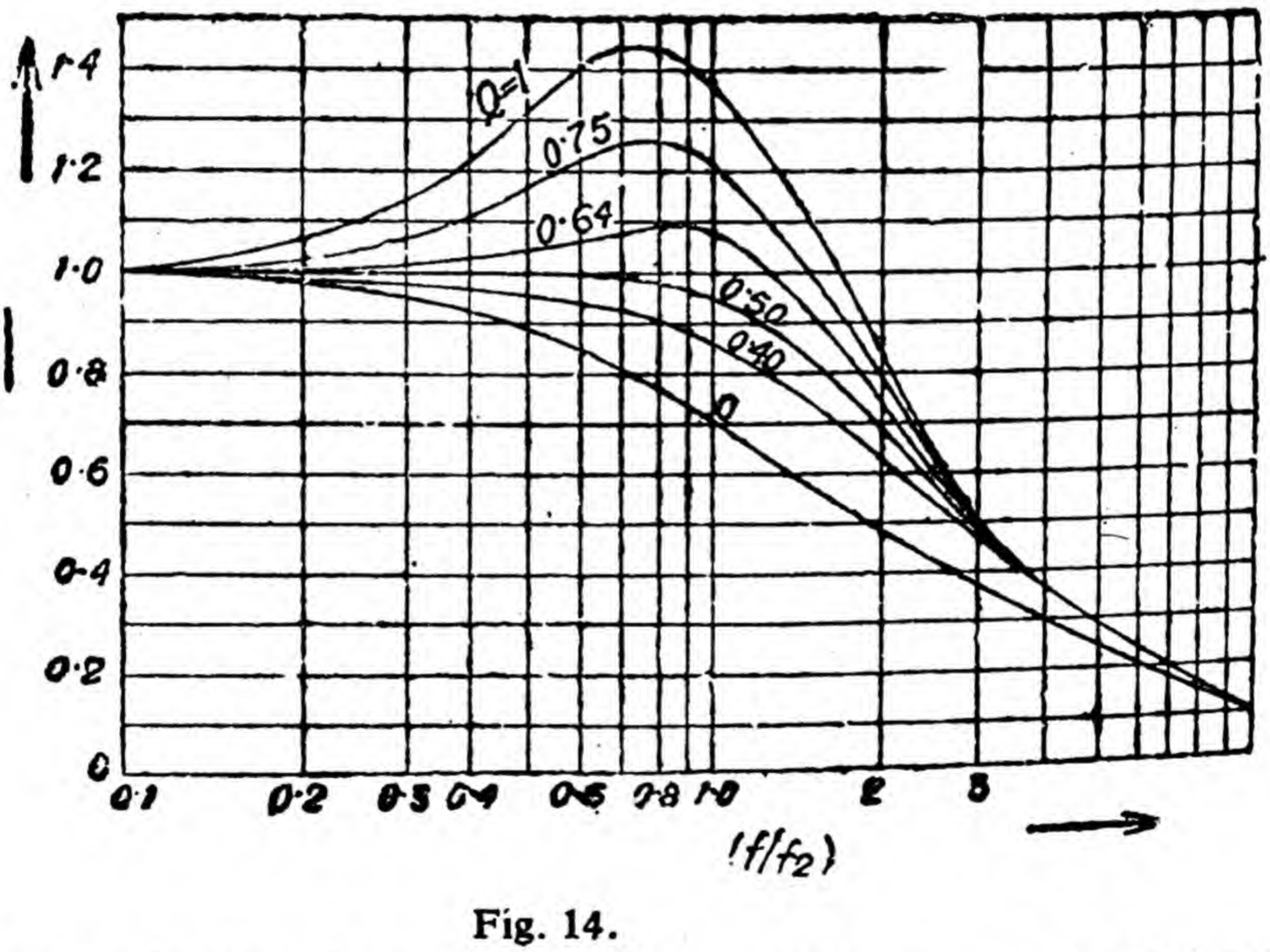


Fig. 14.

The voltage gain may be constant upto a certain designated high frequency. It can be seen from figure 14 that $Q = 0{\cdot}45$ provides an improved band width if a very slight response rise can be tolerated. Thus

$$Q = \frac{L}{C_0 R_l^2} = 0{\cdot}45. \quad ...(7)$$

is a condition of maximum flatness for the response curve. The required value of L can be obtained by the following expression

$$L = 0{\cdot}45 C_0 R_l^2 \quad ...(8)$$

2. Series compensation :

The circuit arrangement is shown in fig. 15.

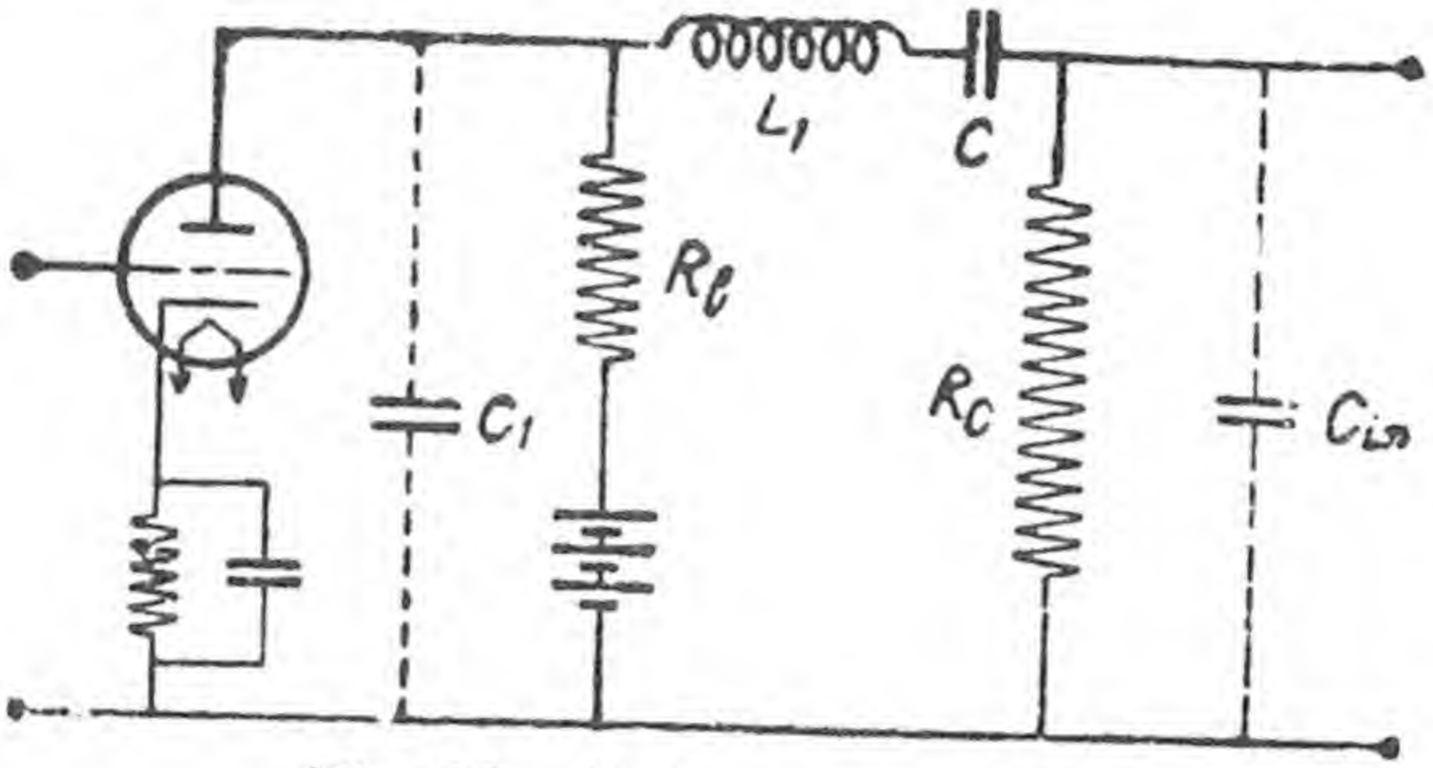


Fig. 15. Series compensation.

In this case a small inductance coil L_1 is connected in series with the coupling capacitor C. At high frequencies, the coil L_1 resonates with the input capacitance C_{in} of the next stage, thus causing an increased voltage across C_{in}. Since both C_{in} and R_c are across the input of the next stage, the input voltage and hence the gain is boosted.

10·8. INDUCTANCE-CAPACITANCE (LC) COUPLED AMPLIFIER

The circuit of the inductance-capacitance coupled amplifier differs from that of the resistance-capacitance coupled amplifier only in the use of an inductor plate load instead of a plate resistor. The use of an inductor instead of a resistor in the plate circuit makes possible the use of a smaller plate supply potential for a particular tube operating conduction, since the d.c. resistance of the inductor is small and so the d.c. potential drop across this inductor will also be small.

The impedance that the coupling coil offers to the signal current is dependent on its inductance and the frequency to the signal. In order to obtain high impedance at low audio frequencies, the inductance of the coupling coil is made as high as is practicable. The inductance of choke coils used as coupling impedance for a.f. amplifier ranges from 10 to 800 henrys. The schematic diagram of the amplifier is give n in figure 16.

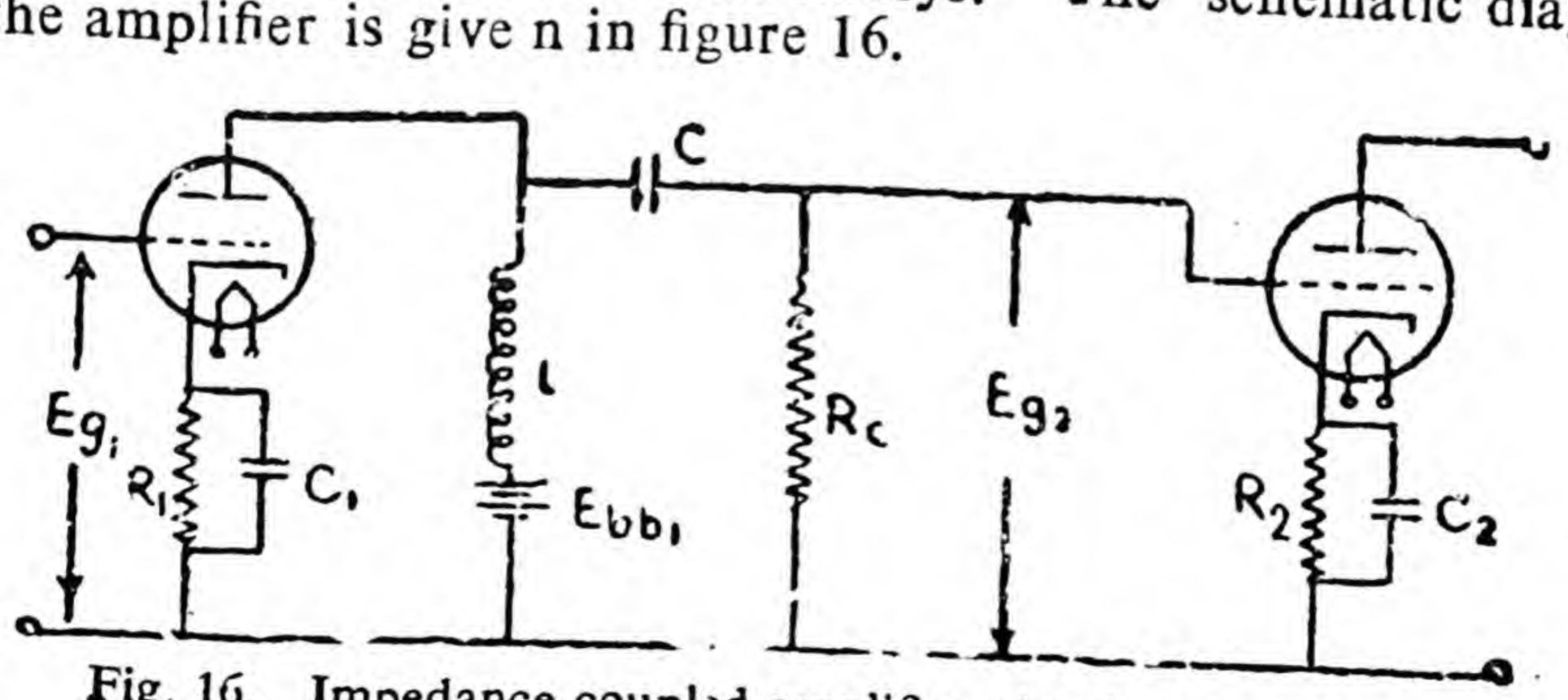
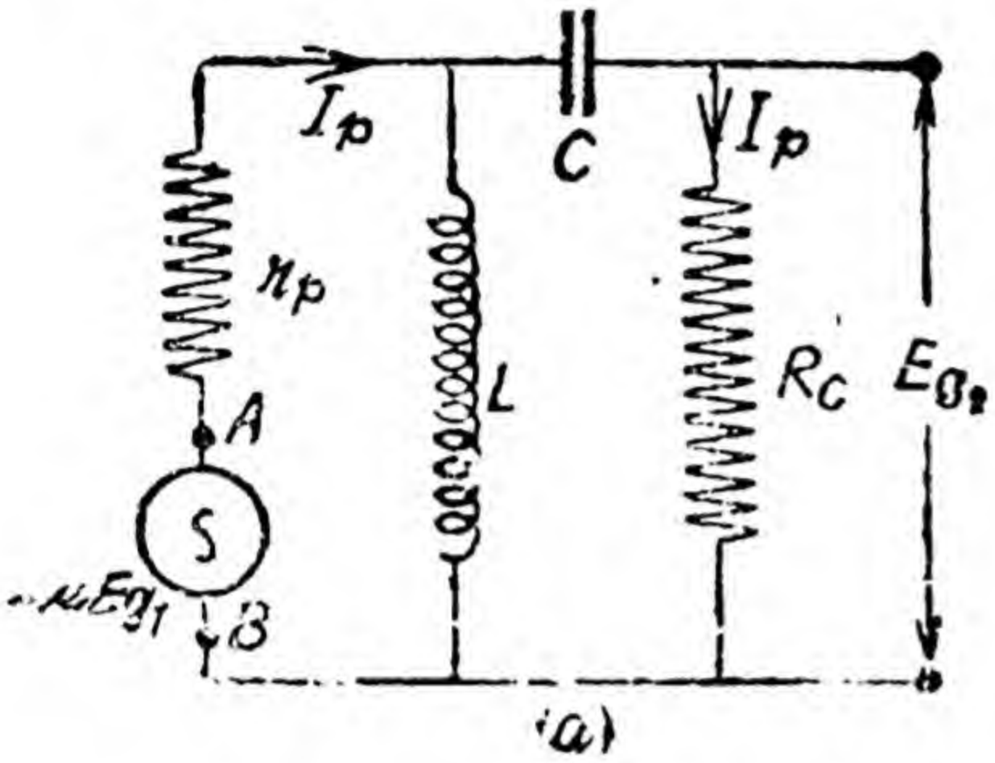
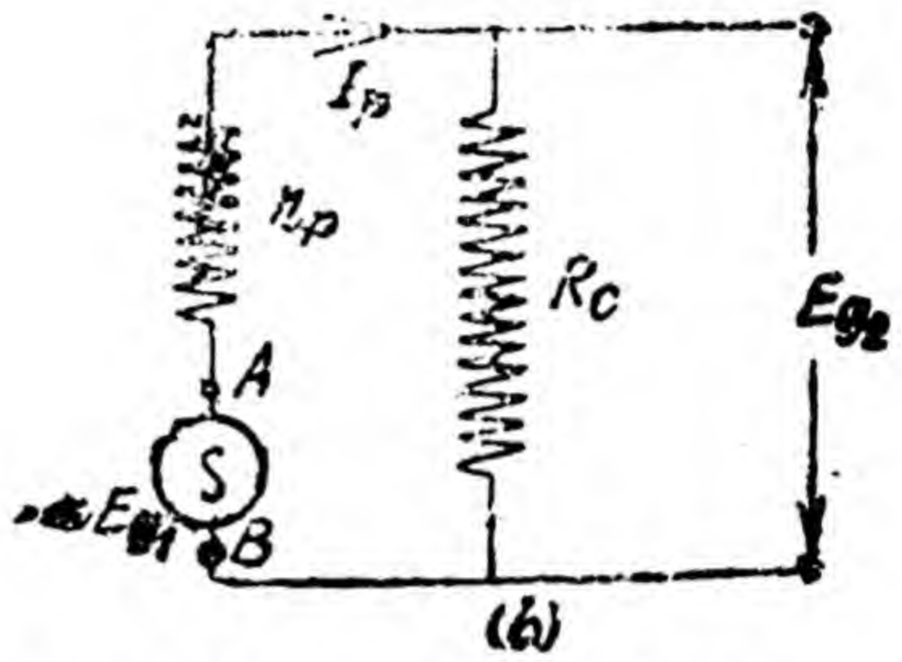


Fig. 16. Impedance coupled amplifier circuit using triodes.

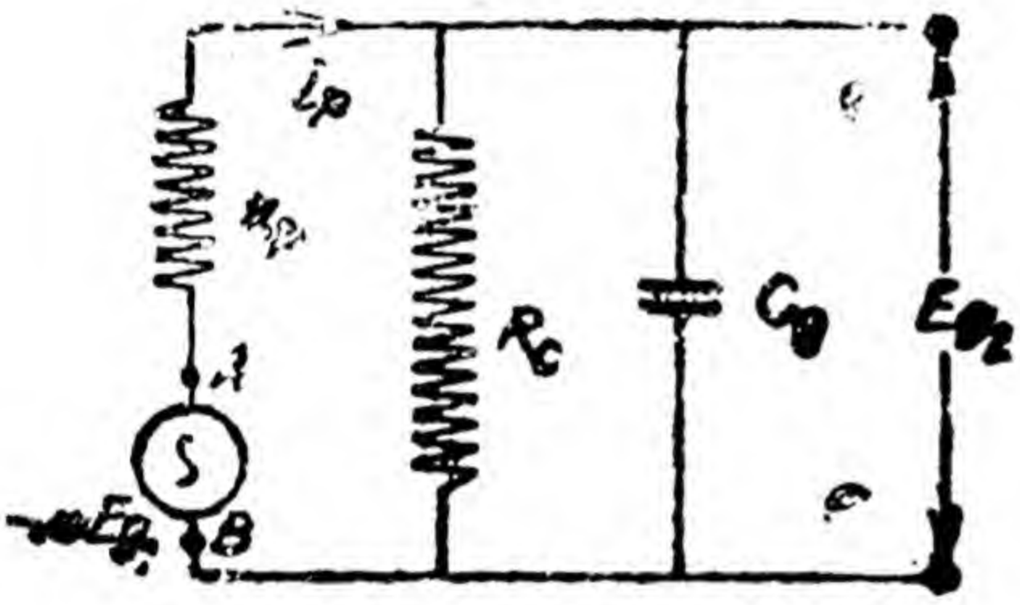
Voltage amplification of the impedance-coupled Amplifier : The method of determining the voltage amplification produced by the impedance coupled amplifier is similar to that used for the *RC* coupled amplifier circuit. The circuit characteristics are again observed at low, intermediate and high frequencies of the audio range. The equivalent circuits for the three frequency ranges are given in figure 17.



(*a*) At low frequencies.



(*b*) At mid frequencies.



(*c*) At high frequencies.

Fig. 17. Equivalent circuits at low, mid and high frequencies.

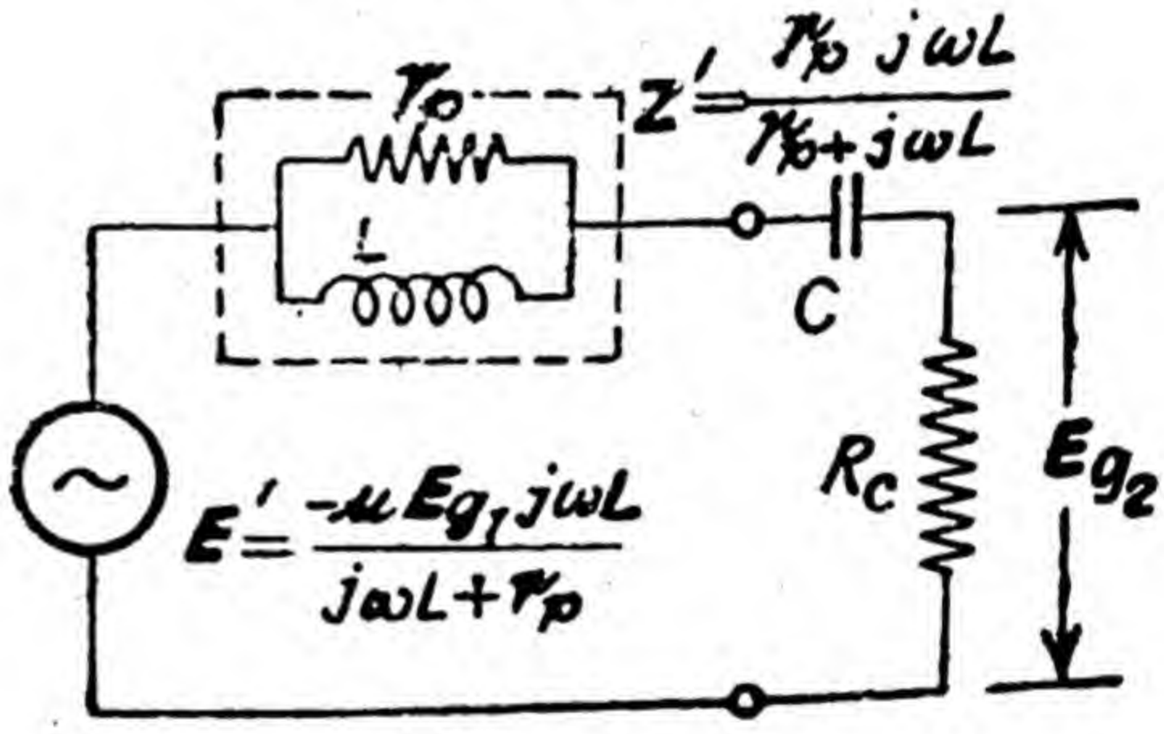


Fig. 18. Thevenin's equivalent at low frequencies.

Expressions for voltage amplification at low, intermediate and high frequencies can be deduced on a similar pattern as done in case of *RC* coupled amplifier. For example, we shall deduce here voltage amplification at low frequencies.

The Thevenin's equivalent circuit of fig. 17 (*a*) at low frequencies is shown in figure 18.

The total impedance of the circuit is given by

$$Z = R_c + \frac{1}{j\omega C} + \frac{r_p . j\omega L}{r_p + j\omega L}$$

The current I flowing in the circuit is

$$\mathbf{I} = \frac{-\frac{\mu \mathbf{E}g_1 j\omega L}{j\omega L + r_p}}{R_c + \frac{1}{j\omega C} + \frac{r_p j\omega L}{r_p + j\omega L}}$$

$$= \frac{\dfrac{-\mu \mathbf{E}g_1 \, j\omega L}{j\omega L + r_p}}{\dfrac{R_c(r_p + j\omega L) + (r_p + j\omega L)/j\omega C + r_p \, j\omega L}{r_p + j\omega L}}$$

$$= \frac{-\mu \mathbf{E}g_1 \, j\omega L}{R_c r_p + j\omega L \, (r_p + R_c) + \dfrac{L}{C} - \dfrac{j r_p}{\omega C}} \quad \ldots(1)$$

The output voltage E_{g2} is given by

$$\mathbf{E}g_2 = \frac{-\mu \mathbf{E}g_1 \, j\omega L \times R_c}{R_c r_p + j\omega L \, (r_p + R_c) + \dfrac{L}{C} - \dfrac{j r_p}{\omega C}} \quad \ldots(2)$$

The complex voltage amplification A_l is given by

$$\mathbf{A}_l = \frac{\mathbf{E}g_2}{\mathbf{E}g_1} = \frac{-\mu j\omega L R_c}{R_c r_p + j\omega L \, (r_p + R_c) + \dfrac{L}{C} - \dfrac{j r_p}{\omega C}}. \quad \ldots(3)$$

From the consideration of equivalent circuit value of A_m and A_h can be deduced.

Frequency response curve of the amplifier : The analysis shows that the response is generally similar to that for the *RC* coupled amplifier except that the flat region is narrower, although the mid-frequency gain is somewhat higher.

The decrease in gain is greater at both low and high frequencies than for *RC* coupled amplifier. This is due to the fact that the impedance of a resistor is fairly uniform over all frequencies and its distributed capacitance is negligible, while the impedance of a coil varies directly with frequency and the coil has an appreciable amount of distributed capacitance. At low audio frequencies, the impedance of coupling coil will be relatively low, thus causing a decrease in the voltage gain at these frequencies. The distributed capacitance of the coil increases the total shunt capacitance of the circuit, and at the high audio frequencies this increase in C_g causes more of the output current to be by-passed, thus further decreasing the voltage gain at these frequencies.

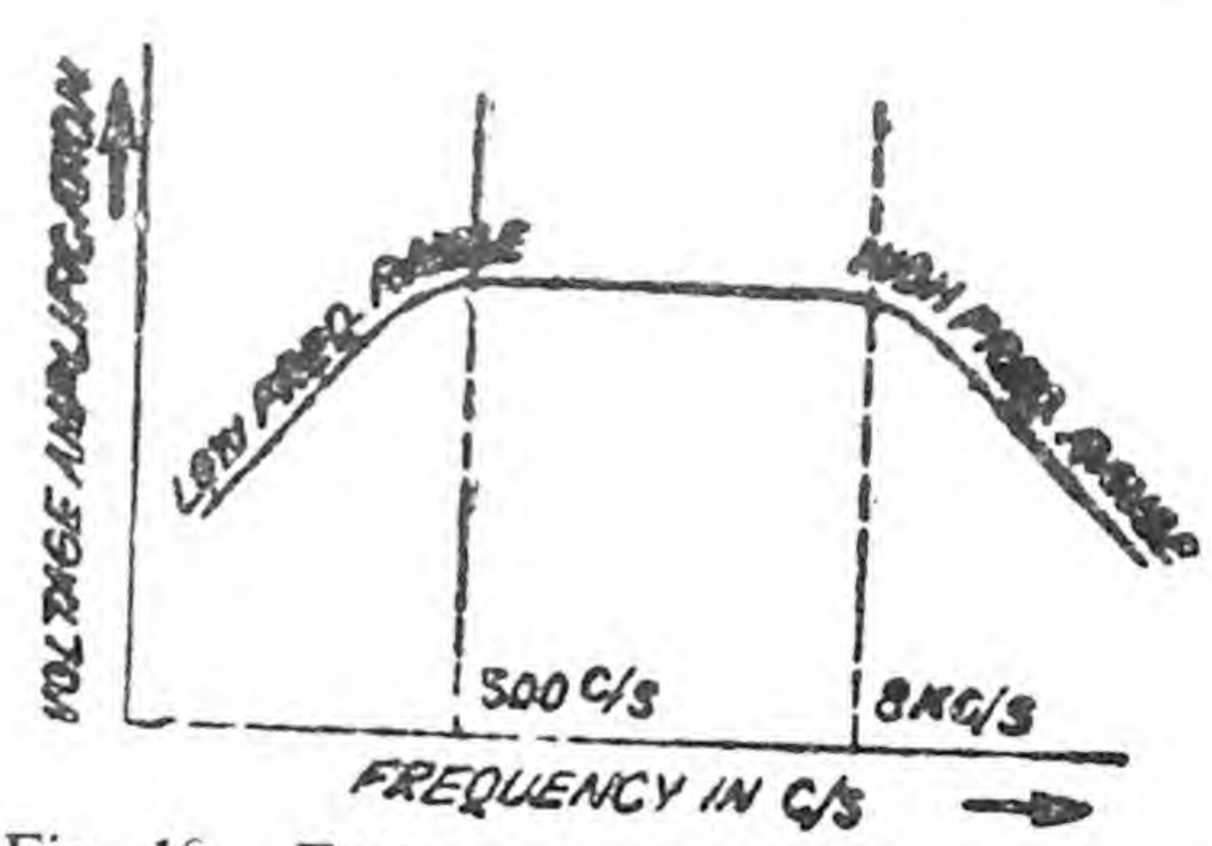


Fig. 19. Frequency response curve of an impedance coupled amplifier.

10·9. TRANSFORMER COUPLED AMPLIFIER :

In case of *RC* coupled and impedance coupled amplifiers, the amplification per stage has been limited to a value approaching to

μ of the tube. In order to increase the gain, it is possible to use a suitable ideal step up transformer. Following are the properties of an ideal transformer :

(*i*) primary and secondary impedances are infinite inductive reactances ; hence there will be no losses by magnetising currents, or in other words winding have no resistance.

(*ii*) whole of the magnetic flux of one winding links with the other winding, *i.e.*, coefficient of coupling is unity. Further

$$M=\sqrt{(L_p L_s)}$$

and the transformer ratio is

$$n=\sqrt{\left(\frac{L_s}{L_p}\right)}.$$

(*iii*) Impedance of the secondary load as viewed from the primary terminals of the transformer is $\frac{Z_s}{n^2}$, *i.e.* when a resistance R is connected across the secondary of a perfect transformer of turns ratio n, its effect on the primary circuit is the same as if the secondary circuit were omitted, and a resistance equal to R/n^2 were connected in parallel with the primary.

Audio amplifier circuits using a low frequency (iron core) transformer as the coupling unit between two successive stages are called transformer coupled audio amplifiers. In the transfosmer coupled amplifier the load impedance of the plate circuit is supplied by a step up transformer which delivers its secondary voltage to the grid of another amplifier. The output of one valve is transferred to the second through electromagnetic induction and the secondary voltage is n times the primary voltage where n is the ratio of number of turns in the secondary to the number of turns in the primary. The transformer serves the following purposes :

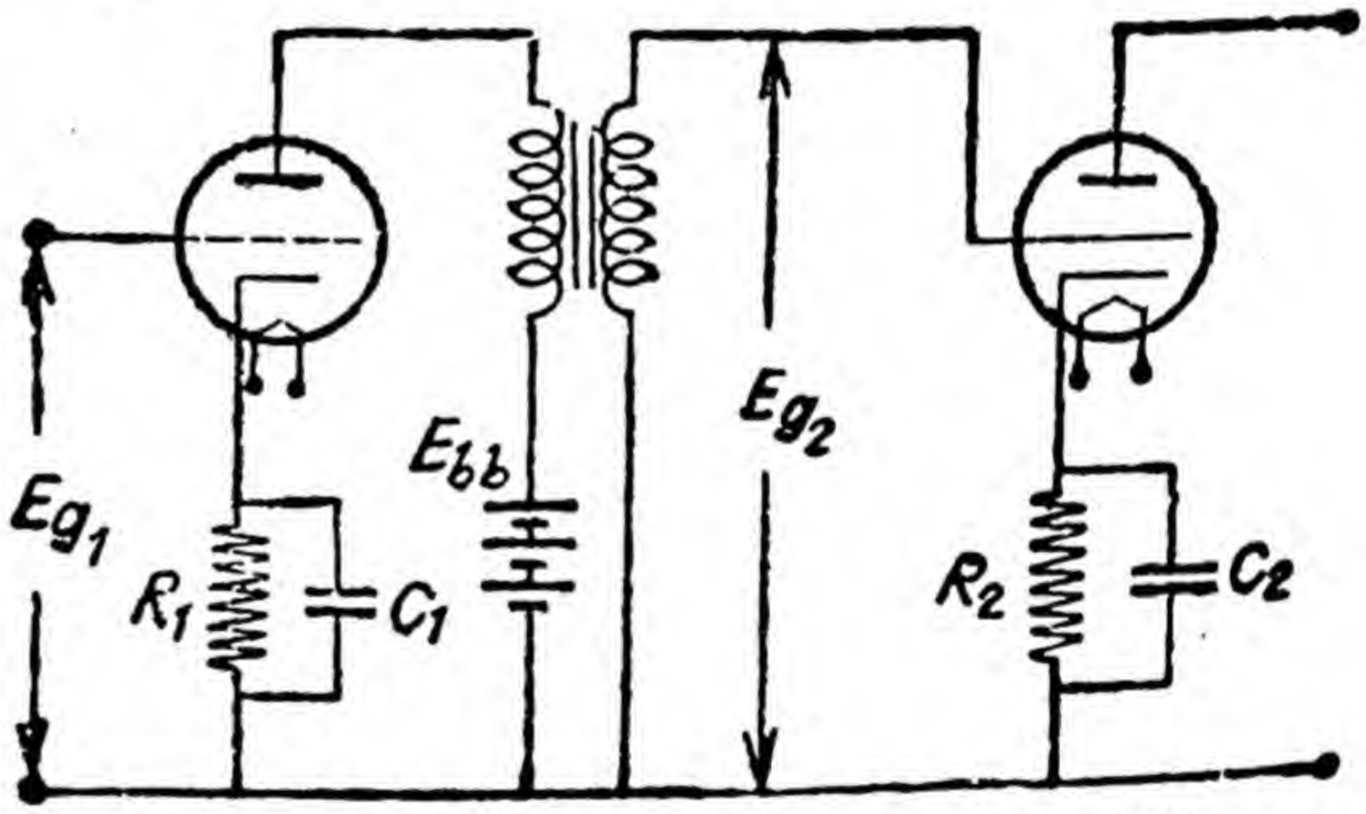


Fig. 20. Transformer-coupled amplifier.

(*i*) it isolates the d.c. components of voltage and current in the adjacent stages,

(*ii*) it couples the a.c. output of one stage to next stage, and

(*iii*) it provides the simple means for stepping up or stepping down voltages etc.

Analysis of a transformer coupled amplifier :

Figure 21 shows the equivalent circuit of an ideal transformer :

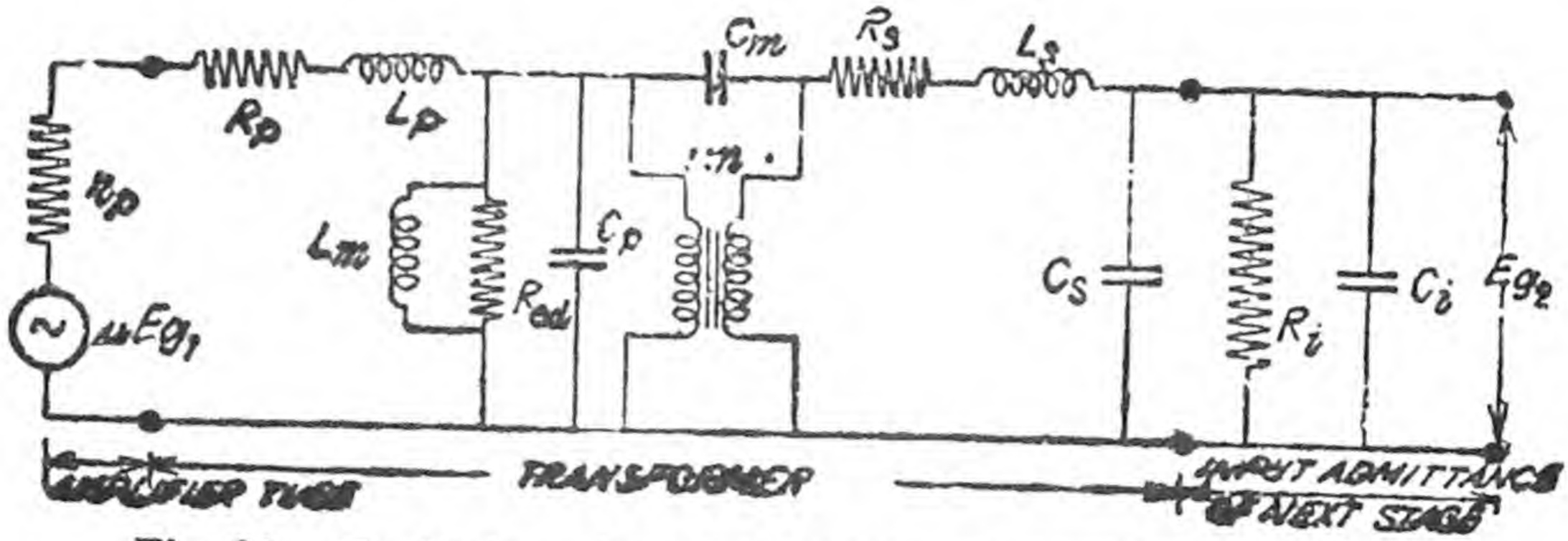


Fig. 21. Equivalent circuit of transformer coupled amplifier.

The name of circuit components are given below :

R_p=effective resistance of primary winding
L_p=primary leakage inductance
L_m=magnetising inductance of primary
R_{ed}=resistance due to eddy current loss
C_p=distributed capacity of the primary winding of the transformer
C_m=winding capacity of the primary and the secondary
R_s=effective resistance of secondary winding
L_s=secondary leakage inductance
C_s=secondary distributed capacity
R_i=input resistance of next stage
C_i=input capacitance of next stage
n=ratio between the number of turns in the secondary and the primary=(N_s/N_p).

Let C' represents the equivalent capacity of C_m placed across the secondary. The value of C' can be determined by considering the stored energy as follows :

Let E_1 is voltage acoss the primary,
E_2 is voltage across the secondary, and
E_m is voltage across the C_m.

Then $$E_m=E_1\pm E_2=E_1\,(1\pm n),$$

positive sign is used when voltages are additive and negative when subtractive.

The energy stored across the capacity C_m per cycle

$$=\tfrac{1}{2}C_mE_m^2$$

or $$=\tfrac{1}{2}C_mE_1^2\,(1\pm n)^2. \qquad \ldots(1)$$

Considering the capacity C' in the secondary, the energy stored is

$$=\tfrac{1}{2}C'\,E_2^2=\tfrac{1}{2}C'\,n^2E_1^2. \qquad \ldots(2)$$

Comparing equations (1) and (2), we get

$$C'=\left(\frac{1\pm n}{n}\right)^2 C_m. \qquad \ldots(3)$$

In other words, we can say that instead of considering the capacity C_m we can consider a capacity $[(1 \pm n)/n]^2 C_m$ in the secondary.

The equivalent circuit of fig. 21 is complicated, hence approximate analysis is made on the following assumptions :

(*i*) In a well designed transformer R_{cd} is usually large in comparison to R_p of the transformer primary and hence it, being in parallel with R_p, can be omitted.

(*ii*) R_i in shunt with the secondary is usually large and hence omitted.

(*iii*) C_p is also omitted since its value is far smaller than the capacity reflected from the secondary to the primary.

The equivalent simplified circuit of fig. 21 as viewed from primary is shown in fig. 22.

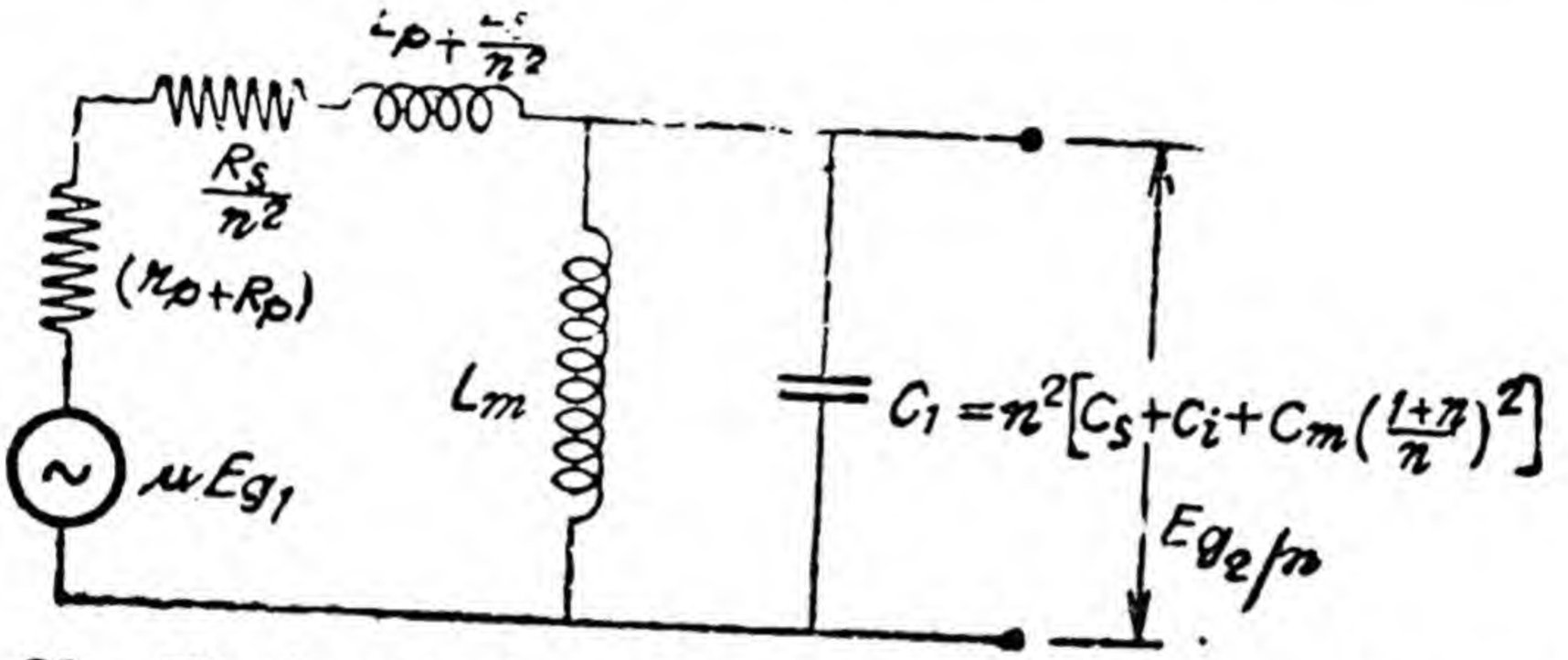


Fig. 22. Final equivalent circuit of transformer coupled amplifier.

Now we shall consider the performance of the amplifier for low, middle and high frequency ranges.

(1) Low frequency range : The low frequency range lies much below the parallel resonant frequency of L_m and C_1. Following points are observed in this range :

(*a*) the reactance of C_1 is so high that it constitutes open circuit and hence it can be omitted.

(*b*) Now the current flowing through

$$\frac{R_s}{n^2} \text{ and } \left(L_p + \frac{L_s}{n^2} \right)$$

is negligible; so the voltage drop across them may be neglected.

The equivalent circuit for this range is now shown in the fig 23.

The plate current $\mathbf{I}_p$ is given by

$$\mathbf{I}_p = \frac{-\mu \mathbf{E}_{g_1}}{(r_p + R_p) + j\omega L_m}$$

Voltage across L_m

$$= j\omega L_m \mathbf{I}_p$$

$$= \frac{-\mu \mathbf{E}_{g1}\, j\omega L_m}{(r_p + R_p) + j\omega\, L_m} \quad \ldots(4)$$

This voltage must be equal to $+\mathbf{E}_{g2}/n$, hence

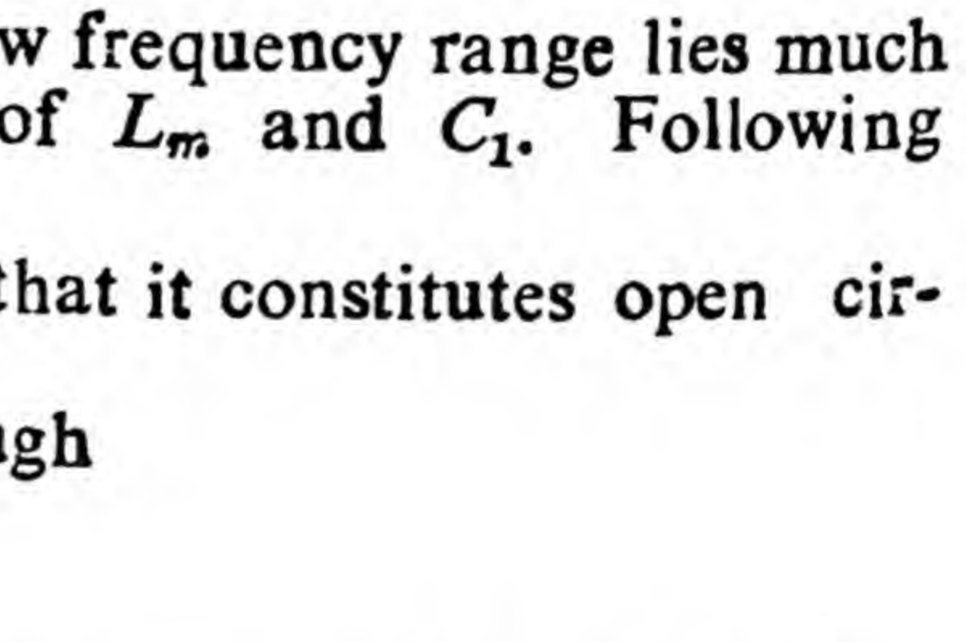


Fig. 23. Equivalent circuit of transformer coupled amplifier in low frequency range.

$$\frac{-\mu \mathbf{E}_{g_1}\, j\omega L_m}{(r_p+R_p)+j\omega L_m} = +\frac{\mathbf{E}_{g_2}}{n}. \qquad \ldots(5)$$

The complex voltage gain at low frequency is given by

$$\mathbf{A}_l = \frac{\mathbf{E}_{g_2}}{\mathbf{E}_{g_1}} = -\mu n \frac{j\omega L_m}{(r_p+R_p)+j\omega L_m}$$

$$= -\frac{\mu n}{1-\dfrac{j\,(r_p+R_p)}{\omega L_m}}. \qquad \ldots(6)$$

The magnitude of $\mathbf{A}_l$ is given by

$$|\,\mathbf{A}_l\,| = \frac{\mu n}{\sqrt{\left(1+\dfrac{(r_p+R_p)^2}{(\omega L_m)^2}\right)}}, \qquad \ldots(7)$$

where the angle $\phi = \tan^{-1}\dfrac{r_p+R_p}{\omega L_m}$.

If we define a frequency f_1 as

$$2\pi f_1 L_m = r_p + R_p, \qquad \ldots(8)$$

then the magnitude of A_l is given by

$$A_l = \frac{\mu n}{\sqrt{\{1+(f_1/f)^2\}}}, \qquad \ldots(9)$$

where f_1 is known as lower half frequency.

(2) Middle frequency range :

From equation (7) it is clear that when $\omega L_m = r_p + R_p$, the voltage gain is $\dfrac{\mu n}{\sqrt{2}}$. Middle frequency range starts from frequency f_1 upwards. In middle frequency range,

(*a*) the frequency is large enough to make reactance ωL_m quite large, so it can be omitted from the equivalent circuit,

(*b*) reactance offered by C_1 is so high that it can be regarded practically equivalent to an open circuit and may be omitted from equivalent circuit.

The equivalent circuit in middle frequency range now becomes as shown in fig. 24.

In middle frequency range, no current flows beyond a negligible core loss current and hence no voltage is lost. Therefore

$$-\mu \mathbf{E}_{g_1} = +\mathbf{E}_{g_2}/n. \qquad \ldots(10)$$

The voltage gain is given by

$$\mathbf{A}_m = \frac{\mathbf{E}_{g_2}}{\mathbf{E}_{g_1}} = -\mu n. \qquad \ldots(11)$$

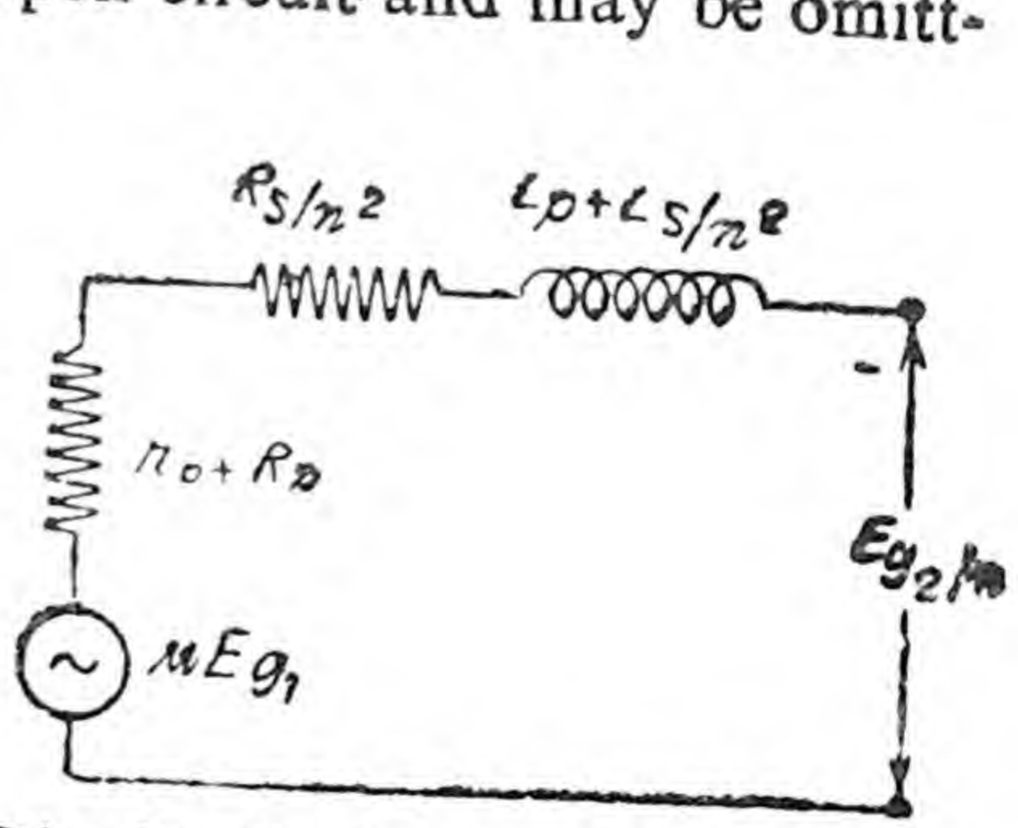


Fig. 24. Equivalent circuit in middle frequency range.

(3) High frequency range :

In high frequency range

(*a*) reactance of L_m is very high so that it is practically equivalent to an open circuit and may be omitted.

(*b*) reactance of C_1 is small at high frequencies and hence can not be omitted.

The equivalent circuit is shown in fig. 25.

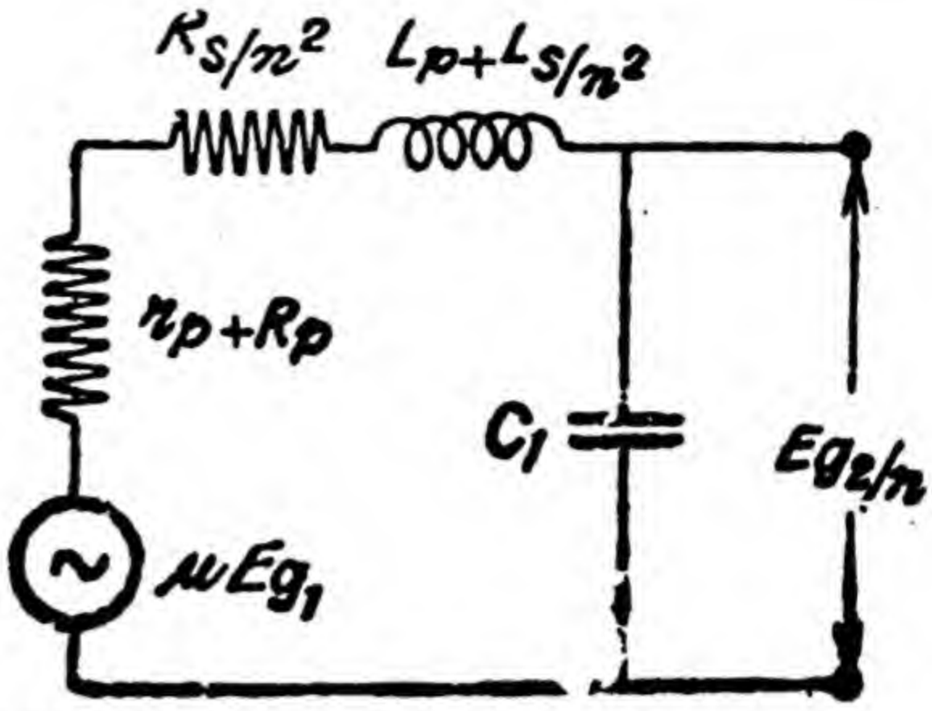


Fig. 25. Equivalent circuit of transformer coupled amplifier at high frequency range.

The plate current is given by

$$\mathbf{I}_p = \frac{-\mu \mathbf{E}_{g_1}}{\left(r_p + R_p + \frac{R_s}{n^2}\right) + j\left[\omega\left(L_p + \frac{L_s}{n^2}\right) - \frac{1}{\omega C_1}\right]} \quad \ldots(12)$$

Voltage across C_1 is given by

$$= \frac{-\mu \mathbf{E}_{g_1} \frac{1}{j\omega C_1}}{\left(r_p + R_p + \frac{R_s}{n^2}\right) + j\left[\omega\left(L_p + \frac{L_s}{n^2}\right) - \frac{1}{\omega C_1}\right]} \quad \ldots(13)$$

This voltage should be equal to $\mathbf{E}_{g_2}/n$, hence

$$\frac{\mathbf{E}_{g_2}}{n} = \frac{-\mu \mathbf{E}_{g_1} \frac{1}{j\omega C_1}}{\left(r_p + R_p + \frac{R_s}{n^2}\right) + j\left[\omega\left(L_p + \frac{L_s}{n^2}\right) - \frac{1}{\omega C_1}\right]}$$

$$\therefore \mathbf{A}_h = \frac{\mathbf{E}_{g_2}}{\mathbf{E}_{g_1}} = \frac{-\mu n \frac{1}{j\omega C_1}}{\left(r_p + R_p + \frac{R_s}{n^2}\right) + j\left[\omega\left(L_p + \frac{L_s}{n^2}\right) - \frac{1}{\omega C_1}\right]} \quad \ldots(14)$$

Due to the reduction of reactance $\frac{1}{\omega C_1}$ and increase of reactance $\omega\left(L_p + \frac{L_s}{n^2}\right)$ at high frequencies, the voltage gain decreases. The frequency of this resonance is given by f_2

$$f_2 = \frac{1}{\pi \sqrt{\left\{\left(L_p + \frac{L_s}{n^2}\right) C_1\right\}}}$$

Frequency Response Curve : Figure 26 shows the variation of voltage gain with frequency in a transformer coupled audio amplifier circuit. Curve *A* represents the voltage gain of the circuit while using a high grade transformer, and curve *B* shows the voltage gain of same circuit while using a low grade transformer.

Insufficient primary inductance is responsible for the drop in the curve at low frequencies and for a high grade transformer in which primary inductance has been increased gives a better low frequency response.

At high frequencies the self capacitance of the secondary winding and the leakage inductance of the primary acts as a series resonant circuit. Thus they produce a peak in the resonant curve.

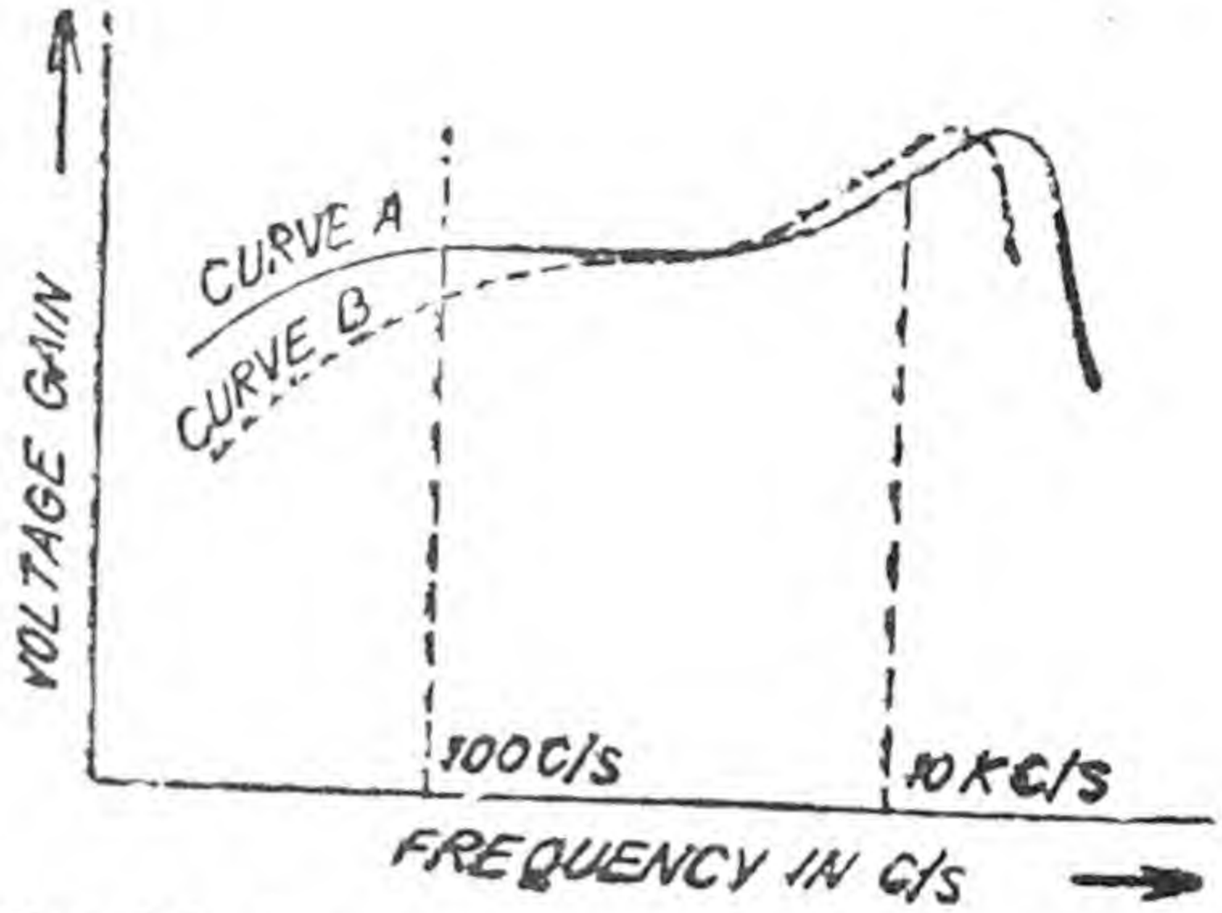


Fig. 26. Frequency response curve of transformer coupled amplifier.

The height of the resonant peak can be materially reduced by increasing the secondary resistance. A high resistance of 0·5 to 10 mega-ohms shunted across the secondary is effective in removing the resonant peak from the characteristic. The amplification is more nearly constant over midfrequency range.

Advantages :

(*i*) They produce higher amplification per stage than the impedance coupled amplifier. In this case very high value of coupling inductance must be used.

(*ii*) The voltage drop across the transformer winding is negligible, thus making it possible to use a low voltage (high tension) battery.

(*iii*) In transformer coupling high frequency cut-off is very sharp.

Disadvantages :

Disadvantages are the large size and the high cost of transformers required to produce high fidelity.

NUMERICAL EXAMPLES

1. *Determine the gain of a resistance-capacitance coupled audio frequency amplifier stage at frequencies* 50 *and* 1000 *kc/s. given the following data :*

Voltage amplification factor $=200$,
Valve anode resistance $=100\ K\Omega$,
Anode load resistance $=250\ K\Omega$,
Capacitance of coupling capacitor $=0{\cdot}01\ \mu F$,
Grid leak resistance $=0{\cdot}5\ M\Omega$.

The frequency 1000 may be considered as middle frequency hence the stage gain is given by

$$A_m = -\frac{\mu R_l R_0}{r_p R_l + r_p R_c + R_l R_0}$$

$$= -\frac{200 \times 250 \times 10^3 \times 0{\cdot}5 \times 10^6}{100 \times 10^3 \times 250 \times 10^3 + 100 \times 10^3 \times 2{\cdot}5 \times 10^6 + 250 \times 10^3 \times 0{\cdot}5 \times 10^6}$$

$$= -\frac{25000 \times 10^9}{2000 \times 10^6}$$

$$= -125$$

or in decibels $= -20 \log_{10} 125 = -$ **42db.**

At 50 cycles/sec. (low frequency),

$$\frac{\text{stage gain at middle frequency}}{\text{stage gain at low frequency}} = 1 + \frac{1}{j\omega CR} = 1 - \frac{j}{\omega CR},$$

where $$R = R_c + \frac{r_p . R_l}{R_l + R_p}$$

$$= 0\ 5 \times 10^6 + \frac{100 \times 10^3 \times 250 \times 10^3}{100 \times 10^3 + 250 \times 10^3} = \frac{40}{3} \times 10^5.$$

Hence $$\frac{\text{stage gain at middle frequency}}{\text{stage gain at low frequency}}$$

$$= 1 - \frac{7j}{2 \times 3{\cdot}14 \times 50 \times 0{\cdot}01 \times 10^{-6} \times 40 \times 10^5}$$

$$= 1 - 0{\cdot}5585j$$

or $$\frac{\text{stage gain at low frequency}}{\text{stage gain at middle frequency}} = \frac{1}{1 - 0{\cdot}5585j}$$

$$= 0{\cdot}76 + 0{\cdot}4252\ j$$

or amplitude $$= \sqrt{\{(0{\cdot}76)^2 + (\cdot 4252)^2\}}$$

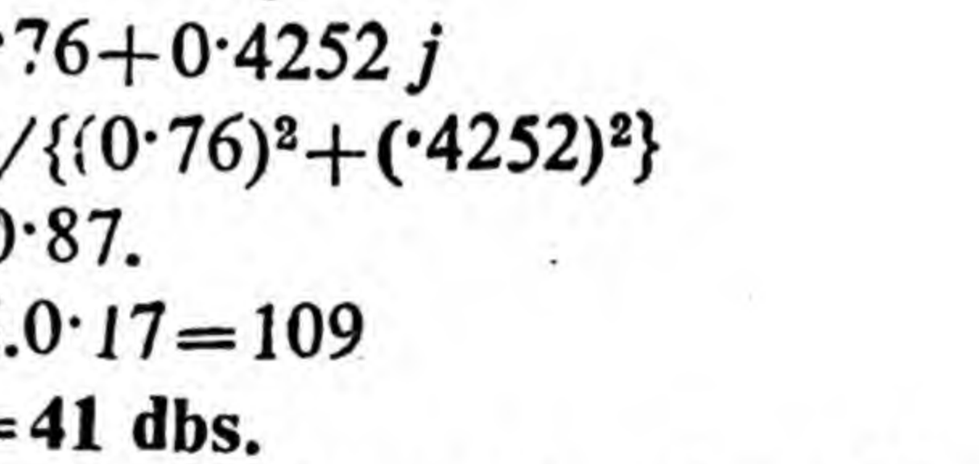
$$= 0{\cdot}87.$$

Stage gain at low frequency $= 115.0{\cdot}17 = 109$

in dbs. $= 20 \log_{10} 109 =$ **41 dbs.**

2. *A valve has an anode slope resistance of 20,000Ω and μ=10. There is a non-inductive resistance of 1,00,000Ω shunted by a capacitance of 100 μμF, in the anode circuit. Calculate the amplification for the voltage of frequency 100 kc/s.*

The equivalent circuit of the given question may be drawn as shown in adjacent figure.



Voltage amplification

$$= -\frac{\mu R_l}{r_p + R_l\,[1 + j\omega C r_p]}.$$

Here $\mu = 10$.

$$R_l = 10.10^4\ \Omega, \quad C = 100\ \mu\mu F = 100 \times 10^{-12}\ F,$$

$$r_p = 2.10^4\ \Omega, \quad \omega = 2.3{\cdot}14.100.10^3.$$

The voltage amplification

$$= -\frac{10.10.10^4}{2 \times 10^4 + 10 \times 10^4\ (1 + j\ 2 \times 3{\cdot}14 \times 100.10^3 .100 \times 10^{-12} .2 \times 10^4)}$$

$$= -\frac{10.10.10^4}{12\times10^4+j\times2\times3\cdot14\times100\times100\times2\times10\times10^{11}\times10^{-12}}$$

$$= -\frac{100.10^4}{12\times10^4+j\ 2\times3\cdot14\times10^4\times2}$$

$$= -\frac{100}{12+12\cdot56j}.$$

Amplification

$$= \frac{100}{\sqrt{\{144+(12\ 56)^2\}}} = \frac{100}{\sqrt{(144+157\cdot65)}}$$

$$= \frac{100}{17\cdot37} = 5\cdot75.$$

3. *Determine the gain of a voltage amplifier with a triode of constant $\mu=48$, $g_m=12$ mA/volt and with an anode load resistance of 160 kilo-ohms for a frequency of 15 kc/s. The output is fed by a coupling condenser of negligible resistance to a grid resistor of 380 kilo ohms and total shunt capacitance of 90 $\mu\mu F$.*

Voltage amplification at high frequency

$$= -\frac{\mu R_l R_c}{r_p R_c + r_p R_l + R_c\ R_l(1+j\omega C r_p)}$$

$$= -\frac{\mu}{\left[\left\{1+r_p\left(\frac{R_c+R_l}{R_l.R_c}\right)\right\}^2+\{\omega^2C^2r_p^2\}\right]^{1/2}}.$$

Here $\mu=48$,

$$r_p=\frac{\mu}{g_m}=\frac{48}{1\cdot2\times10^{-3}}=40000\ \Omega,$$

$\omega=2\pi.15.10^3$,

$C=90\times10^{-12}\ F$.

$R_l=160$ kilo ohms $=160\times10^3$ ohms,

$R_c=380\times10^3$ ohms.

$$\text{V.A.} = -\frac{48}{\left[\left\{1+40{,}000\left(\frac{160\times10^3+380\times10^3}{380\times160\times10^6}\right)\right\}^2+\{(2\times3\cdot14\times15\times10^3)^2.(90\times10^{-12})^2\ (40{,}000)^2\}\right]^{1/2}}$$

$$= -\frac{48}{\left[\left\{1+40{,}000\left(\frac{540}{380\times160\times10^3}\right)\right\}^2+49766\cdot4\times10^{-8}\right]^{1/2}}$$

$$= -\frac{48}{[(1\cdot35)^2+49766\cdot4\times10^{-8}]^{1/2}}$$

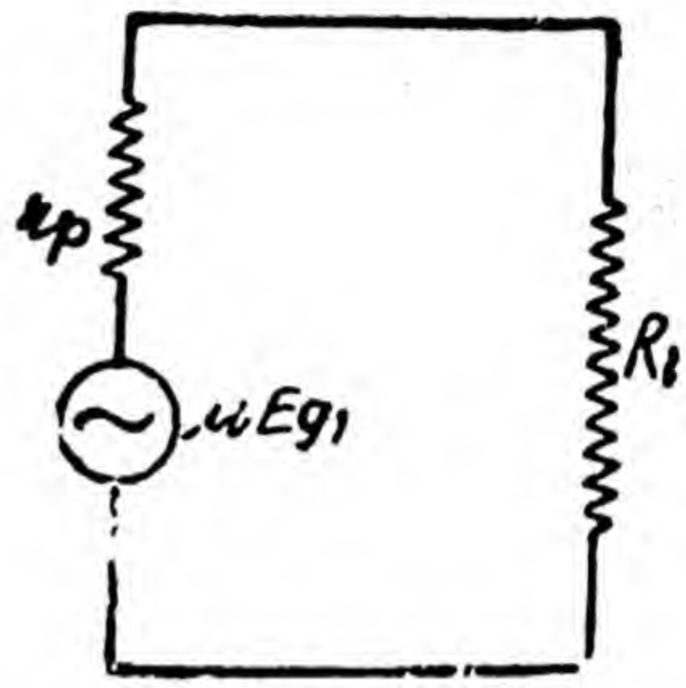
$$= -\frac{48}{1\cdot354} = -35\cdot6.$$

4. *A three stage RC coupled amplifier has the following constants :*

plate resistance $r_p = 10{,}000$
valve amplification $= 15$,
external resistance $= 80{,}000\Omega$,
loss in coupling condenser $= 5\%$.

Find the voltage across the last output resistance for an input of 3000 *microvolts.*

Let the following circuit represent the first stage of the amplifier.



Voltage across R_l

$$= \frac{\mu Eg_1 R_l}{R_l + r_p}$$

$$= \frac{15 \times 3 \times 10^{-3} \times 80 \times 10^3}{10 \times 10^3 + 80 \times 10^3}$$

$$= \frac{45 \times 80 \times 10^{-3}}{90}$$

$$= 40 \times 10^{-3}.$$

Now we take into account the loss across coupling condenser

40×10^{-3} would include a loss of $\frac{5 \times 40 \times 10^{-3}}{100}$

$\therefore$ loss in condenser $= 2 \times 10^{-3}$.

Input for the second stage $= 40 \times 10^{-3} - 2 \times 10^{-3}$

$= 38 \times 10^{-3}$ volts.

Drop across second stage $R_l = \frac{15 \times 38 \times 10^{-3} \times 80 \times 10^3}{90 \times 10^3}$

$$= \frac{38 \times 80 \times 10^{-3}}{6}$$

$$= \frac{19 \times 80 \times 10^{-3}}{3} \text{ volts.}$$

Loss in the second stage $= \frac{5}{100} \times \frac{19 \times 80 \times 10^{-3}}{3}$

$$= \frac{76 \times 10^{-3}}{3}$$

Hence input for the third stage $= \frac{19 \times 80 \times 10^{-3}}{3} - \frac{76 \times 10^{-3}}{3}$

$$= \frac{(19 \times 80 - 76) \times 10^{-3}}{3}$$

$$= \frac{1444 \times 10^{-3}}{3} \text{ volts.}$$

For third stage voltage drop $= \frac{15 \times 1444 \times 10^{-3} \times 80 \times 10^3}{3 \times 90 \times 10^3}$

$$= \frac{1444 \times 40}{9} \times 10^{-3} \text{ volts.}$$

$$\text{Loss in this stage} = \frac{5}{100} \times \frac{1444 \times 40}{9} \times 10^{-3} \text{ volts.}$$

$$= \frac{2888 \times 10^{-3}}{9} \text{ volts.}$$

$$\text{Input for the next stage} = \left[\frac{1444 \times 40}{9} - \frac{1444 \times 2}{9} \right] 10^{-3} \text{ volts.}$$

$$= \frac{1444}{9} \times 38 \times 10^{-3}$$

$$= 160{\cdot}44 \times 38 \times 10^{-3} \text{ volts.}$$

5. *The slope of the anode current grid potential curve of a three electrode valve is 1·05 mA/V. and the slope of the anode current anode potential curve is 0·04 mA/V. What voltage amplification will the valve give across an anode resistance af 15000Ω ?*

We know that the slope of the anode current to grid potential *i.e.*, $\frac{\partial I_p}{\partial V_g}$ is equal to mutual transconductance (g_m) *i.e.*

$$g_m = 1{\cdot}05 \text{ mA./V.}$$
$$= 1{\cdot}05 \times 10^{-3} \text{ A./V.}$$

Also the slope of the anode current anode potential curve gives the admittance, *i.e.*

$$\frac{1}{r_p} = 0{\cdot}40 \times 10^{-3}.$$

$$r_p = \frac{10^3}{0{\cdot}40} = 2500\Omega.$$

Now again $\mu = r_p \times g_m = 1{\cdot}05 \times 10^{-3} \times 2500 = 2{\cdot}625.$

$$\text{Amplification or stage gain} = \frac{\mu R_l}{R_l + r_p}$$

$$= \frac{2{\cdot}625 \times 15000}{15000 + 2500}$$

$$= \mathbf{2{\cdot}25.}$$

6. *A two stage RC coupled amplifier uses triode valve in the first stage of amplification factor 20 and plate resistance 20KΩ. Load resistance is 100KΩ. Grid leak resistance of the second stage is 1MΩ. Calculate the value of the coupling condenser so that the stage gain at 50 c/s is 2 db below the medium frequency range voltage gain.*

The stage gain of *RC* amplifier at low frequencies is given by

$$A_l = - \frac{\mu R_l R_c}{r_p R_l + r_p R_c + R_l R_c + \frac{(r_p + R_l)}{j\omega C}},$$

The stage gain at mid frequencies is given by

$$A_m = - \frac{\mu R_l R_c}{{}_p R_l + r_p R_c + R_l R_c}.$$

$$\therefore \quad \frac{\mathbf{A}_l}{\mathbf{A}_m}=\frac{1}{1+\frac{1}{\omega C}\left[\frac{r_p+R_l}{r_pR_l+r_pR_c+R_lR_c}\right]}, \qquad \ldots(1)$$

$$\left|\frac{\mathbf{A}_l}{\mathbf{A}_m}\right|=\frac{1}{\sqrt{\left\{1+\frac{1}{(\omega CR_e)^2}\right\}}},$$

where $R_e=\dfrac{r_pR_l+r_pR_c+R_lR_c}{r_p+R_l}$

$$=\frac{20\times10^3\times100\times10^3+20\times10^3\times1\times10^6+100\times10^3\times10^6}{20\times10^3+100\times10^3}$$

$=101{\cdot}667\times10^4$ ohms.

At 50 c/s, the gain is $2db$ below the medium frequency range gain

$$\therefore \quad 2=20\log_{10}\left|\frac{A_m}{A_l}\right|$$

or $$\left|\frac{A_m}{A_l}\right|=\text{Anti log}\frac{2}{20}=\text{Anti log } 0{\cdot}1=1{\cdot}259. \qquad \ldots(2)$$

Substituting equation (2) in equation (1), we have

$$1{\cdot}259=\sqrt{\left\{1+\frac{1}{(\omega CR_e)^2}\right\}}$$

or $$(1{\cdot}259)^2=1+\frac{1}{(\omega CR_e)^2}$$

or $$\frac{1}{(\omega CR_e)^2}=(1{\cdot}259)^2-1=0{\cdot}58$$

or $$\frac{1}{\omega CR_e}=\sqrt{\{0{\cdot}58\}}=0{\cdot}76$$

or $$\frac{1}{2\pi fC\times101{\cdot}667\times10^4}=0{\cdot}76.$$

Solving it for $f=50$ c/s, we have $\mathbf{C=0{\cdot}004}\mu f$.

7. A triode is used in RC coupled amplifier. The load resistance is 30 kilo-ohms, coupling condenser is 0·005 μf, grid leak resistance of the next stage is 1 Meg. ohm., shunt capacity is 100 PF, plate resistance of the valve is 8,000 ohms and its amplification factor is 20. Calculate (a) the voltage gain of the amplifier at the medium frequency range, (b) lower and upper half power frequencies (c) value of the coupling condenser so as to have a lower half power frequency of 15 c/s.

(a) Voltage gain of the amplifier at mid frequency range is

$$\mathbf{A}_m=-\frac{\mu R_lR_c}{r_pR_l+r_pR_c+R_lR_c}=-\frac{\mu}{\frac{r_p}{R_c}+\frac{r_p}{R_l}+1}$$

$$\therefore \quad A_m=\frac{-20}{\frac{8000}{1\times10^6}+\frac{8000}{30\times10^3}+1}=-15{\cdot}4$$

(b) Lower half power frequency is given by

$$f_1=\frac{1}{2\pi C}\left[\frac{1}{\frac{r_pR_l}{r_p+R_l}+R_c}\right] \text{ (equation 2 art. } 10{\cdot}2)$$

$$=\frac{1}{2\pi\times{\cdot}005\times10^{-6}}\left[\frac{1}{\frac{8000\times30\times10^3}{8000+30\times10^3}+1\times10^6}\right]$$

$$=31{\cdot}64 \text{ c/s.}$$

Upper half power frequency is given as

$$f_2=\frac{1}{2\pi C_g}\left(\frac{1}{r_p}+\frac{1}{R_l}+\frac{1}{R_c}\right) \text{ (see equation 7 art. } 10{\cdot}2)$$

$$=\frac{1}{2\pi\times100\times10^{-12}}\left(\frac{1}{8000}+\frac{1}{30\times10^3}+\frac{1}{10^6}\right)$$

$$=25{\cdot}11\times10^4 \text{ c/s.}$$

(c) $$C=\frac{1}{2\pi f_1}\left[\frac{1}{\frac{r_pR_l}{r_p+R_l}+R_c}\right]$$

$$=\frac{1}{2\pi\times15}\left[\frac{1}{\frac{8000\times30\times10^3}{8000+30\times10^3}+1\times10^6}\right]$$

$$=\frac{1}{2\times3{\cdot}14\times15\times100{\cdot}63\times10^4}=0{\cdot}105\times10^{-7} \text{ farad}$$

$$=0\ 01\ \mu\text{f.}$$

8. *A resistance capacitance coupled triode amplifier has anode load of* 25 *kilo-ohms, coupling condenser of* 0·01 *micro-farad, dynamic plate resistance of* 20 *kilo-ohms, amplification factor of tube of* 40, *grid leak resistance of* 500 *kilo-ohms and total shunt capacitance of* 100 *micro-micro farad. Calcu ate* (a) *mid band gain* (b) *lower and upper high power angular frequencies,* (c) *magnitude and phase angle of complex vol tage gain at angular frequencies of* 100, 5000 *and* 600×10^3 *radians/sec.*

(a) Voltage gain of the amplifier at mid frequency range is

$$A_m=\frac{-\mu R_lR_c}{r_pR_l+r_pR_c+R_lR_c}=\frac{-\mu}{\frac{r_p}{R_c}+\frac{r_p}{R_l}+1}$$

$$=\frac{-40}{\frac{20\times10^3}{500\times10^3}+\frac{20\times10^3}{25\times10^3}+1}=-21{\cdot}74$$

(b) Lower half power angular frequencies ω_1 is given by

$$\omega_1=\frac{1}{C}\left[\frac{1}{\frac{r_p R_l}{r_p+R_l}+R_c}\right]$$

$$=\frac{1}{0{\cdot}01\times10^{-6}}\left[\frac{1}{\frac{20\times10^3\times25\times10^3}{20\times10^3+25\times10^3}+500\times10^3}\right]$$

$$=195{\cdot}7 \text{ radians/sec.}$$

Upper half power angular frequency ω_2 is given by

$$\omega_2=\frac{1}{C_g}\left(\frac{1}{r_p}+\frac{1}{R_l}+\frac{1}{R_c}\right)$$

$$=\frac{1}{100\times10^{-12}}\left(\frac{1}{20\times10^3}+\frac{1}{25\times10^3}+\frac{1}{500\times10^3}\right)$$

$$=920\times10^3 \text{ radians/sec.}$$

(c) Voltage gain at angular frequency of 100 radians/sec is given by

$$A=\frac{A_m}{1-j\,f_1/f} \quad \text{or} \quad |A|=\frac{A_m}{\sqrt{(1+(f_1/f)^2}}$$

$$\therefore \quad A=\frac{21{\cdot}74}{\sqrt{[1+(195{\cdot}7/100)^2]}}=9{\cdot}9$$

Phase angle $\phi=180°+\tan^{-1} f_1/f$

$$=180°+\tan^{-1}\ 1{\cdot}957=243°55'$$

Angular frequency of 5000 radians/sec lies well within the mid band. Hence voltage gain is equal to A_m *i.e.*, (21·74) and phase angle is 180°.

Angular frequency 600×10^3 radians/sec lies close to the upper half power frequency. Hence voltage gain at this frequency is given by

$$A=\frac{A_m}{1+j\,f/f_2} \quad \text{or} \quad |A|=\frac{A_m}{\sqrt{\{1+(f/f_2)^2\}}}$$

$$\therefore \quad A=\frac{21{\cdot}74}{\sqrt{\left\{1+\left(\frac{600\times10^3}{920\times10^3}\right)^2\right\}}}=18{\cdot}2$$

The phase angle ϕ is given by

$$\phi=180°-\tan^{-1} f/f_2=180°-\tan^{-1}\ (600/920)$$

$$=146°54'$$

9. (a) *Each stage of a* 3 *stage RC amplifier has its lower and upper cut off frequencies to be* 100 *c/s and* 100 *kc/s respectively. All stages are identical. Find the* 3*db bandwidth of the amplifier.*

(b) *The first and the second stage of a two stage RC coupled amplifier have the lower cut off frequencies to be* 100 *cycles and* 200 *cycles respectively. The upper cut off frequencies of them are* 110 *kc/s and* 100 *kc/s respectively. Find the over all* 3*db bandwidth of the amplifier.*

(a) Let f_{1n} be the lower cut off frequency of the cascade amplifier and f_1 be the lower cut off frequency of each individual stage. Then according to equation 2 of article 10·4, we have

$$f_{1n}=\frac{f_1}{(2^{1/n}-1)^{1/2}}=\frac{100}{(2^{1/3}-1)^{1/2}}=196 \text{ cycles.}$$

Let f_{2n} be the upper cut of frequency of the cascade amplifier and f_2 be the upper cut off frequency of individual stage. Then according to equation (3) of article 10·4, we have

$$f_{2n}=f_2\,(2^{1/n}-1)^{1/2}$$
$$=100\times10^3\,(2^{1/3}-1)^{1/2}=100\times10^3\times5\cdot1=51 \text{ kc/s.}$$

The overall bandwidth of the amplifier

$$=f_{2n}-f_{1n}$$
$$=51\times10^3-196=\mathbf{50\cdot804\ kc/s.}$$

(b) We know that $\frac{1}{\sqrt{2}}=\frac{1}{[1+(f_1/f_{1n})^2]^{n/2}}$

$$\therefore\quad \frac{1}{\sqrt{2}}=\frac{1}{\sqrt{\left\{1+\left(\frac{100}{f_{1n}}\right)^2\right\}}}\times\frac{1}{\sqrt{\left\{1+\left(\frac{200}{f_{1n}}\right)^2\right\}}}$$

or
$$2=\left\{1+\left(\frac{100}{f_{1n}}\right)^2\right\}.\left\{1+\left(\frac{200}{f_{1n}}\right)^2\right\}$$

Solving it we have $f_{1n}=210$ cycles.

The upper cut off frequency f_{2n} of the amplifier is given by

$$\frac{1}{\sqrt{2}}=\frac{1}{\sqrt{\left\{1+\left(\frac{f_{2n}}{150}\right)^2\right\}}}\cdot\frac{1}{\sqrt{\left\{1+\left(\frac{f_{2n}}{100}\right)^2\right\}}}$$

or
$$2=\left[1+\left(\frac{f_{2n}}{150}\right)^2\right].\left[1+\left(\frac{f_{2n}}{100}\right)^2\right]$$

Solving we have $f_{2n}=78$ kc/s.

The overall bandwidth of the two stage amplifier

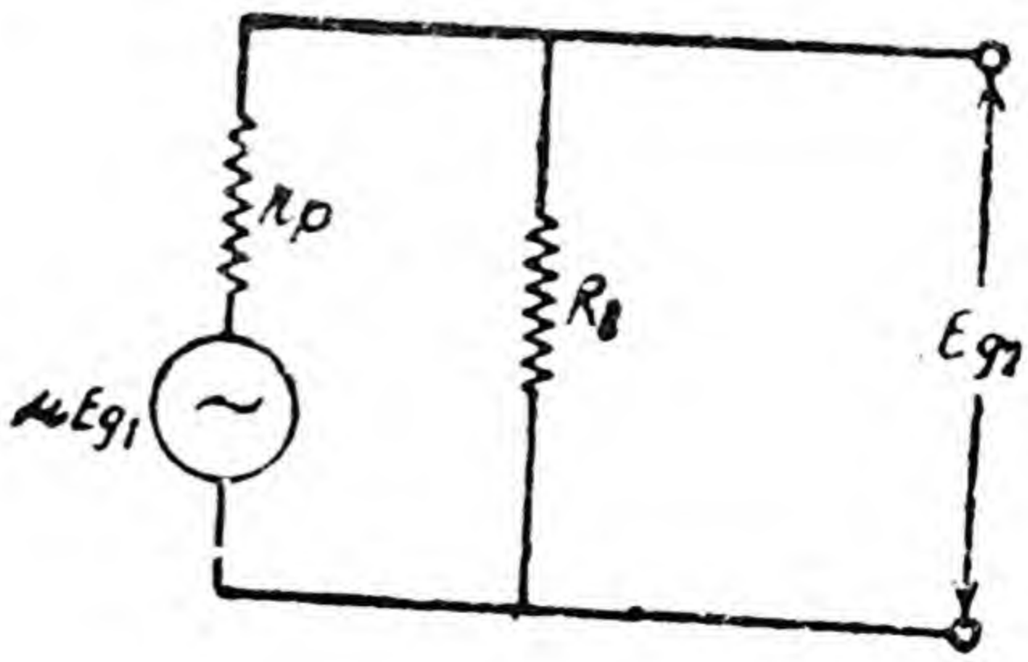
$$\Delta=f_{2n}-f_{1n}=(78-\cdot21)=\mathbf{77\cdot79\ kc/s.}$$

10. *A valve amplifier consists of a valve having an internal anode circuit impedance of* 3000 Ω *and an amplification factor of* 25. *An external impedance is placed in the anode circuit. What is the amplification, if this impedance consists of*

(*a*) *resistance of* 30,000 Ω.

(*b*) *a coil of* 5*mH and* 25 Ω *resistance shunted by a capacitor of* ·001μF, *the amplifier in this case being operated at the resonant frequency of this circuit.*

(a) The equivalent circuit is shown in adjacent figure.

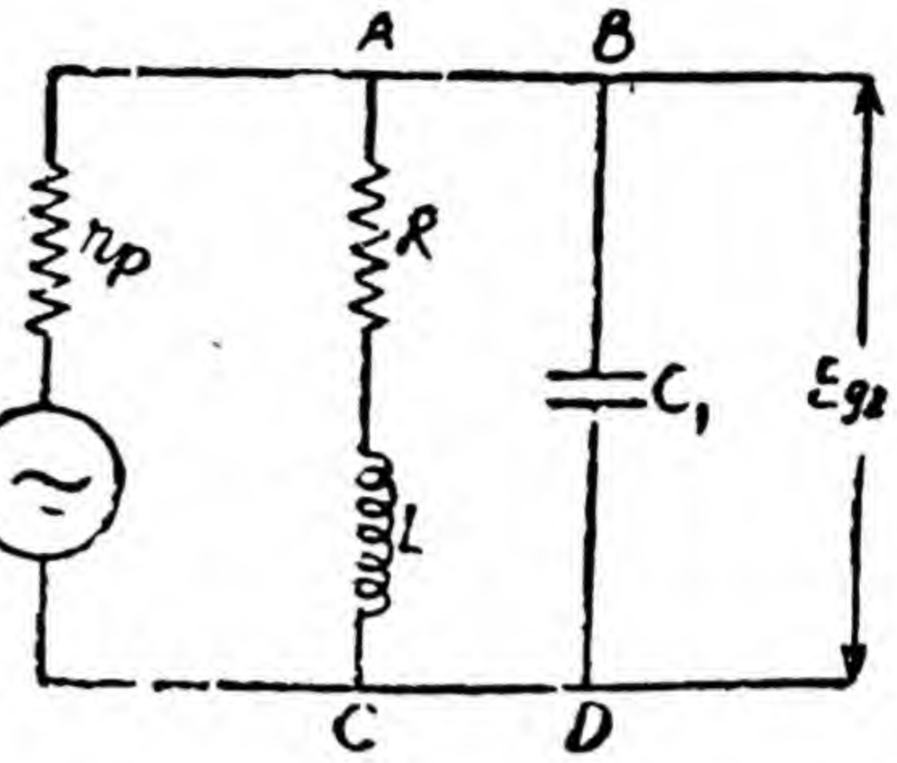


$$\text{Amplification}=-\frac{\mu R_l}{R_l+r_p}$$

$$=-\frac{25\times30{,}000}{30{,}000+30{,}000}$$

$$=-\frac{25}{2}=\mathbf{-12\cdot5.}$$

(b) The equivalent circuit in this case is shown in the following figure :



If Z is the impedance of $ABCD$ then

$$\frac{1}{Z}=\frac{1}{R+j\omega L}+j\omega C$$

$$=\frac{R-j\omega L}{R^2+\omega^2L^2}+j\omega C$$

$$=\frac{R}{R^2+\omega^2L^2}$$

$$+j\omega\left(C-\frac{L}{R^2+\omega^2L^2}\right)$$

At resonance j term should be zero; therefore

$$C-\frac{L}{R^2+\omega^2L^2}=0$$

or

$$C=\frac{L}{R^2+\omega^2L^2},$$

$$R^2+\omega^2L^2=\frac{L}{C},$$

$$\omega^2L^2=\frac{L}{C}-R^2 \text{ or } \omega L=\sqrt{\left(\frac{L}{C}-R^2\right)}.$$

Hence

$$Z=\frac{L}{RC}$$

Now

$$E_{g_2}=\frac{-\mu E_{g_1}Z}{r_p+\frac{L}{RC}}=-\frac{\mu E_{g_1}\frac{L}{RC}}{r_p+\frac{L}{RC}}$$

Amplification will be

$$\frac{E_{g_2}}{E_{g_1}}=-\frac{\mu L}{r_pRC+L}$$

$$=-\frac{25\times5\times10^{-3}}{30\times10^3\times25\times1\times10^{-9}+5\times10^{-3}}=-\frac{25\times5}{30\times25\times0\cdot001+5}$$

$$=-\frac{125}{5\cdot75}=-21\cdot74.$$

11. *A valve having an amplification factor of 7 and an anode slope resistance of 7000Ω is to be used as a low frequency amplifier. The anode circuit transformer has a transformation ratio of 3, and its low resistance winding has an inductance of 10H. Calculate the amplification for voltage of frequency 100 c/s.*

$$\mathbf{I}_p = -\frac{\mu E g_1}{r_p + j\omega L}$$

$$= -\frac{\mu E g_1}{\sqrt{(r_p^2 + \omega^2 L^2)}}$$

$$\therefore \quad \mathbf{E}g_2 = \mathbf{I}_p \, j\omega L$$

$$= -\frac{\mu \mathbf{E} g_1 \omega L}{\sqrt{(r_p^2 + \omega^2 L^2)}}$$

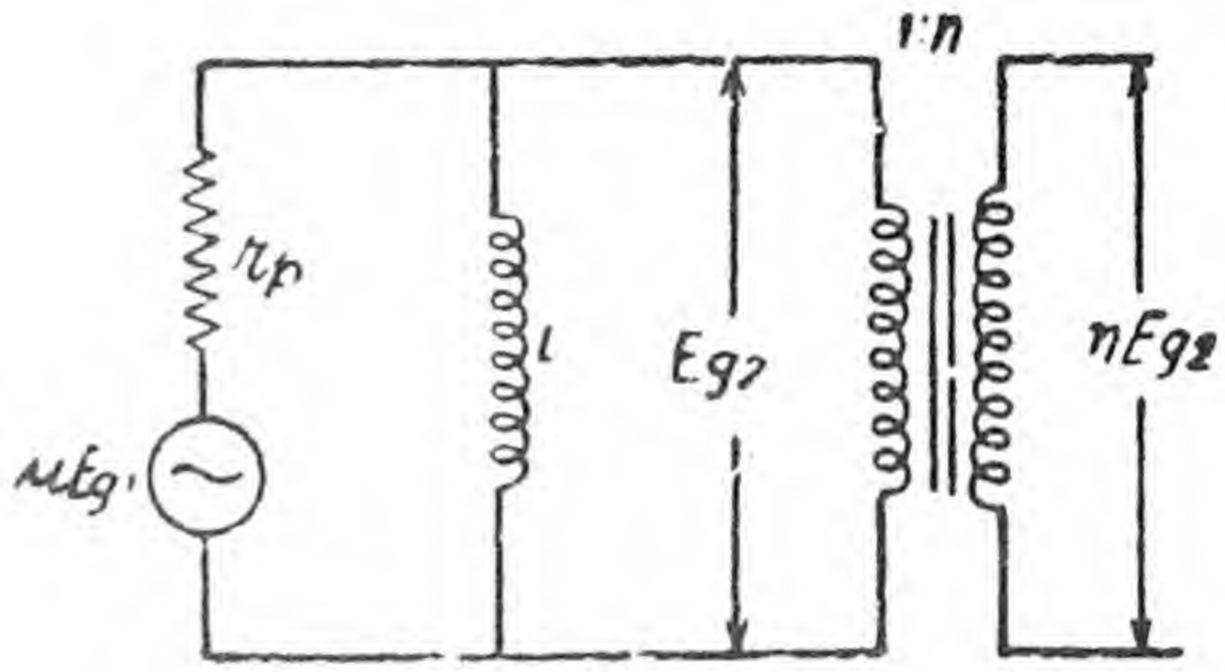
Amplification

$$= -\frac{n\mathbf{E}g_2}{\mathbf{E}g_1} = -\frac{n\mu\omega L}{\sqrt{(r_p^2 + \omega^2 L^2)}}$$

$$= -\frac{n\mu}{\sqrt{\left\{1 + \left(\frac{r_p}{\omega L}\right)^2\right\}}} = -\frac{n\mu}{\sqrt{\left\{1 + \frac{r_p^2}{(2 \times 3{\cdot}14 \times 100)^2 \times (10)^2}\right\}}}$$

$$= -\frac{3 \times 7}{\sqrt{\left\{1 + \left(\frac{7000}{628 \times 10}\right)^2\right\}}}$$

$$= -14.$$



12 *A transformer coupled amplifier has a triode with amplification factor* 20 *and plate resistance* 10,000 *ohm. Effective primary resistance*=1000 *ohm, effective secondary resistance*=9000 *ohm; magnetising inductance of primary*=10 *henrys; leakage inductance referred to the primary*=0·1 *henry, effective total capacitance across the primary*=400 *pf. and turn ratio* $\frac{N_s}{N_p}=3$. *Calculate* (*i*) *mid band gain.* (*ii*) *lower half power frequency.* (*iii*) *resonant frequency in high frequency range.* (*iv*) *gain at a frequency of* 20 *cycle per second.*

(*i*) Mid band gain $= \mu \frac{N_s}{N_p} = 20 \times 3 = 60$

(*ii*) $r_p' = R_p + r_p = 10{,}000 + 1{,}000 = 11{,}000\Omega$.

Let f_1 be the lower half power frequency, then

$$2\pi f_1 L_p = r_p'$$

$$f_1 = \frac{r_p'}{2\pi L_p} = \frac{11000}{2\pi \times 12} = \mathbf{175 \text{ c/s}}$$

(*iii*) Resonant frequency f_2 in high frequency range is given by

$$f_2 = \frac{1}{2\pi\sqrt{(L_s C_s)}} = \frac{1}{2\pi\sqrt{(\cdot 1 \times 400 \times 10^{-12})}}$$

$$= \mathbf{7{\cdot}9 \times 10^4 \text{ cycles/sec.}}$$

(*iv*) The gain A at a frequency of 20 cycles per sec. *i.e.* A_{20} is given by

$$A_{20} = \frac{A_m}{\sqrt{\left\{1 + \left(\frac{r'_p}{\omega L_p}\right)^2\right\}}}$$

$$= \frac{60}{\sqrt{\left\{1+\left(\frac{11000}{256{\cdot}64}\right)^2\right\}}} = \frac{60}{\sqrt{(1+75{\cdot}69)}} = \mathbf{3{\cdot}4.}$$

13. *A transformer coupled amplifier has the following values of circuit constants :*

Valve 6J5 constants; $r_p = 8{\cdot}8k\Omega$, $\mu = 15$, $C_{gp} = 3pf$, $C_{gk} = 3pf$ *and* $C_{pk} = 4pf$. *Transformer constants* $R_{d.c} = 200$ *ohms* $R_s' = 2700$ *ohms,* $L_p = 45$ *henry,* $C_m = 300$ *pf,* $C_p = 10$ *pf,* $C_s = 50$ *pf,* $k = 99{\cdot}8\%$, $n = 3$. *Both polarity dots are at the top. The following stage is a 6J5 with a resistance load and a gain of* $15 \angle 180°$. *The stray and wiring capacitance on the secondary side of the transformer is 50 pf. Compute all values necessary to draw a graph between gain and frequency.*

Assume $L_p' = (1-k)\ L_p$ *and* $L_s' = n^2\ L_p'$ *for an iron core transformer.*

Gain at mid frequency range $= \mu n = 15 \times 3 = \mathbf{45.}$

Lower cut off frequency $f_1 = \dfrac{r_p + R_{d-c}}{2\pi L_p}$

$$= \frac{(8{\cdot}8 + {\cdot}2)\ 10^3}{2\pi \times 45} = \mathbf{31{\cdot}5\ cycles.}$$

Total shunting capacitv

$$= n^2\left[C_s + C_w + C_{in} + \left(\frac{1-n}{n}\right)^2 C_m\right]$$

$$= 3^2\left[50 + 50 + \{3 + (1+15)\ 3\} + \frac{2^2}{3^2} \times 300\right] \times 10^{-12}$$

$$= \mathbf{2557\ pf.}$$

Total leakage inductance $= L'_p + \dfrac{L'_s}{n^2}$

$$= (1-k)\ L_p + n^2\ (1-k)\ L_s/n^2$$

$$= (1 - {\cdot}998)\ 45 + 3^2\ (1 - {\cdot}998)\ 45/3^2$$

$$= \mathbf{0{\cdot}180\ henry.}$$

Resonant frequency $= \dfrac{1}{2\pi\sqrt{(C_s L_s)}}$

$$= \frac{1}{2\pi\sqrt{(2557 \times {\cdot}180 \times 10^{-12})}}$$

$$= \mathbf{7400\ cycles.}$$

$R'_s = 2700$ ohms

$\therefore \quad R_s = \dfrac{R'_s}{n^2} = \dfrac{2700}{9} = \mathbf{300\ ohms.}$

$$Q_a = \frac{2\pi f_0 L_s}{R_p + R_s} = \frac{2\pi \times 7{\cdot}400 \times {\cdot}18}{8{\cdot}8 + {\cdot}2 + {\cdot}3} = \mathbf{0{\cdot}90.}$$

EXERCISES AND PROBLEMS

1. Derive an expression for the voltage amplification of *RC* coupled triode amplifier for :
   (*a*) low frequency range,
   (*b*) intermediate frequency range, and
   (*c*) high frequency range.
2. Discuss frequency response curve of *RC* coupled amplifier.
3. What is the effect of cathode resistance and condenser on the frequency response curve of *RC* coupled amplifier.
4. Draw and explain the gain and phase shift curves for *RC* coupled amplifier in low and high frequency ranges.
5. What do you mean by gain bandwidth product of an amplifier ? Obtain an expression of the gain bandwidth in case of *RC* coupled amplifier and show that it is independent of the circuit parameters R_l, R_c and C.
6. Draw the circuit diagram of low frequency compensation of *RC* amplifier. If the filter ciruit consists of a resistance R_f and a condenser C_f, and the time constant of the coupling circuit is equal to the time constant $C_f R_l$, then prove that the gain at low frequencies is independent of frequency.
7. Obtain an expression for the high frequency compensation of *RC* amplifier. Plot a graph between relative gain $| A_h | / | A_m |$ as a function of (f/f_2) for different values of Q. The symbols have their usual meanings.
8. (*a*) What are the advantages of impedance coupled amplifier over *RC* coupled amplifier ?
   (*b*) Derive the expression for the voltage amplification of an impedance coupled amplifier for intermediate frequency range.
9. Discuss the frequency response curve of an impedance coupled amplifier.
10. (*a*) Prove that in case of transformer instead of considering a capacity C in primary, we can consider a capacity $C[(1 \pm n)/n]^2$ in the secondary where n is the turns ratio.
    (*b*) Draw a circuit diagram and its equivalent circuit in case of a transformer coupled amplifier showing all the capacitances existing between different components.
11. Derive the simple expression for the approximate voltage gain at the mid frequency range of a transformer coupled amplifier ?.
12. A triode is used in an *RC* coupled amplifier. The load resistance is 25 kilo-ohm, coupling condenser is $\cdot 02 \mu F$, grid leak resistance of the next stage is 10^6 ohms, total shunt capacitance is $140 \mu\mu F$. The amplification factor of the tube in the region of operation is 20 and dynamical plate resistance is 15,000 ohms.
    Calculate the voltage gain of one stage of this amplifier in the middle frequency range and also the lower and upper cross-over frequencies. What should be the value of coupling condenser to have a lower half power frequency of 10 c/s ?
13. A resistance-capacitance coupled triode A_1 amplifier stage has plate load 100 K ohm and coupling capacitance of $0{\cdot}05 \mu F$. The amplification factor of the valve is 100 and dynamic plate resistance of 50 K onm. Grid leak

resistance of next stage is 150 K ohm. Calculate (*a*) mid-band gain, (*b*) lower half power frequency and (*c*) magnitude and phase angle of gain at a frequency of $20/2\pi$ c/s.

14. A *RC* coupled amplifier uses pentode having amplification factor $\mu=100$, dynamic plate resistance 500 K ohm, load resistance 75 *K* ohm, grid leak resistance of each stage 400 K ohm, coupling condenser $0{\cdot}2\mu F$, plate to cathode capacitance $5\mu\mu F$, input capacitance $10\mu\mu F$, and stray capacitance $2\mu\mu F$. Calculate lower and upper half power frequencies, mid-band gain and gain bandwith product.

15. A transformer coupled cascade amplifier uses triodes having amplification factor $\mu=20$ and plate resistance $r_p=10{,}000$ ohm. The parameters of transformer are ; effective primary resistance=1000 ohm, effective secondary resistance=10,000 ohm ; incremental magnetising inductance of primary=20 henry ; leakage inductance referred to the primary=0·5. henry ; effective total capacitance referred to the primary=$80\mu\mu F$; turns ratio=5. Calculate (*a*) voltage gain in the middle frequency range, (*b*) lower half power frequency, (*c*) series resonant frequency, and (*d*) voltage gain at a frequency of 50 c/s.

11 FEED BACK AMPLIFIERS

When a part of output signal is combined with the input signal, feed back is said to exist. When the voltage returned to the input circuit is in phase with the signal voltage, it will increase the strength of the signal input and consequently, will increase the gain ; such a feed back is referred to as positive feed back or regeneration. When the voltage returned to the input circuit is 180 degrees out of phase with the signal voltage, it will decrease the strength of the signal voltage and thereby the gain; such a feed back is known as negative feed back, or inverse feed back, or degeneration.

Characteristics of positive feed back :

(*i*) Frequency response curve becomes sharpened or the range of uniform amplification decreases.
(*ii*) Gain of the amplifier increases.
(*iii*) If the feed back is enough, oscillations may take place.

Characteristics of negative feed back :

(*i*) Tends to flatten the frequency response curve and extends the range of uniform amplification.
(*ii*) Practically reduces the frequency and phase shift distortion.
(*iii*) Increases the stability of an amplifier.
(*iv*) Reduces the gain.
(*v*) Higher fidelity.
(*vi*) Less amplitude distortion.
(*vii*) Lower ratio of noise level.
(*viii*) Less harmonic distortion.

11·0. GAIN WITH FEEDBACK :

The principle of feed back is illustrated in fig. 1.

In the diagram, a voltage E_g is applied to the input terminals of the amplifier. Suppose that the voltage drop at the output terminals is V_0. Let a fraction of it, αV_0, be fed back in series with input signal in such a way that the resultant signal that appears between grid-cathode terminals has the form

$$E_g' = E_g + \alpha V_0. \qquad \ldots(1)$$

By definition, normal gain A of an amplifier is

$$A = \frac{\text{Output voltage}}{\text{Input potential between grid and cathode}} = \frac{V_0}{E_g'}$$

or $$AE_g' = V_0.$$

Putting the value of E_g' from (1),

$$A(E_g + \alpha V_0) = V_0$$

or $$AE_g = V_0(1 - A\alpha). \quad \ldots(2)$$

But the resultant gain of the amplifier including the effect of feed back is defined as

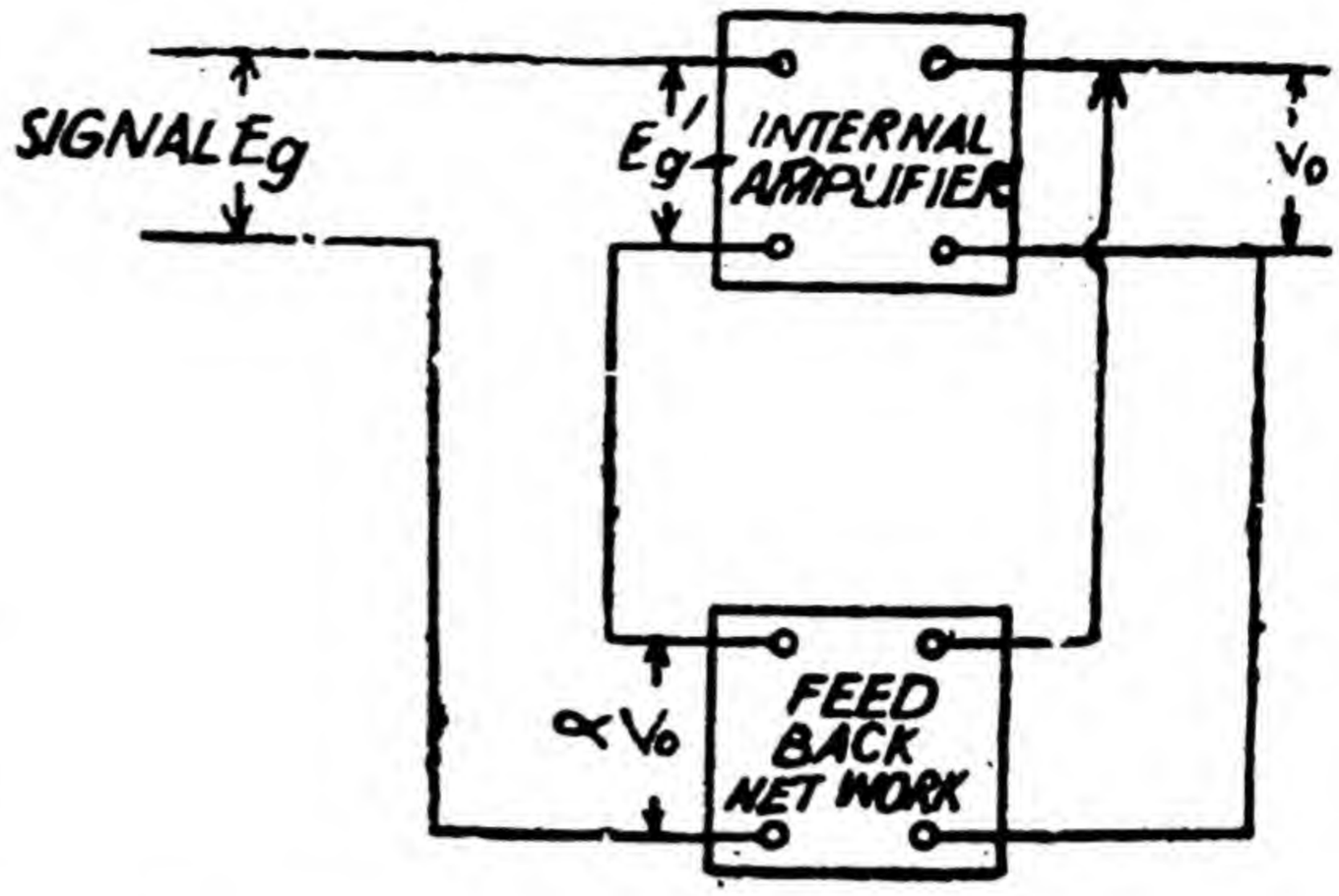


Fig. 1. Negative feed back principle

$$A_f = \frac{\text{output voltage}}{\text{input signal voltage}} = \frac{V_0}{E_g} = \frac{A}{1 - A\alpha}. \quad \ldots(3)$$

This equation expresses the resultant gain of the amplifier with feed back in terms of the nominal gain A of the amplifier without feed back.

We draw following significant inferences from equation (3) :

(1) If the numerical value $| 1 - A\alpha | > 1$, then $A_f < A$ and the feed back is known as degenerative or negative.

(2) If the numerical value $| 1 - A\alpha | < 1$, then $A_f > A$ and feed back is positive or regenerative.

(3) If $A\alpha = +1$, $1 - A\alpha = 0$ and A_f becomes infinite. This shows that the amplifier gives an output with zero input. The amplifier then works as an oscillator *i.e.* a device for generating alternating voltage.

(4) If $| A\alpha | \gg 1$, then $1 - A\alpha \approx -A\alpha$.

$$A \approx \frac{A}{-A\alpha} = -\frac{1}{\alpha}.$$

We now infer that :

(*a*) If $A\alpha \gg 1$, then the gain is inversely proportional to α and the frequency response characteristic of the feed back amplifier is inverse of the response characteristic of the feed back network. Thus with a proper design of the feed back network, the amplifier can be made to have any desired frequency response.

(*b*) If $| A\alpha | \gg 1$, then amplification is independent of A and hence the tube factor. Hence the gain stability of the amplifier is improved by feed back provided the feed back factor is large.

(*c*) If $| A\alpha | \gg 1$, and load does not form a part of the feed back network, then the amplification is independent of load impedance in output. If α is made dependent upon load impedance, then the amplification can be made to vary in a desired manner with load impedance.

Thus if feed back is positive, *i.e.*, α is positive, then

$$A_f = \frac{A}{1 - A\alpha} \quad \ldots(4)$$

If feed back is negative, *i.e.* α is negative, then

$$A_f = \frac{A}{1 + A\alpha} \qquad \ldots(5)$$

The presence of the negave feed back in an amplifier results in a number of desirable characteristics as discussed below :

(1) **Stability of Amplification** : Suppose that the feed back is negative and the feed back factor $A\alpha$ is made large in comparison to unity ; then resultant gain will take the form

$$A_f = \frac{A}{A\alpha} = \frac{1}{\alpha} \qquad \ldots(6)$$

From equation (6), we infer that $A\alpha$ is large compared to unity : the actual amplification with negative feed back is a function of the characteristic of the feed back network only. In particular, if α is independent of frequency, A_f will also be independent of frequency. This will substantially reduce the frequency and phase distortion of amplifier.

If $A\alpha \gg 1$, then $\frac{1}{\alpha} \ll A$

and $$A_f \approx \frac{1}{\alpha} \ll A. \qquad \ldots(7)$$

Thus the overall gain of the amplifier with negative feed back is less than the normal gain without feed back. This is the price that must be paid to secure the advantages of inverse feed back. This is not a serious price to pay since the loss in gain can be overcome by the use of addjtional tubes.

Even if equation (6) is not completely valid, there is considerable improvement in general stability. Taking log of eq. (5),

$$\log A_f = \log A - \log (1 + A\alpha).$$

Differentiating, we get

$$\frac{dA_f}{A_f} = \frac{dA}{A} - \frac{\alpha\, dA}{(1 + A\alpha)} = \frac{dA}{A}\left(1 - \frac{\alpha A}{1 + A\alpha}\right) = \frac{dA}{A}\left(\frac{1}{1 + A\alpha}\right) \qquad \ldots(8)$$

$\frac{dA_f}{A_f}$ represents the fractional change in the overall gain and $\frac{dA}{A}$ gives the fractional change in A. If, for example, in a particular amplifier $(1 + A\alpha) = 5$, the variation in any parameter that might cause a variation of 5% change in the nominal gain will result in a change of only 1% in the resultant gain of the amplifier.

(2) **Reduction of frequency and phase distortion :** It follows from equation (4) that the overall gain of the amplifier is almost independent of frequency, provided that α is frequency independent. In such cases the frequency and phase distortion of an amplifier are materially reduced below the non-feed back value.

(3) Reduction of non-linear distortion : We have so far assumed that waveshapes of output and input potentials are the same. If an appreciable non-linear distortion exists, then the output contains harmonic components in addition to the signal of fundamental frequency.

Suppose B_2 is the first harmonic generated in the tube when large voltage is impressed on the input and B_2' is the first harmonic component that appears at the output and differs from B_2; because of feed back, we have

$$B_2' = B_2 - B_2' A\alpha$$

or

$$B_2' (1 + A\alpha) = B_2,$$

$$B_2' = \frac{B_2}{(1 + A\alpha)}. \qquad \ldots(9)$$

If $(1 + A\alpha) = 10$, then first harmonic distortion with feed back is only one-tenth of its value without feed back.

11·1. INCREASE IN BANDWIDTH OF AN AMPLIFIER DUE TO NEGATIVE FEED BACK :

(1) High frequency performance :

In case of negative feed back, the gain of the amplifier is given by

$$A_f = \frac{A}{1 + \alpha A}. \qquad \ldots(1)$$

At high frequencies, the gain of *RC* coupled amplifier without feed back is expressed as

$$A_h = \frac{A_m}{1 + jf / f_2}, \qquad \ldots(2)$$

where f_2 is the upper cut off frequency. A_h and A_m are the gains of the amplifier at high and middle frequencies respectively. Let the gain at high frequencies with feedback be denoted by $(A_h)_f$, then from equation (1), we have

$$(A_h)_f = \frac{A_h}{1 + A_h \alpha}. \qquad \ldots(3)$$

Substituting the value of A_h in equation (3) from equation (2),

$$(A_h)_f = \frac{\dfrac{A_m}{1 + jf / f_2'}}{1 + \alpha \dfrac{A_m}{1 + jf / f_2}}$$

$$= \frac{A_m}{1 + \alpha A_m + jf / f_2} = \frac{\dfrac{A_m}{1 + \alpha A_m}}{1 + \dfrac{jf}{f_2 (1 + \alpha A_m)}} \qquad \ldots(4)$$

In case of feed-back amplifier, let f_2' represent the upper cut off frequency then

$$(A_h)_f=\frac{\frac{A_m}{1+\alpha A_m}}{1+\frac{jf}{f_2'}} \qquad \ldots(5)$$

Comparing equations (4) and (5), we get

$$f_2'=f_2\,(1+\alpha\, A_m), \qquad \ldots(6)$$

It is clear from equation (6) that the upper cut off frequency is increased by a factor $(1+\alpha\, A_m)$ which is, of course, greater than 1.

(2) Low frequency performance :

At low frequencies, the gain of RC coupled amplifier without feed back is given by

$$A_l=\frac{A_m}{1-jf_1/f}, \qquad \ldots(7)$$

where f_1 is the lower cut off frequency.

Now $$(A_l)_f=\frac{A_l}{1+\alpha A_l}. \qquad \ldots(8)$$

Substituting the value of A_l in equation (8) from equation (7), we have

$$(A_l)_f=\frac{\frac{A_m}{1-jf_1/f}}{1+\alpha\,\frac{A_m}{1-jf_1/f}}=\frac{A_m}{1+\alpha\, A_m-jf_1/f}$$

$$=\frac{\frac{A_m}{1+\alpha A_m}}{1-\frac{jf_1}{f(1+\alpha A_m)}}. \qquad \ldots(9)$$

In case of feed back amplifier let f_1' denote the lower cut off frequency then

$$(A_l)_f=\frac{\frac{A_m}{1+\alpha A_m}}{1-j\,f_1'/f} \qquad \ldots(10)$$

Comparing equations (9) and (10), we get

$$f_1'=\frac{f_1}{(1+\alpha A_m)} \qquad \ldots(11)$$

This expression shows that the lower cut off frequency of the feed back amplifier is reduced by $(1+\alpha A_m)$.

Conclusions drawn from eqs. (6) and (11) predict increase in bandwidth as shown in fig. 2.

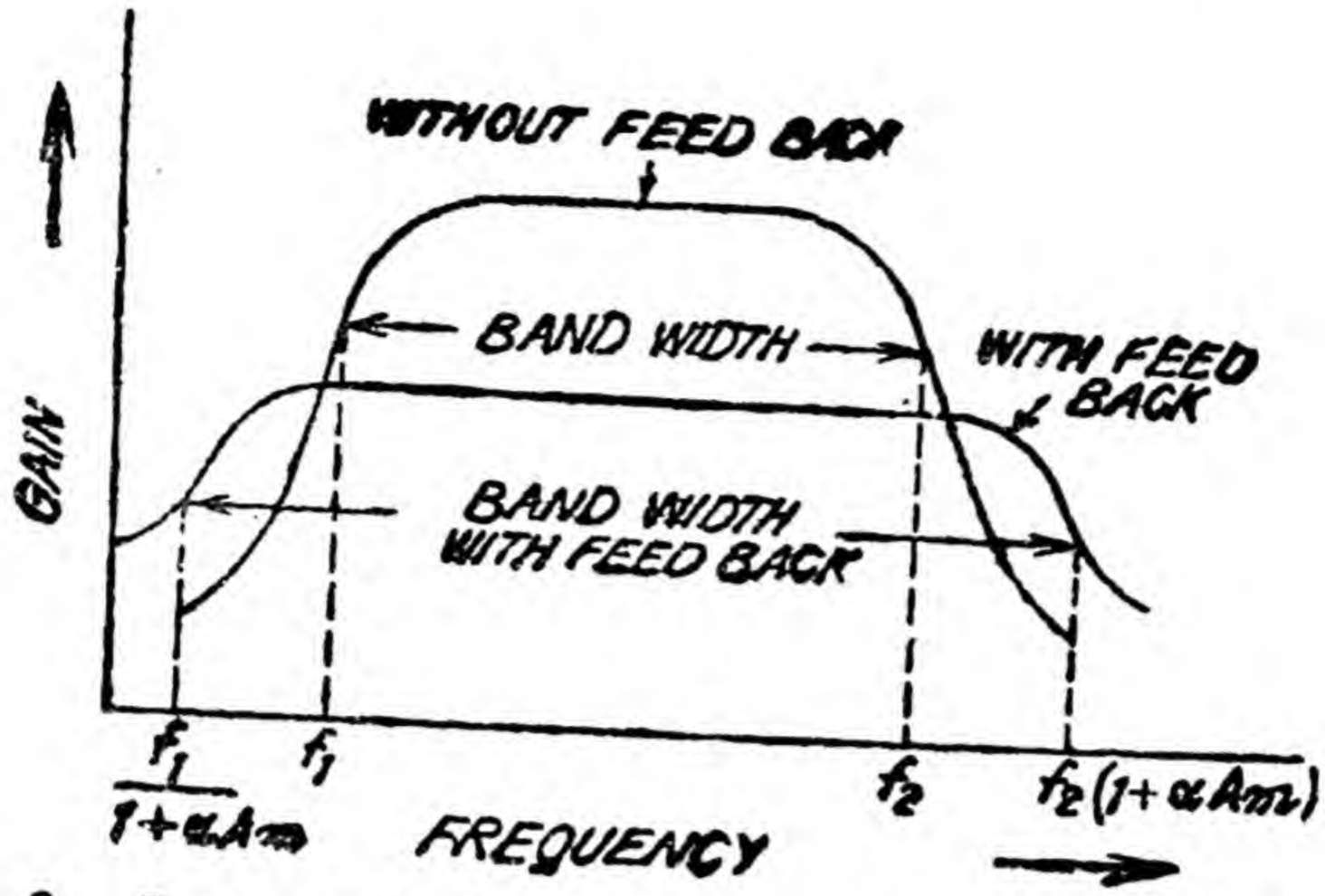


Fig. 2. Increase in bandwidth with negative feed back.

11·2. FEED BACK CIRCUITS :

(i) **Current feed back circuit :** It is the system in which the voltage fed back from the output of the tube is proportional to the current through the load. This type of feed back is known as 'current feed back'.

(ii) **Voltage feed back circuit :** It is th e system in which the voltage fed back from the output of the amplifier is proportional to the voltage across the load. This type of feed back is called 'voltage feed back'.

(iii) **Compound feed back circuit :** This utilises a combination of current and voltage feed back.

(i) **Current feed back :** In most of the amplifi e rs self biasing is used. The self biasing is achieved by connecting a parallel combination of a resistance and capacitance in cathode lead. The value of the bypass condenser is so chosen that even for the lowest frequency of the input signal, the reactance of the condenser is very small. Under this condition the total impedance of the cathode may be neglected for a.c. operation. In this case no feed back is achieved. In order to achieve the inverse current feed back, the condenser is omitted from the cathode lead. Now only the resistor being in the cathode lead. The feed back ratio is the ratio of this

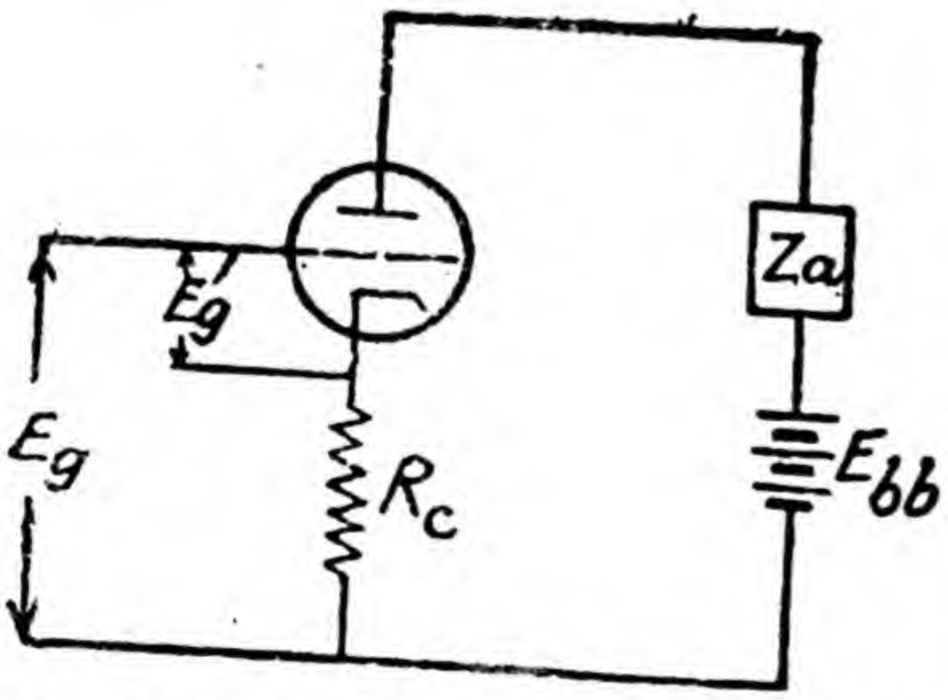


Fig. 3. (a) Cathode feed back amplifier.

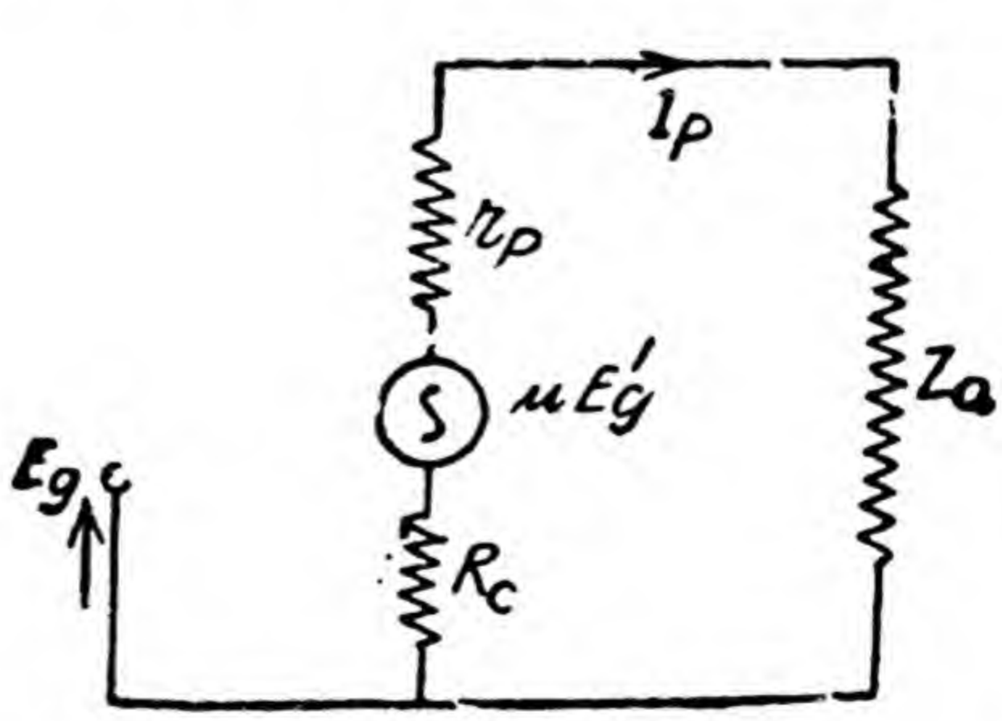


Fig. 3(b) Equivalent circuit.

resistor to the load in the plate circuit. Figure 3 (*a*) shows the circuit diagram of cathode feedback amplifier. The a.c. plate current flowing through the cathode impedance R_c produces an a.c. voltage drop which is superimposed on the input voltage $\mathbf{E}_g$ to constitute the voltage $\mathbf{E}_g'$. Polarity of this voltage is such as to reduce the signal voltage $\mathbf{E}_g$ and hence overall gain is reduced. The circuit, therefore, provides current inverse feed back. The feed back is affected through the resistor R_c, the feed factor being $R_c/\mathbf{Z}_a$.

The equivalent circuit may be analysed to get an expression for voltage gain as follows :

Voltage feed back $= R_c.\mathbf{I}_p$.

Effective signal voltage $\mathbf{E}_g' = \mathbf{E}_g - R_c\mathbf{I}_p$.

From the equivalent circuit,

$$\mu\mathbf{E}_g' = \mathbf{I}_p (r_p + R_c + \mathbf{Z}_a)$$

or
$$\mu (\mathbf{E}_g - R_c\mathbf{I}_p) = \mathbf{I}_p (r_p + R_c + \mathbf{Z}_a)$$

$$\mathbf{I}_p = \frac{\mu\mathbf{E}_g}{(r_p + R_c + \mathbf{Z}_a) + R_c.\mu}$$

$$= \frac{\mu\mathbf{E}_g}{r_p + R_c(1+\mu) + \mathbf{Z}_a} = \frac{\mu\mathbf{E}_g}{r_p' + \mathbf{Z}_a}, \qquad \ldots(1)$$

where $r_p' = r_p + R_c (1+\mu)$.

Thus the gain of this amplifier

$$[\mathbf{A}_f] = \frac{\mathbf{V}_0}{\mathbf{E}_f} = \frac{-\mathbf{Z}_a\mathbf{I}_p}{\mathbf{E}_g} = \frac{-\mu\mathbf{Z}_a}{r_p' + \mathbf{Z}_a}. \qquad \ldots(2)$$

Equation (1) gives the current in the plate circuit and this may be interpreted to show that the effect of negative feed back is to increase effective internal resistance of the tube from r_p without feed back to r_p' given by $r_p' = r_p + R_c (1+\mu)$.

From equation (1) it is possible to draw the alternate a.c. equivalent circuit of fig. 3 (*a*) as shown in fig. 4.

The internal impedance without feed back $= r_p + R_c$.

The internal impedance with feed back $= r_p + R_c (1+\mu)$.

The ratio of internal impedance with feed back and without feed back is given by

$$\frac{r_p + R_c (1+\mu)}{r_p + R_c} = 1 + \mu \frac{R_c}{r_p + R_c} \qquad \ldots(3)$$

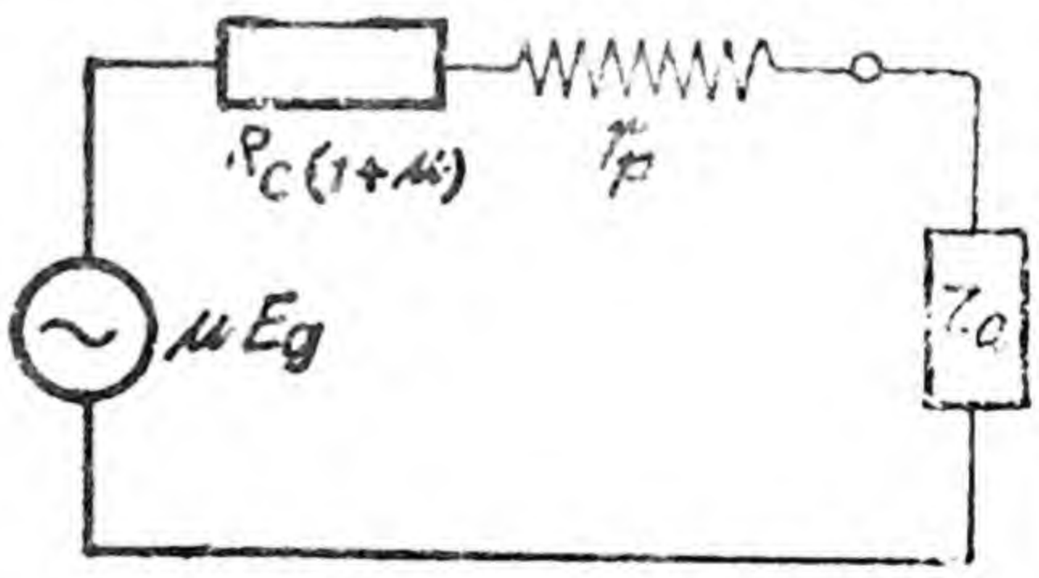


Fig. 4. A. c. equivalent circuit of fig. 3 (*a*)

So, if the effective internal resistance of the tube is high compared with the impedance of the load, the current is sufficiently independent of plate load.

†(ii) **Voltage Fe ed-back :** Figure 5 shows a *RC* coupled amplifier with voltage inverse feed·back. Voltage developed across load resistance Z_a is coupled by coupling condenser C to shunt resistors R_1 and R_2. Voltage developed across R_1 is feed back to input in a way as to be subtracted from the input E_g.

The normal gain of amplifier is given by $[A]=\frac{-\mu Z_a}{r_p+Z_a}$ and α in this circuit is $\frac{R_1}{R_2+R_1}$.

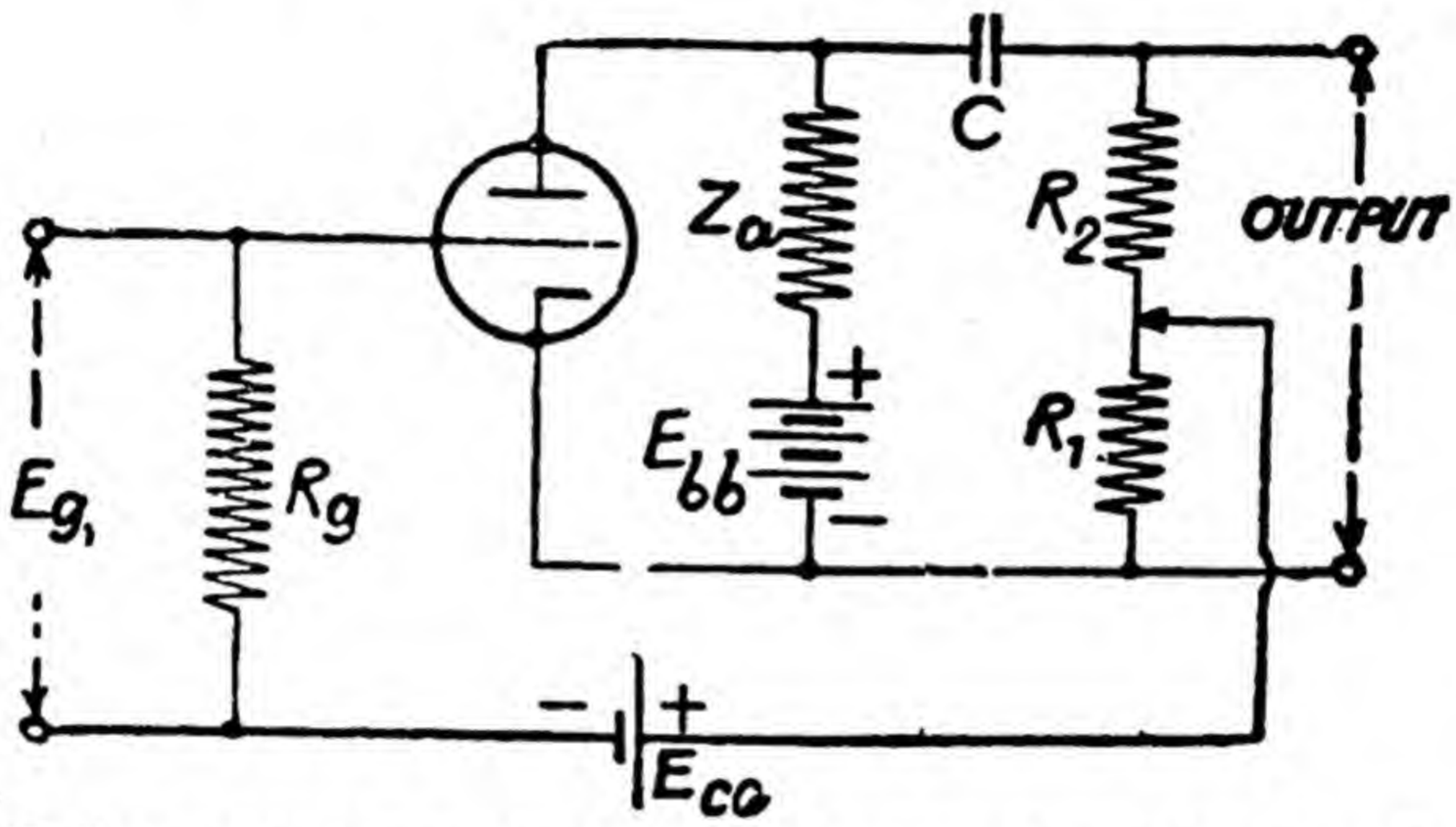


Fig. 5. RC coupled amplifier with inverse feed-baek voltage.

The resultant gain with feed back is

$$|A_f|=\frac{A}{1-A\alpha}=\frac{-\mu Z_a/(r_p+Z_a)}{1+\alpha\left(\frac{\mu Z_a}{r_p+Z_a}\right)}$$

$$=\frac{-\mu Z_a}{(r_p+Z_a+\alpha\mu Z_a}$$

$$=\frac{-\mu Z_a}{r_p+Z_a\,(1+\alpha\mu)}$$

$$=\frac{\frac{-\mu}{(1+\alpha\mu)}Z_a}{\frac{r_p}{(1+\alpha\mu)}+Z_a}=\frac{-\mu' Z_a.}{r_p'+Z_a} \qquad ...(3)$$

where $\mu'=\frac{\mu}{1+\alpha\mu}$ and $r_p'=\frac{r_p}{1+\alpha\mu}$.

Hence the resultant gain is that which would be obtained from the tube whose amplification factor is μ' and the plate resistance is r_p'. The tube behaves as if it has effective amplification factor of

†If there is phase difference of 180° between output and input, the feed back is negative as shown in fig. 5.

In case of double stage *RC* coupled amplifier the phase difference will be 360° and hence the feed back will be positive.

$\mu' = \frac{\mu}{1+\alpha\mu}$ and dynamical plate resistance $r_p' = \frac{r_p}{1+\alpha\mu}$. Thus both the effective amplification factor and dynamic plate resistance have been reduced by the factor $\frac{1}{(1+\alpha\mu)}$ as shown in fig 6. This indicates that a tube possessing high plate resistance, can be effectively converted into a low plate resistance tube and thereby permits an impedance match to a low impedance load.

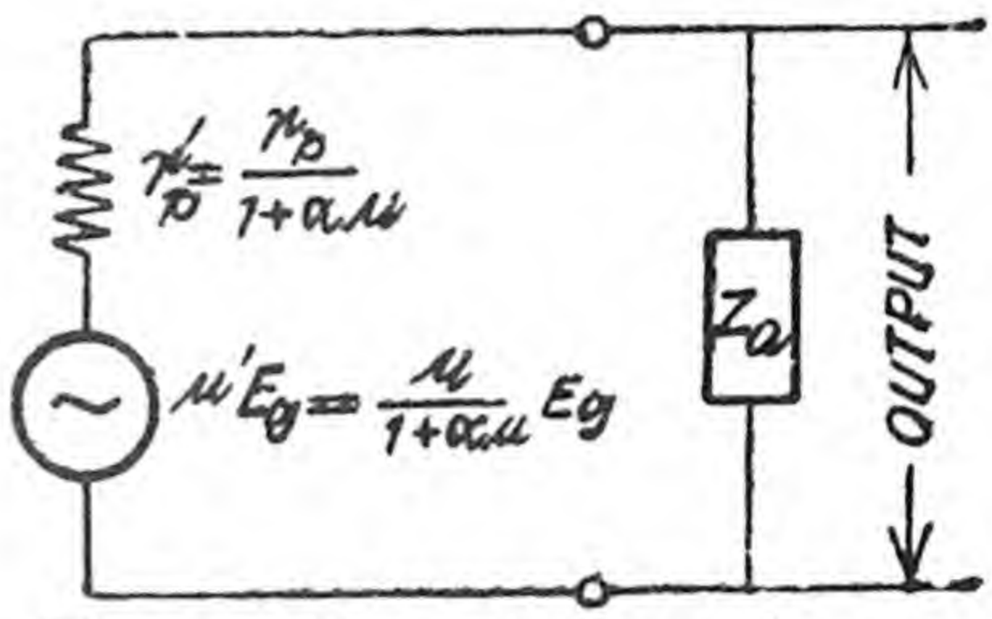


Fig. 6. Alternate form of a.c. equivalent of fig. 5.

(iii) Compound Feed Back :

The circuit showing the compount feed back is given in figure 7.

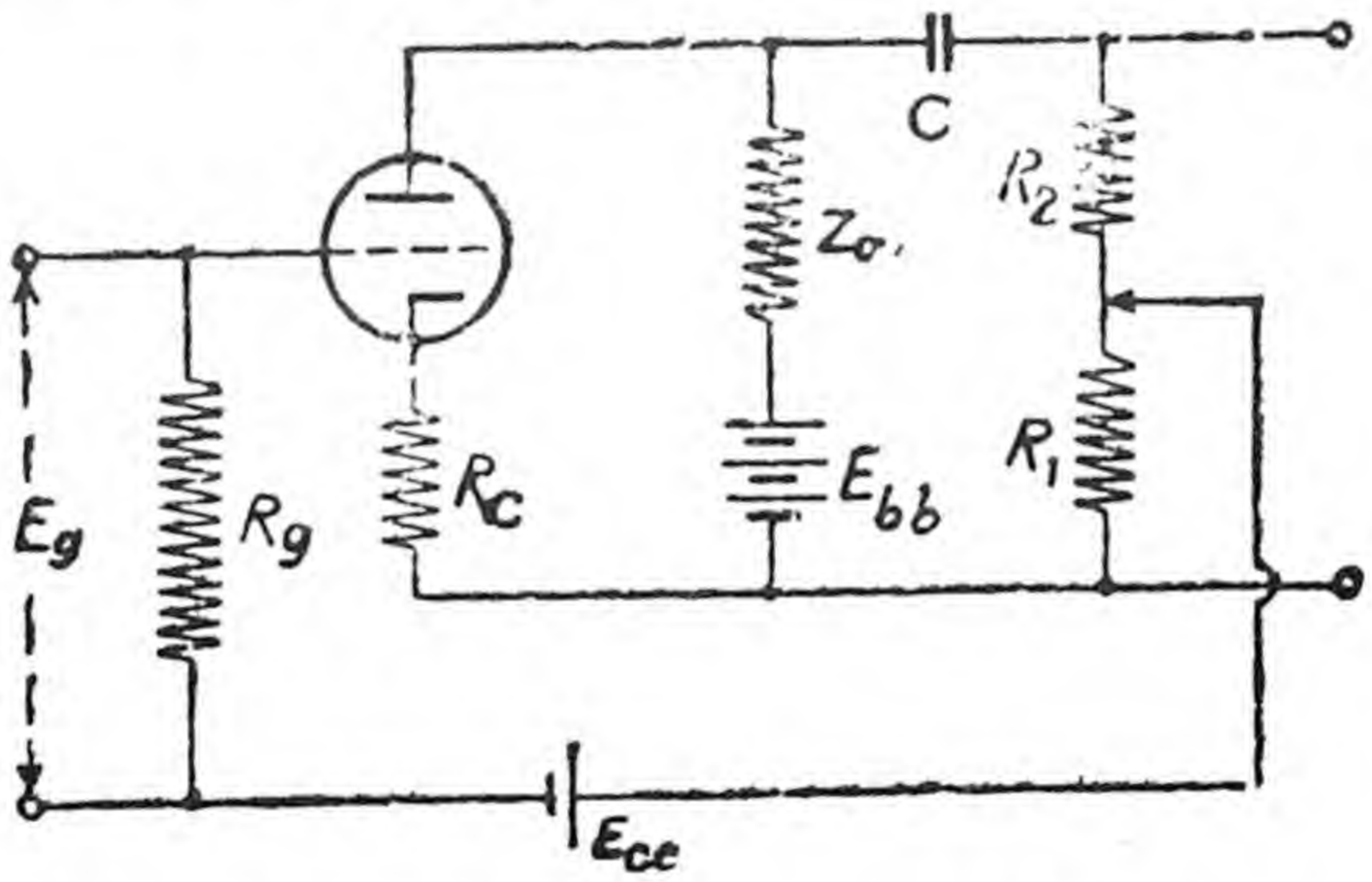


Fig. 7 Circuit showing compound feed back.

A.c. plate current flowing through cathode resistance R_c produces a voltage drop which gets added to the input voltage E_g in such a polarity as to reduce it. This, therefore, provides current inverse feed back. The feed back factor in this case is R_c/Z_a. Fraction $\frac{R_1}{R_2+R_1}$ of a.c. output voltage is fed-back to the input to provide voltage feed back.

Thus the compound feed back factor will be

$$\alpha' = \frac{R_c}{Z_a} + \frac{R_1}{R_2+R_1}.$$

Here we assume that $(R_2+R_1) \gg Z_a$ and the reactance of capacitor is negligible over the frequency range of operation.

The voltage gain with current feed-back is given by

$$A_c = \frac{-\mu Z_a}{r_p + Z_a + R_c(1+\mu)}.$$

Now the voltage amplfication with current feed-back as well as voltage feed-back is given by

$$A_f = \frac{A_c}{1 - A_c\alpha} = \frac{-\mu Z_a/r_p + Z_a + R_c\,(1+\mu)}{1 - [-\mu Z_a/r_p + Z_a + R_c(1+\mu)]\alpha}$$

$$= \frac{-\mu Z_a}{r_p + Z_a + R_c\,(1+\mu) + \mu Z_a\alpha},$$

$$\therefore \quad [A_f] = \frac{-\mu Z_a}{r_p + R_c\,(1+\mu) + Z_a\,(1+\alpha\mu)},$$

which may be written in the form

$$[A_f] = \frac{-\mu'' Z_a}{r_p'' + Z_a}, \qquad \ldots(4)$$

where $\mu'' = \frac{\mu}{(1+\alpha\mu)}$ and $r''_p = \frac{r_p + R_c\,(1+\mu)}{(1+\alpha\mu)}$

Equation (4) suggests that the circuit of figure 7 may be represented bv a.c. equivalent circuit of figure 8.

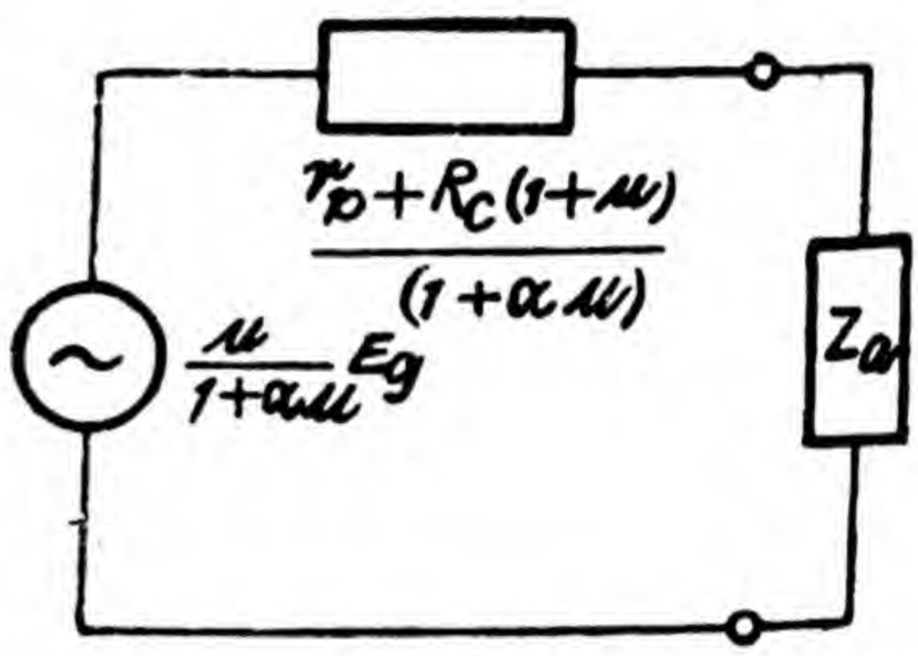


Fig. 8. A.c. equivalent circuit of figure 7.

We know that current inverse feed back increases the internal impedance whereas voltage feed back reduces it. Thus the internal impedance in compound feed back is a function of relative effects in two types of feed back, *i.e.*, it may be increased or decreased.

11·3. EFFECT OF FEED BACK ON THE INPUT TERMINAL IMPEDANCE :

In this article we shall examine that how the input impedance of the amplifier is affected by the presence of the feed back. A generalized treatment of the problem will be considered. It will be found that the effective input impedance increases for both voltage and current feed back.

(a) Voltage feed back : The block diagram of voltage feed back amplifier is shown in figure 9.

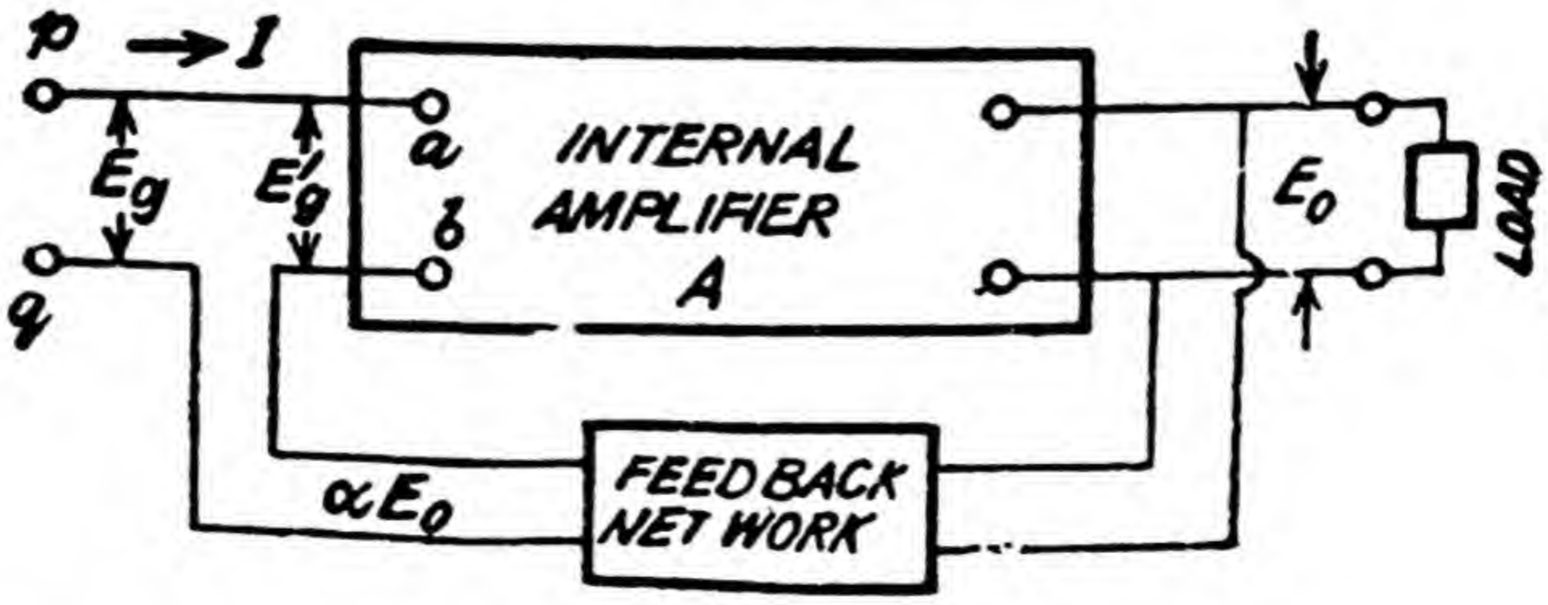


Fig. 9. Block diagram of voltage feed back amplifier.

The input impedance with feed back is given by

$$Z_{1f}=\frac{E_g}{I}. \qquad \ldots(1)$$

From figure

$$E_g=E_g'-\alpha E_0.$$

$$\therefore\ Z_{1f}=\frac{E_g'-\alpha E_0}{I}$$

$$=\frac{E_g'-\alpha A E_g'}{I}=\frac{E_b'}{I}(1-\alpha A). \qquad \ldots(2)$$

The input impedance without feed back is

$$Z_1=\frac{E_g'}{I}.$$

$$\therefore\ Z_{1f}=Z_1(1-\alpha A). \qquad \ldots(3)$$

For negative feed back $|(1-\alpha A)|>1$, therefore, the input impedance with feed back is greater than the input impedance without feed back.

(b) Current feed back : The block diagram of current feed back is shown in fig. 10.

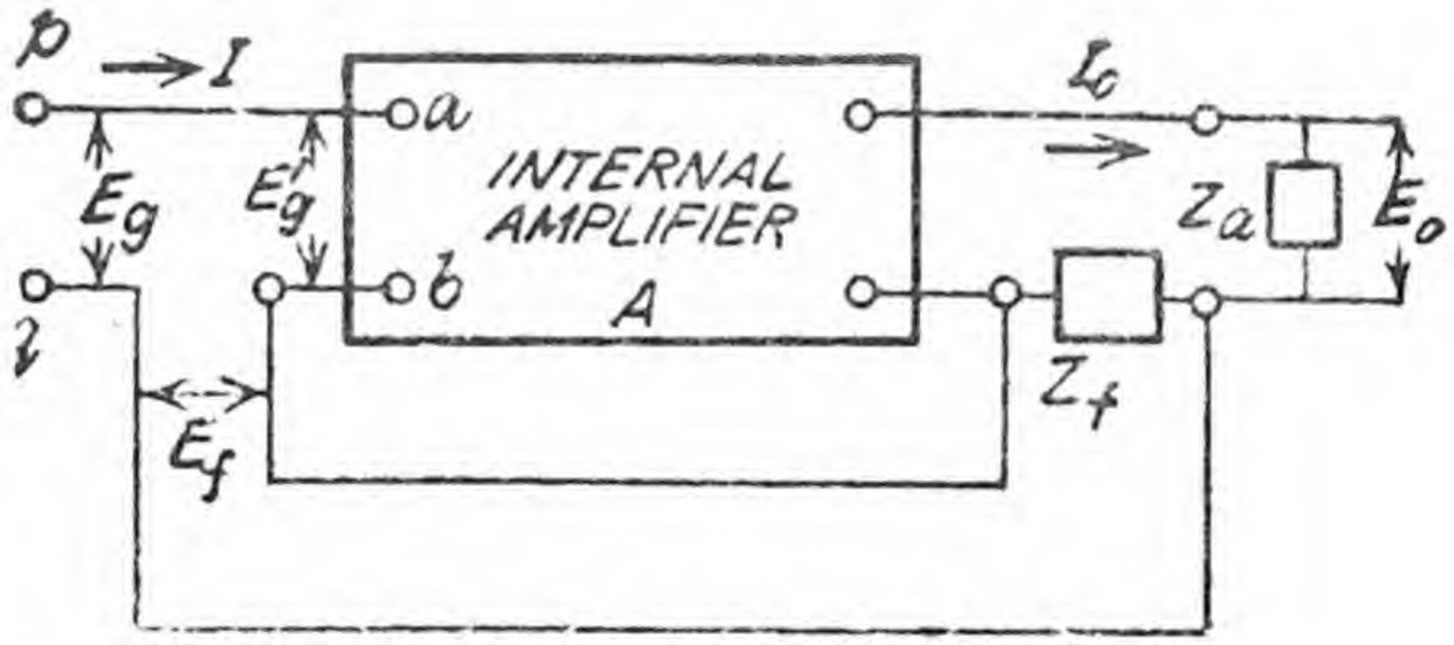


Fig. 10. Block diagram of current feed back.

In this case

$$Z_{1f}=\frac{E_g}{I}$$

$$=\frac{E_g'-E_f}{I}$$

$$=\frac{E_g'-I_0Z_f}{I}$$

$$=\frac{1}{I}\left(E_g'-\frac{E_0}{Z_a}\times Z_f\right). \qquad \ldots(4)$$

But $\alpha=\dfrac{Z_f}{Z_a}$ and $E_0=E_g'A.$

$$\therefore\quad Z_{1f}=\frac{1}{I}(E_g'-E_g'\,A\alpha)=\frac{E_g'}{I}(1-A\alpha)$$

or $$Z_{1f}=Z_1(1-A\alpha). \qquad \ldots(5)$$

Equation (5) shows that input impedance with current feed back is greater than the input impedance without feed back.

11·4. EFFECT OF FEED BACK ON THE OUTPUT TERMINAL IMPEDANCE :

The output terminal impedance of a circuit is the impedance looking back from the output terminals into the energy source sueh that load impedance is connected in the circuit but the input potential is reduced to zero. Clearly, the output impedance of the amplifier is nothing but a parallel combination of the effective internal impedance Z_{tf} and the load impedancc Z_a. We know that Z_{tf} depends upon the type of feed back, hence the output impedance will also depend upon the type of feed back.

Figure 11 shows the output terminal impedance of the general feed back amplifier. The output impedance Z_{of} is the ratio of the current I_0 into the output terminals when a potential E_0 is impressed. Hence

$$Z_{0f}=\frac{Z_{tf}.Z_a}{Z_{tf}+Z_a}. \quad ...(1)$$

The output impedance without feed back Z_0 is given by

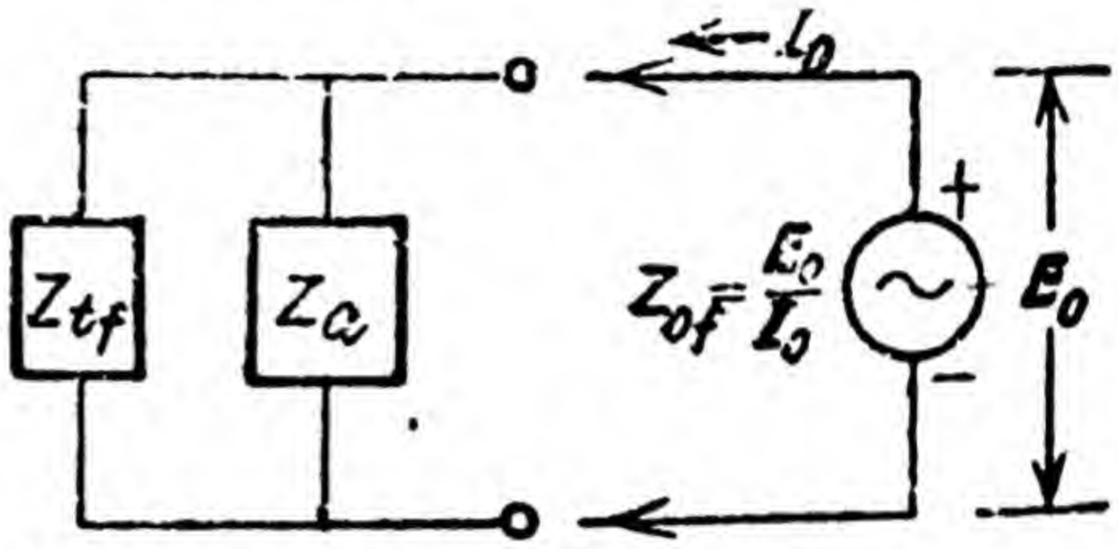


Fig. 11. The output terminal impedance of the general feed back amplifier.

$$Z_0=\frac{Z_t\ Z_a}{Z_t+Z_a} \quad ...(2)$$

The value of Z_{tf} in case of voltage feed back is given by

$$Z_{tf}=\frac{Z_t}{1-\alpha A_t}. \quad ...(3)$$

In case of current feed back

$$Z_{tf}=Z_t-A_tK_a. \quad .. (4)$$

(a) Voltage feed back : In order to find out the output terminal impedance with voltage feed back, the input source to the general feed back amplifier is reduced to zero and a voltage source E_0 is applied to the output terminals. The situation is shown in figure 12. The internal amplifier circuit is represented by an impe-

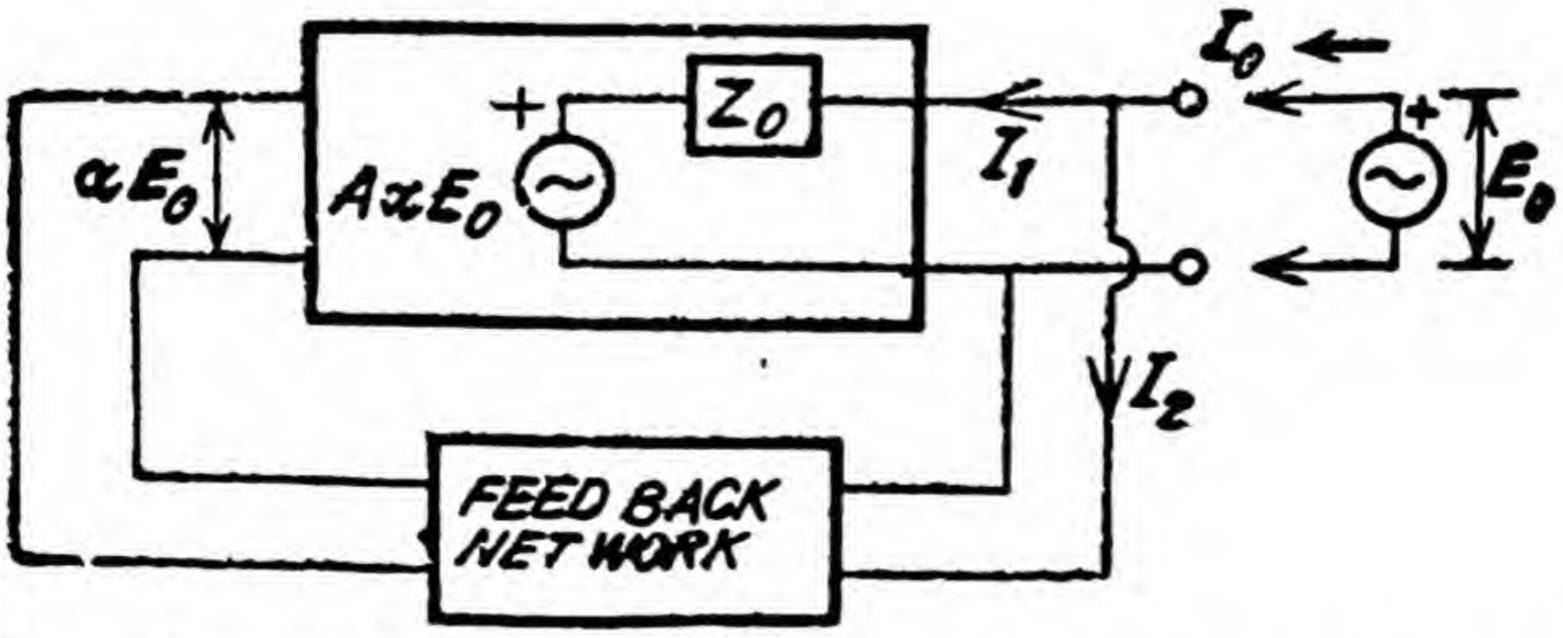


Fig. 12. General voltage feed back amplifier for calculating the output terminal impedance.

dance Z_0. Due to the voltage feed back, an internal voltage source of $A\alpha E_0$ volts appears in series with Z_0.

The current I_0 from the applied source is seen to be

$$I_0 = \frac{E_0 - A\alpha E_0}{Z_0}. \qquad \ldots(5)$$

Now effective output impedance with feed back is

$$Z_{0f} = \frac{E_0}{I_0} = \frac{Z_0}{1 - A\alpha}. \qquad \ldots(6)$$

The equation (6) shows that the output impedance is reduced by a factor $(1 - A\alpha)$.

Alternate treatment : When the feed back loop is closed, let the output current I_0 from the generator has two parts : One part I_1 flows through the amplifier while the other part I_2 enters into the feed back loop. Thus

$$I_0 = I_1 + I_2,$$

where $$I_1 = \frac{E_0 - A\alpha E_0}{Z_0} \text{ and } I_2 = \frac{E_0}{Z_\alpha},$$

where Z_α is the input impedance of the feed back circuit.

$$\therefore\ I_0 = \frac{E_0 - A\alpha E_0}{Z_0} + \frac{E_0}{Z_\alpha}$$

or $$\frac{I_0}{E_0} = \frac{1 - A\alpha}{Z_0} + \frac{1}{Z_\alpha}.$$

In general Z_α is far greater than Z_0, *i.e.* $Z_\alpha > Z_0$.

$$\therefore\ \frac{I_0}{E_0} = \frac{1 - A\alpha}{Z_0}$$

or $$Z_{0f} = \frac{Z_0}{1 - A\alpha}.$$

(b) Current feed back : In this case also the voltage source is reduced to zero as before, and a voltage source is applied at the output terminals. The situation is shown in fig. 13. I_f is the current flowing through the feed back impedance Z_f.

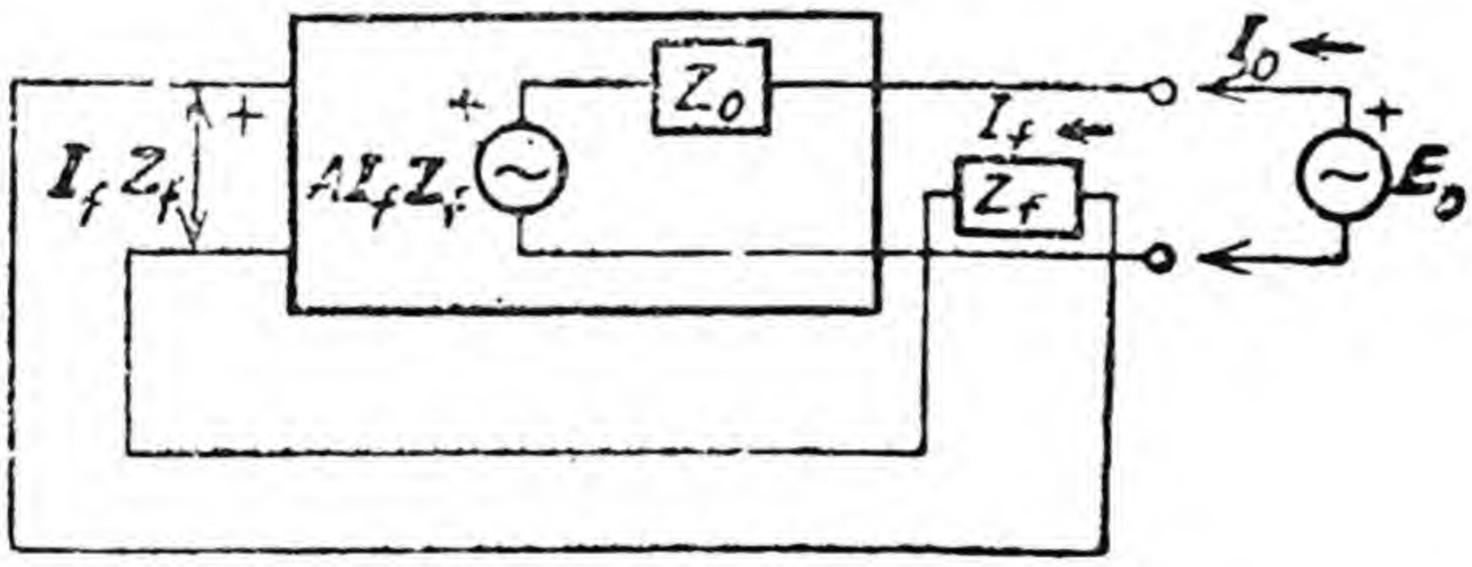


Fig. 13. General current feed back amplifier for calculating the output terminal impedance.

Applying Kirchoff's current law to the output circuit

$$\mathbf{E_0}=\mathbf{I_0Z_0}-\mathbf{AI}_f\mathbf{Z}_f$$

$$=\mathbf{I_0Z_0}-\mathbf{AZ}_f\left(\mathbf{I_0}-\frac{\mathbf{E_0}}{\mathbf{Z}_a}\right).$$

This gives

$$\mathbf{I_0}\,(\mathbf{Z_0}-\mathbf{AZ}_f)=\mathbf{E_0}\left(1-\mathbf{A}\frac{\mathbf{Z}_f}{\mathbf{Z}_a}\right).$$

$$\therefore\ \mathbf{Z}_{0f}=\mathbf{E}_i/\mathbf{I_0}=(\mathbf{Z_0}-\mathbf{AZ}_f)/(1-\mathbf{AZ}_f/\mathbf{Z}_a=\mathbf{Z_0}\,\frac{1-\mathbf{A}\,(\mathbf{Z}_f/\mathbf{Z_0})}{1-\mathbf{A}\,(\mathbf{Z}_f/\mathbf{Z}_a)}.$$

11·5. CATHODE FOLLOWER :

All the amplifiers we have dealt with so far are of the grounded cathode type in which the signal is applied between grid and cathode, the latter being grounded. These amplifiers are most commonly used. There, however, exist other configurations for vacuum tube amplifier, *i.e.* cathode follower.

Fig. 14 shows the circuit arrangement of a cathode follower.

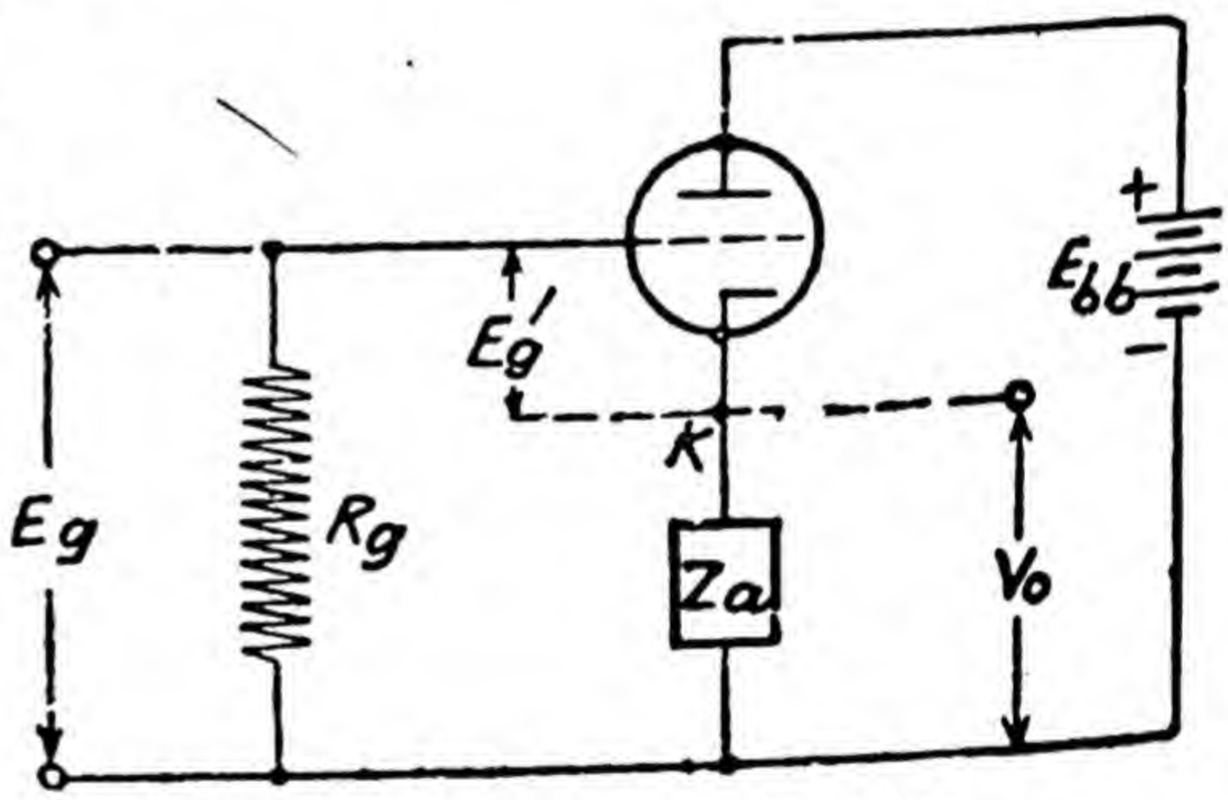


Fig. 14. Circuit arrangement for cathode follower.

In this amplifier an increase in grid to ground voltage $\mathbf{E}_g$, increases the plate current $\mathbf{I}_p$ and thus increases cathode to ground voltage. Thus the cathode follows the grid in its change of potential with respect to ground and hence the name 'cathode follower'.

When an alternating signal voltage $\mathbf{E}_g$ is applied between the grid and ground we see from the figure that whole of the alternating output voltage $\mathbf{V}_0$ is returned to the input circuit. Since the voltage is opposite in phase to E_g, the feed back fraction α is -1.

The equivalent circuit of cathode follower is shown in figure 15.

From a.c. equivalent circuit,

$$\mathbf{I}_p\,(r_p+\mathbf{Z}_a)=\mu\mathbf{E}_g'=\mu\,(\mathbf{E}_g-\mathbf{V}_0)=\mu\,(\mathbf{E}_g-\mathbf{I}_p\mathbf{Z}_a).$$

$$\mathbf{I}_p\,[r_p+\mathbf{Z}_a\,(1+\mu)]=\mu\mathbf{E}_g,$$

$$\mathbf{I}_p=\frac{\mu\mathbf{E}_g}{r_p+\mathbf{Z}_a\,(1+\mu)}.$$

Thus the voltage gain

$$\mathbf{A}_f=\frac{\mathbf{V}_0}{\mathbf{E}_g}=\frac{\mu\mathbf{Z}_a}{r_p+\mathbf{Z}_a\,(1+\mu)}$$

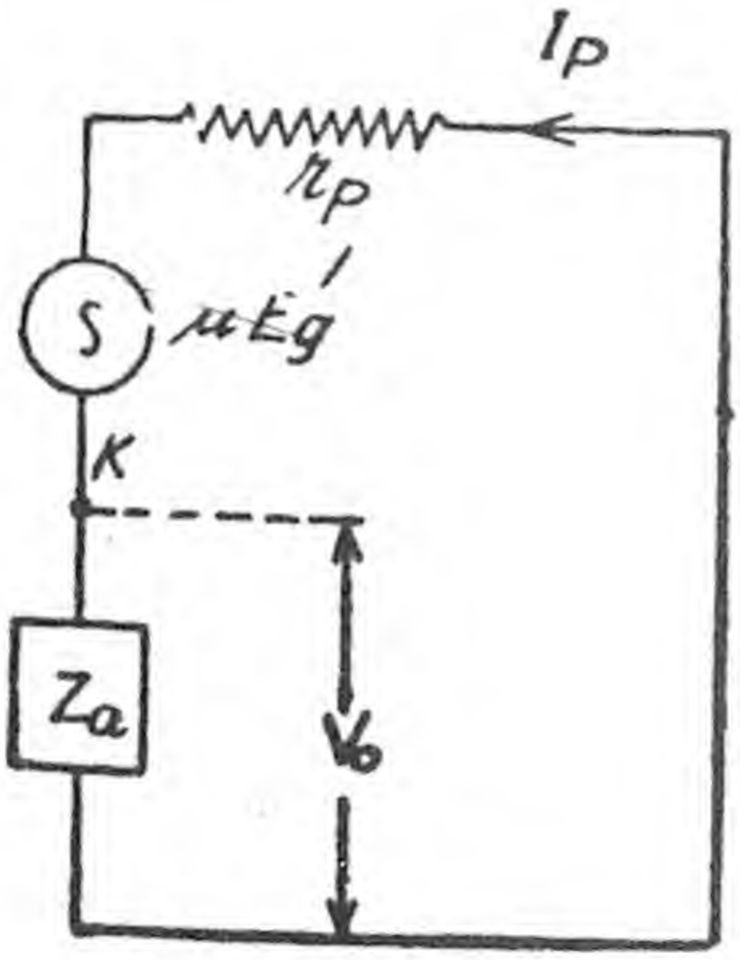


Fig. 15. Equivalent circuit of cathode follower.

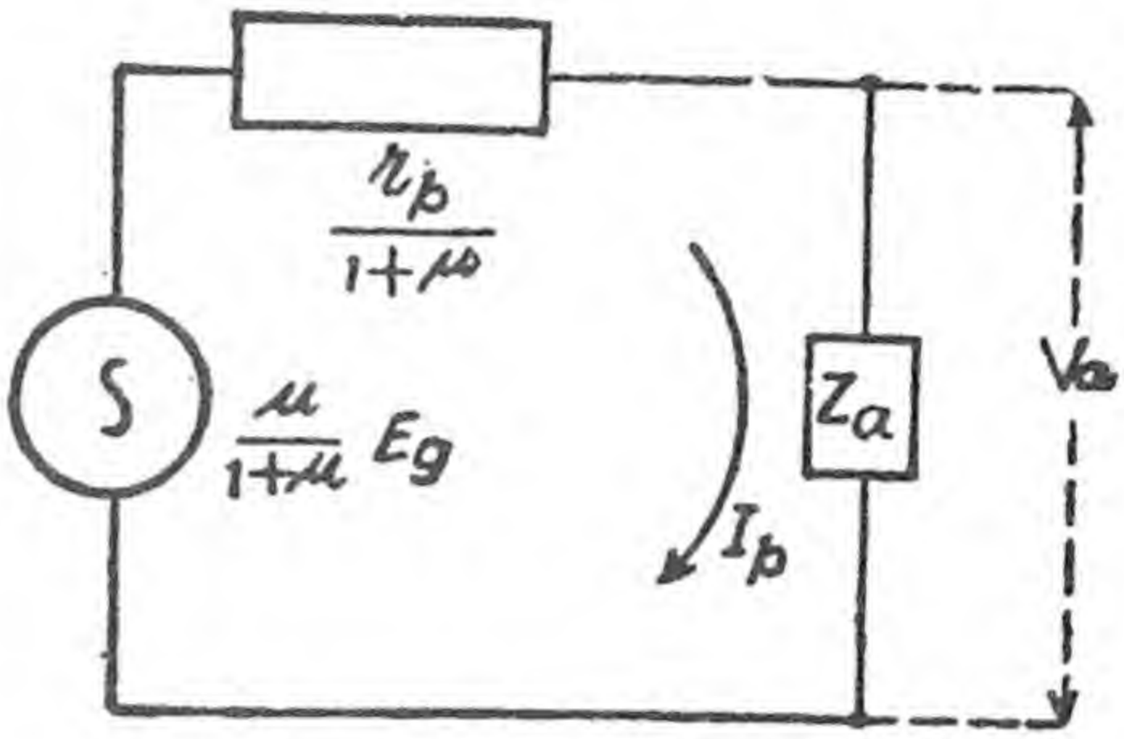


Fig. 16. Equivalent circuit.

Since $\mathbf{V_0}=\mathbf{I_p}.\mathbf{Z_a}$, the voltage gain is always less than unity and in the limit approaches the value $\frac{\mu}{(1+\mu)}$ as the ratio $\frac{r_p}{Z_a}$ approaches zero. Thus for tubes with large value of amplification factor μ and with $Z_a \gg r_p$, the voltage gain approaches unity. Since the gain is almost unity, cathode and grid rise and fall together in potential by equal amount, justifying the name 'cathode follower'.

The expression for current can be written as

$$\mathbf{I_p}=\frac{\frac{\mu}{1+\mu}\mathbf{E_g}}{\frac{r_p}{1+\mu}+\mathbf{Z_a}}$$

This equation suggests that the circuit of cathode follower may be represented as shown in fig. 16. From figure effective internal potential $\mathbf{E_f}$ and effective internal impedance $\mathbf{Z_f}$ are given by

$$\mathbf{E_f}=\frac{\mu}{1+\mu}\mathbf{E_g},$$

$$\mathbf{Z_f}=\frac{r_p}{1+\mu}.$$

Output terminal impedance

$$\mathbf{Z_{0f}}=\frac{\mathbf{Z_f}.\mathbf{Z_a}}{\mathbf{Z_f}+\mathbf{Z_a}}=\frac{\{r_p/(1+\mu)\}.\mathbf{Z_a}}{r_p/(1+\mu)+\mathbf{Z_a}}$$

$$=\frac{r_p.\mathbf{Z_a}}{r_p+\mathbf{Z_a}\,(1+\mu)}=\frac{\mu\mathbf{Z_a}}{g_m\,[r_p+\mathbf{Z_a}\,(1+\mu)]}$$

$$=\frac{\mathbf{A_f}}{g_m}$$

Also input impedance

$$= Z_i (1 - A\alpha) = Z_i (1 + A),$$

where Z_i is the input impedance of the valve alone without feedback.

Effective internal impedance is very low if $\mu \rhd 1$. Thus cathode follower has a high input impedance and low output impedance and may, therefore, be conveniently used as a coupling device between a high impedance source and a low impedance load.

Main advantage of cathode follower :

(*i*) It is used as a power amplifier with less amplitude distortion.

(*ii*) Its input impedance is very high and so the power drawn from the source is small.

(*iii*) It can be used as impedance matching device between a high impedance source and a low impedance load.

(*iv*) Output is in phase with input.

(*v*) Value of output voltage can be changed without the change in load.

11·6. RC COUPLED CATHODE FOLLOWER :

Cathode follower circuit is widely used as a coupling amplifier between stages. Circuit for such use is shown in fig. 17.

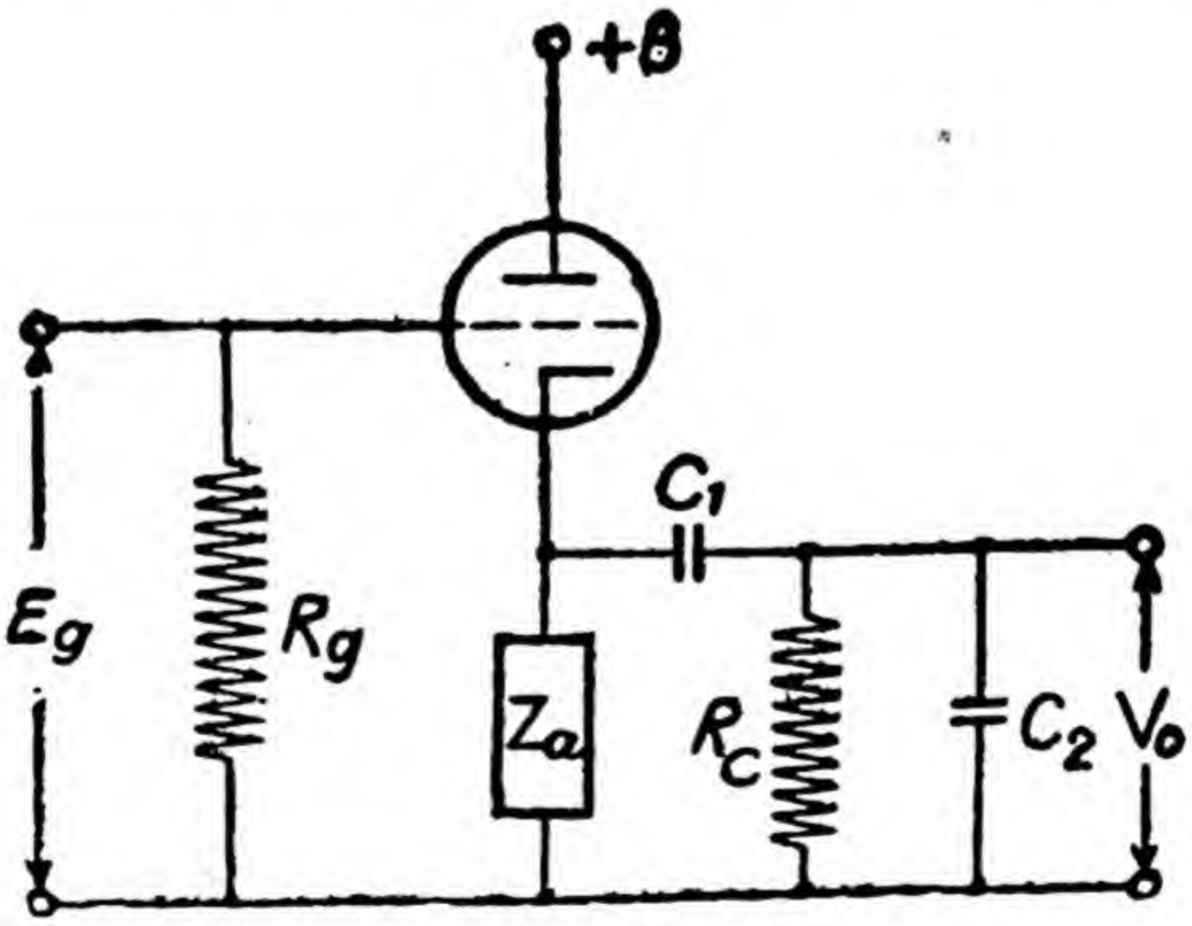


Fig. 17. RC coupled cathode follower.

In this circuit C_1 is the coupling condenser, R_e is the grid leak resistance of the next stage ond C_2 is sum of the effective output capacitance of cathode follower stage and the effective input capacitance of next stage.

Fig. 18 shows the equivalent circuit of *RC* coupled cathode follower.

It is generally more convenient and useful to divide the entire

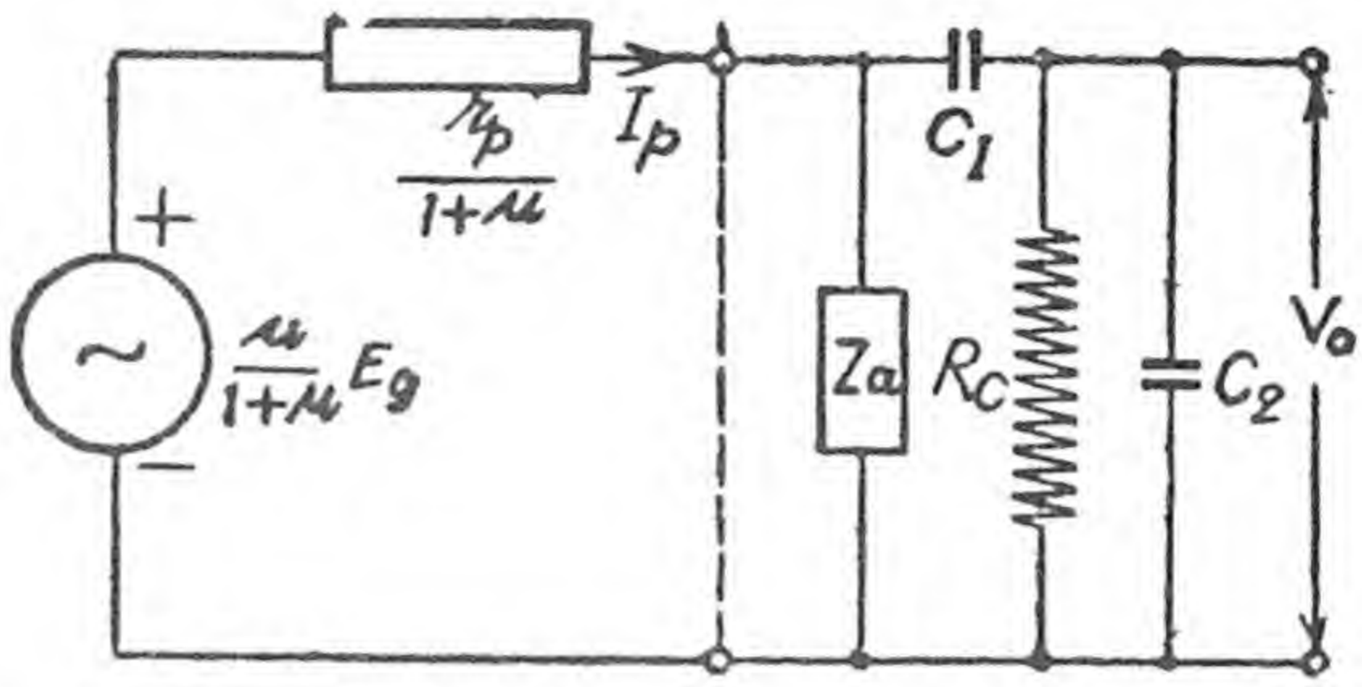


Fig. 18. Equivalent circuit of RC coupled cathode follower.

frequency range into three arbitrary frequency ranges and then to calculate voltage gain of the corresponding range.

(i) Mid Frequency Range : In this frequency range, frequency is small enough to make the reactance of C_2 large compared to Z_a and hence C_2 may be omitted. Again the frequency is large enough to make the reactance of coupling condenser C_1 small compared with Z_a and may be ignored. Thus the equivalent circuit reduces to as shown in fig. 19. The impedance Z of the parallel combination of Z_a and R_c is

$$Z_m = \frac{Z_a R_c}{Z_a + R_c}$$

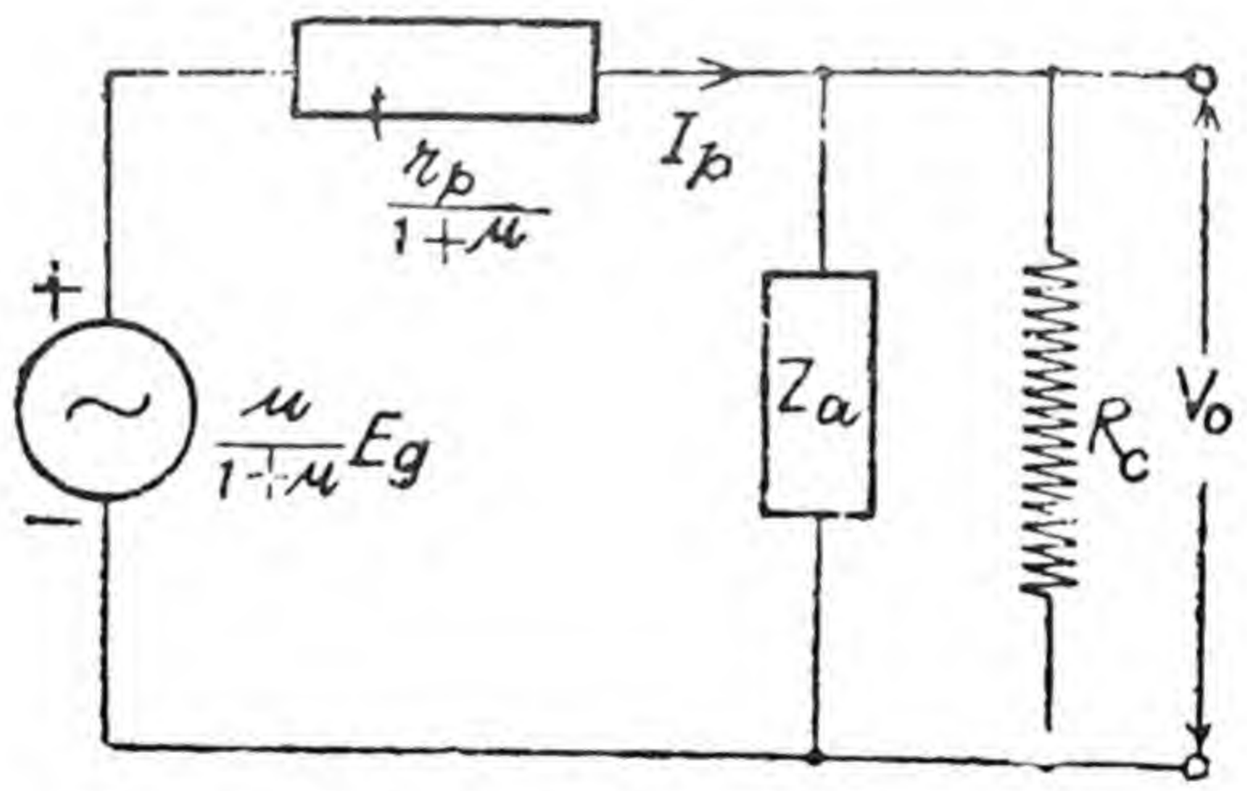


Fig. 19. Equivalent circuit at mid frequencies.

Plate current at this frequency range will be given by

$$I_p = \frac{\frac{\mu}{1+\mu} E_g}{\frac{r_p}{1+\mu} + Z_m}$$

or

$$I_p = \frac{\left(\frac{\mu}{1+\mu}\right) E_g}{\left(\frac{r_p}{1+\mu}\right) + \frac{Z_a . R_o}{Z_a + R_o}} \qquad \text{...(1)}$$

Therefore the voltage gain

$$A_{fm}=\frac{V_0}{E_g}=\frac{I_p.Z_m}{E_g}=\frac{I_p.[Z_a\ R_c/(Z_a+R_c)]}{E_g}$$

$$=\frac{\frac{\mu}{1+\mu}.E_g\ \frac{Z_a\ R_c}{(Z_a+R_c)}}{\left[\frac{r_p}{1+\mu}+\frac{Z_a\ R_c}{(Z_a+R_c)}\right]E_g}$$

$$=\frac{\mu}{r_p\left(\frac{Z_a+R_c}{Z_a\ R_c}\right)+(1+\mu)} \qquad ...(2)$$

(ii) High frequency range : At high frequencies, the reactance of coupling condenser C_1 is negligible compared with R_c. Reactance of C_2 becomes significant. Thus the equivalent circuit can be drawn as shown in fig. 20.

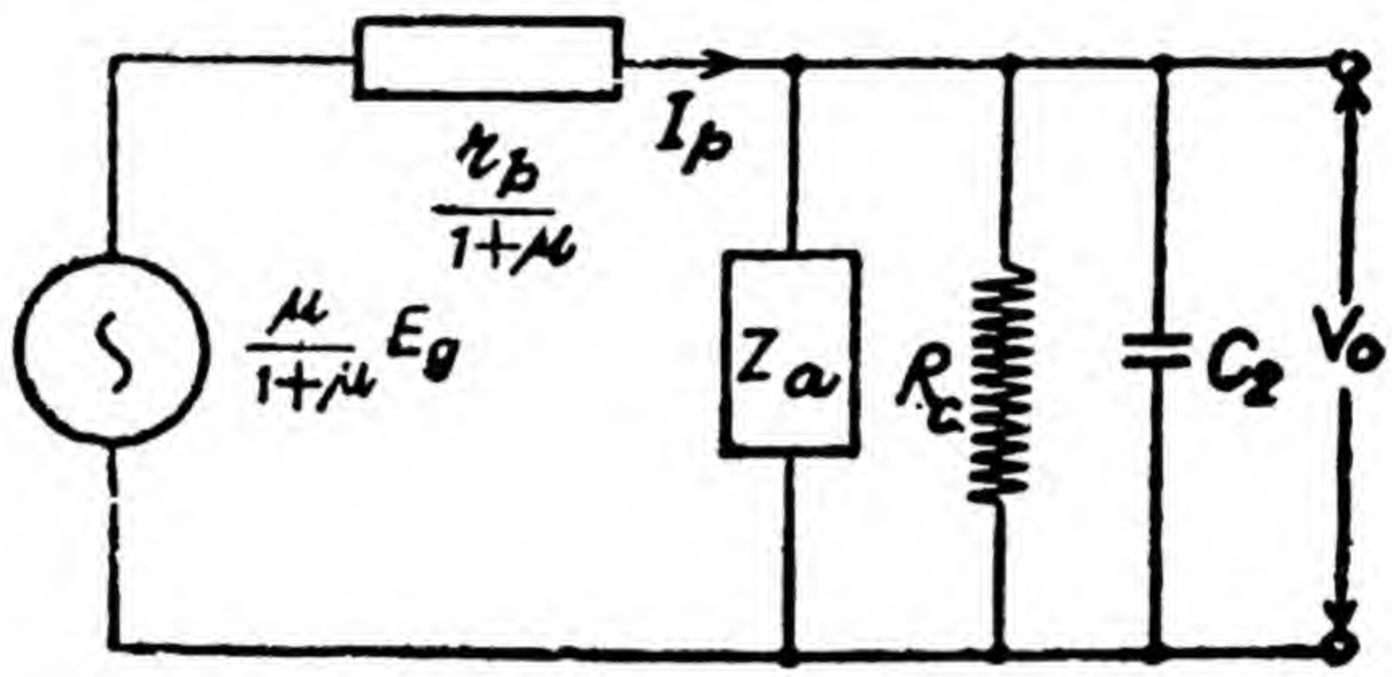


Fig. 20. Equivalent circuit at high frequencies.

We have

$$\frac{1}{Z_h}=\frac{1}{Z}+\frac{1}{R_c}+j\omega C_2. \qquad ...(3)$$

The plate current is given by $I_p=\dfrac{\frac{\mu}{1+\mu}.E_g}{\left(Z_h+\frac{r_p}{1+\mu}\right)}.$ ...(4)

Therefore the voltage gain will be given by relation

$$A_{fh}=\frac{V_0}{E_g}=\frac{I_p\ Z_h}{E_g}$$

$$=\frac{\frac{\mu}{1+\mu}.E_g Z_h}{E_g\left(Z_h+\frac{r_p}{1+\mu}\right)}=\frac{\mu\ Z_h}{Z_h\ (1+\mu)+r_p}$$

$$=\frac{\mu}{(1+\mu)+r_p/Z_h}, \qquad ...(5)$$

where Z_h is given in previous equation

Now $$\frac{A_{fh}}{A_{fm}}=\frac{\mu/\{(1+\mu)+r_p/Z_h\}}{\mu/\{(1+\mu)+r_p/Z_m\}}$$

$$= \frac{(1+\mu)+r_p/Z_m}{(1+\mu)+r_p/Z_h}$$

$$= \frac{(1+\mu)+r_p/Z_m}{1+\mu+r_p\,(1/Z_m+j\omega C_2)}$$

$$= \frac{1}{1+\dfrac{j\omega C_2}{1+\mu+r_p/Z_m}}$$

$$= \frac{1}{1+j\omega C_2 r'_p} = \frac{1}{1+j\,\dfrac{f}{f_2}} \qquad \ldots(6)$$

where $r'_p = r_p/\{\mu+1+r_p/Z_m\}$ and $f_2 = 1/\{2\pi r'_p C_2\}$

Here f_2 is the upper half power frequency. The form of equation (6) is the same for conventional *RC* coupled amplifier. The only difference is that $r'_p C_2$ is much smaller than the corresponding value in conventional *RC* coupled amplifier. Due to this fact the upper half power frequency f_2 in this case is much larger than the usual *RC* coupled amplifier. The effective internal impedance $r_p/(1+\mu)$ and the output terminal impedance is a parallel combination of $r_p/(1+\mu)$ and Z_h.

(iii) Low Frequency Range : In low frequency range, frequency is low enough to make the reactance of C_1 very large compared with Z_a and C_2 may be omitted. Then equivalent circuit is of the form as shown in figure 21.

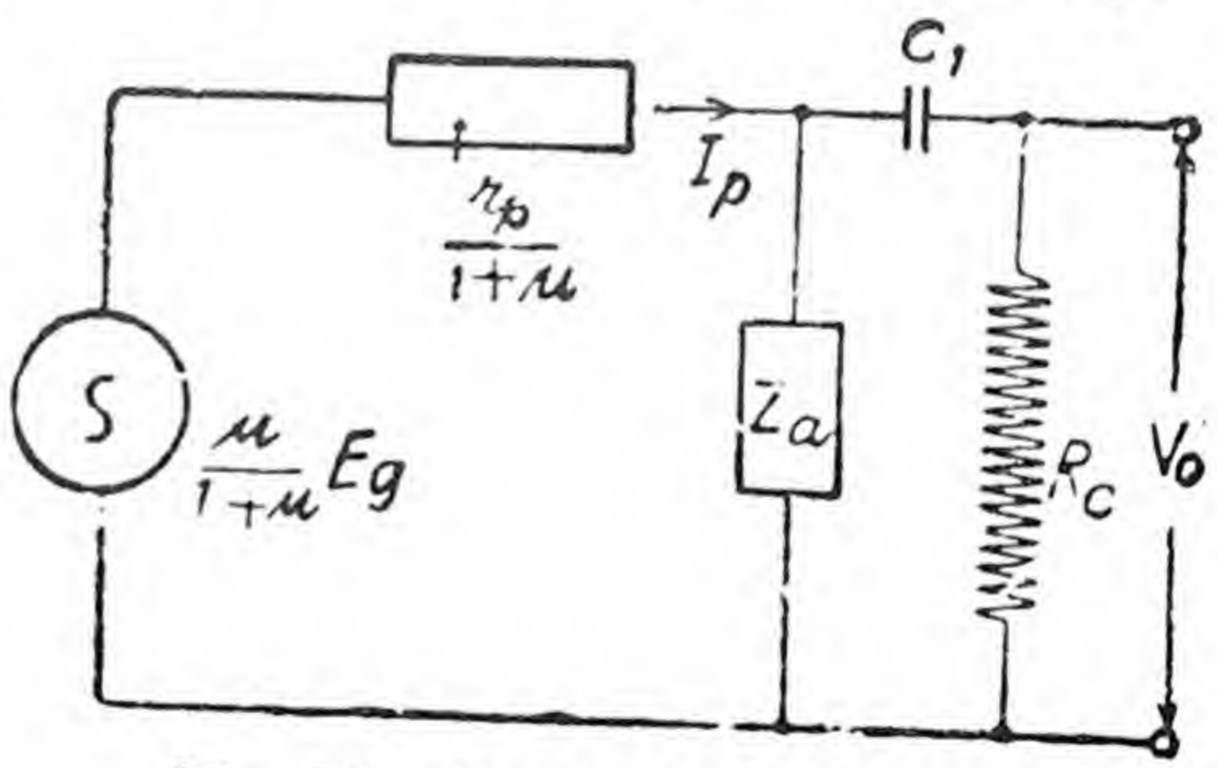


Fig. 21. Equivalent circuit at low frequencies.

The impedance of elements Z_a, C_1, R_c is

$$\frac{1}{Z_l} = \frac{1}{Z_a} + \frac{1}{R_o + \dfrac{1}{j\omega C_1}}$$

$$\text{Plate current } I_p = \frac{\dfrac{\mu}{1+\mu} E_g}{\dfrac{r_p}{1+\mu} + Z_l} \qquad \ldots(7)$$

Voltage gain at low frequency range is

$$A = \frac{I_p Z_l}{E_g} = \frac{\dfrac{\mu}{1+\mu} \cdot E_g Z_l}{\left[\left(\dfrac{r_p}{1+\mu}\right) + Z_l\right] E_g}$$

$$= (1 \frac{\mu}{(1+\mu)+r_p/Z_l} \qquad \ldots(8)$$

Now the ratio of low frequency gain to medium frequency gain can be obtained. The expression is given as follows :

$$\frac{\mathbf{A}_{fl}}{\mathbf{A}_{fm}} = \frac{1}{1 - jf_1/f} \qquad \ldots(9)$$

where $$f_1 = \frac{1}{2\pi C_1 R_o}$$

Obviously f_1 is the lower half power frequency. For typical cathode follower amplifier, the value of f_1 is of the same order of magnitude as in a conventional *RC* coupled amplifier.

11·7. GROUNDED GRID AMPLIFIER :

Sometimes this amplifier is also called as cathode input amplifier. The amplifiers are either grounded cathode (grid input) or grounded plate (cathode follower) or grounded grid (cathode input). The basic circuit of grounded grid amplifier is shown in figure 22. The input voltage is fed from a voltage source of voltage E_i and internal impedance Z_i. In this amplifier, any current flowing through the plate grid capacity will return to the plate circuit through ground instead of returning through the input circuit as in the conventional amplifiers. The control grid in this amplifier acts in a dual capacity *i.e.*, as a control grid and screen grid. It thus helps to prevent unwanted oscillations. No neutralization is needed even at very high frequencies, and so it is often used in high frequency applications. The path for the plate current in this amplifier is through the source of input voltage as well as through the load. So internal impedance of the input plays an important part in the circuit operation and should be included in the circuit. Figure 23 shows the a.c. equivalent circuit of figure 22.

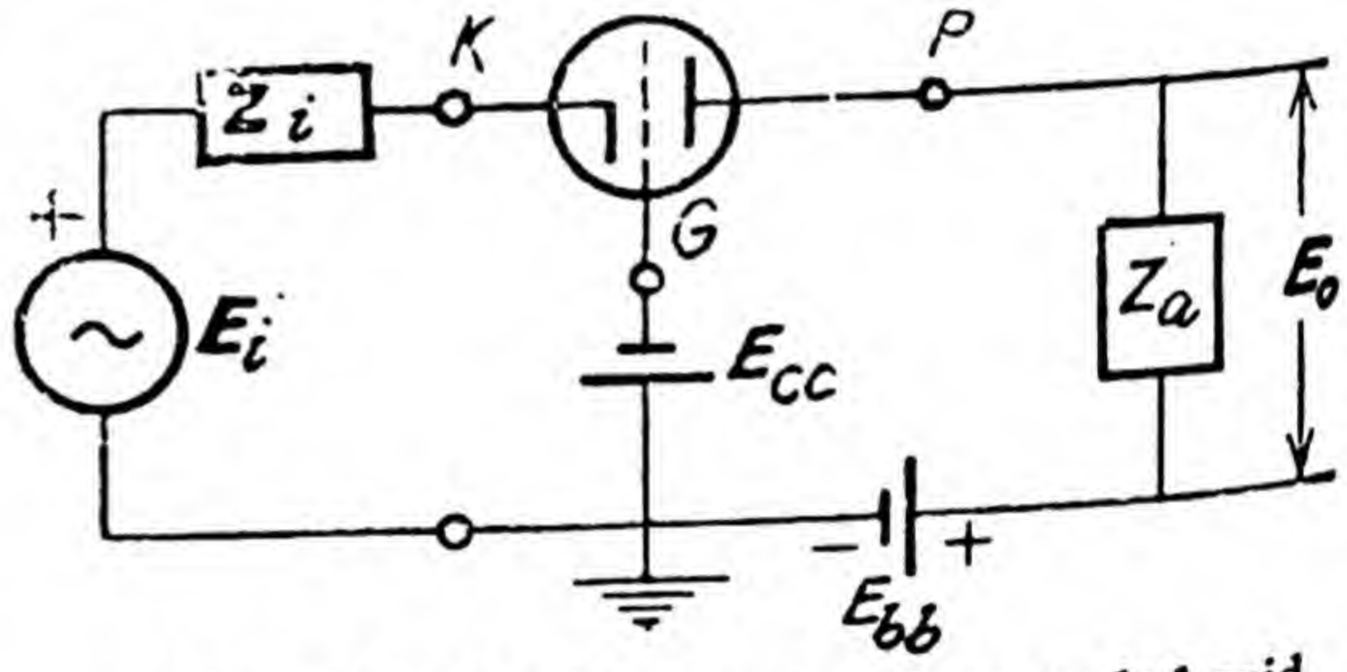


Fig. 22. Basic circuit of grounded grid amplifier.

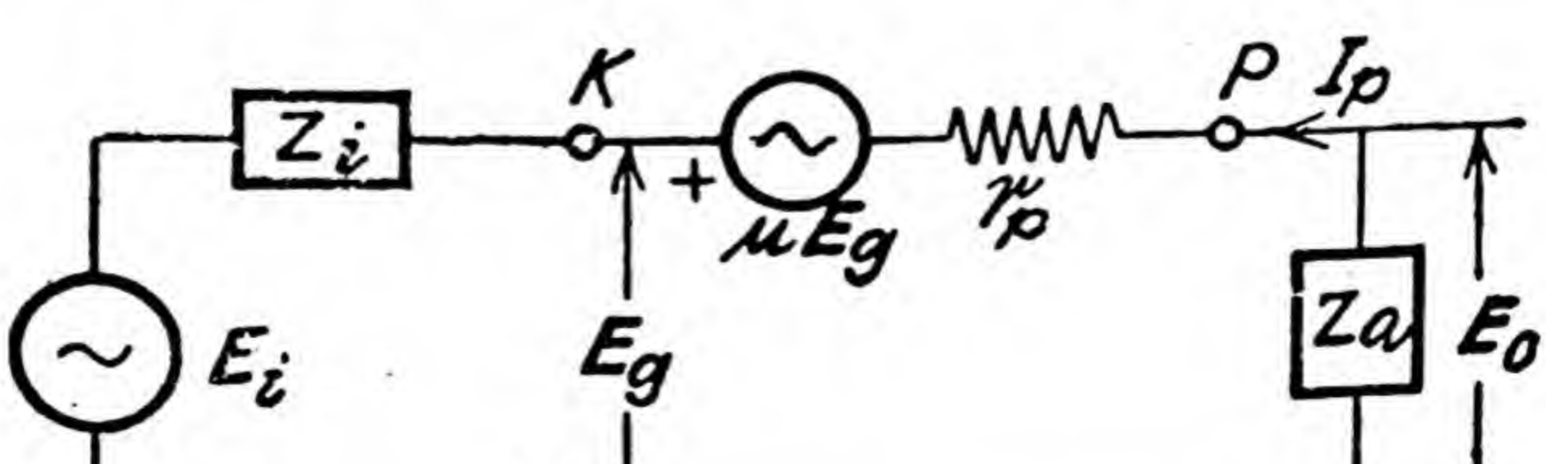


Fig. 23. a.c. equivalent circuit of figure 22.

From the equivalent circuit, the plate current $\mathbf{I}_p$ is given by

$$\mathbf{I}_p=\frac{\mu\mathbf{E}_g-\mathbf{E}_i}{\mathbf{Z}_i+r_p+\mathbf{Z}_a}$$

But in the grid circuit

$$\mathbf{E}_g=-(\mathbf{E}_i+\mathbf{I}_p\mathbf{Z}_i)$$

$$\therefore \quad \mathbf{I}_p=\frac{-\mu\,(\mathbf{E}_i+\mathbf{I}_p\mathbf{Z}_i)-\mathbf{E}_i}{\mathbf{Z}_i+r_p+\mathbf{Z}_a}$$

or $\mathbf{I}_p\,(\mathbf{Z}_i+r_p+\mathbf{Z}_a)=-(\mu+1)\,\mathbf{E}_i-\mu\mathbf{I}_p\mathbf{Z}_i$

$\mathbf{I}_p\,[(\mathbf{Z}_i+r_p+\mathbf{Z}_a)+\mu\mathbf{Z}_i]=-(\mu+1)\,\mathbf{E}_i$

or $$\mathbf{I}_p=\frac{(\mu+1)\,\mathbf{E}_i}{\mathbf{Z}_a+r_p+\mathbf{Z}_i\,(\mu+1)}. \qquad \ldots(1)$$

Equation (1) suggests that the input circuit may be represented by Thevenin's equivalent circuit as shown in figure 24. The output voltage $\mathbf{E}_0$ is given by

$$\mathbf{E}_0=-\mathbf{I}_p\mathbf{Z}_a$$

The voltage gain of the amplifier is

$$\mathbf{A}_f=\frac{\mathbf{E}_0}{\mathbf{E}_i}=\frac{-\mathbf{I}_p\mathbf{Z}_a}{\mathbf{E}_i}$$

$$=-\frac{(\mu+1)\,\mathbf{Z}_a}{\mathbf{Z}_a+r_p+\mathbf{Z}_i\,(\mu+1)} \qquad \ldots(2)$$

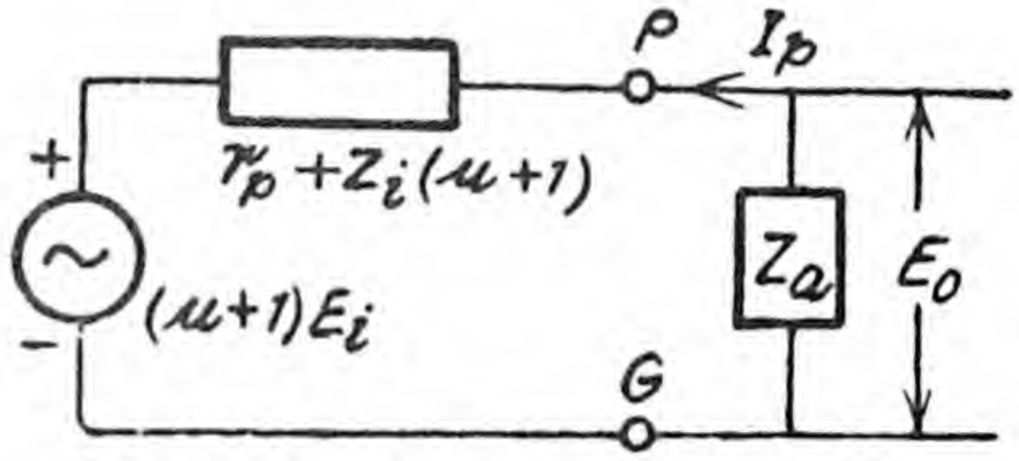


Fig. 24. Thevenin's form of a.c. equivalent circuit of grounded grid amplifier.

The input impedance $\mathbf{Z}_{in}$ is given by

$$\mathbf{Z}_{in}=\frac{\mathbf{E}_i}{-\mathbf{I}_p}+\frac{\mathbf{Z}_a+r_p+\mathbf{Z}_i\,(\mu+1)}{(\mu+1)}$$

$$=\frac{\mathbf{Z}_a+r_p}{(\mu+1)}+\mathbf{Z}_i \qquad \ldots(3)$$

Thus the input impedance presented to the input source at the grid cathode terminals is $(\mathbf{Z}_a+r_p)/(\mu+1)$ as shown in figure.

The output terminal impedance $(\mathbf{Z}_{out})_f$ of this amplifier is a parallel combination of $\mathbf{Z}_i$ and $\mathbf{Z}_a$, hence

$$(\mathbf{Z}_{out})_f=\frac{\mathbf{Z}_i\mathbf{Z}_a}{\mathbf{Z}_i+\mathbf{Z}_a} \qquad \ldots(4)$$

where $(\mathbf{Z}_{out})_f$ is the output terminal impedance with feed back.

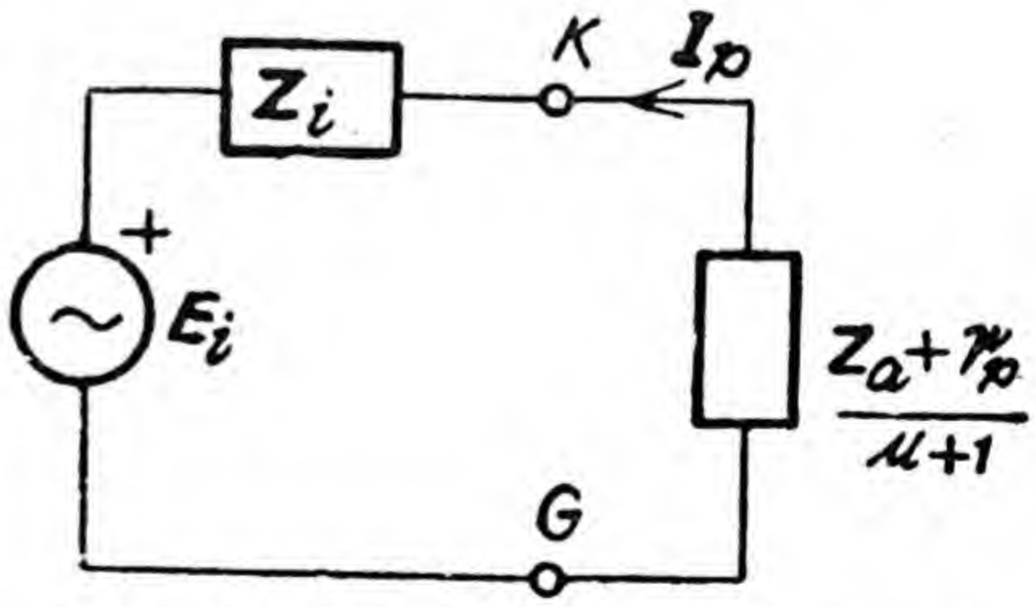


Fig. 25. Equivalent input circuit of grounded grid amplifier.

Advantage of grounded grid amplifier :

1. The grounded grid provides an electrostatic screen between the input and output of the amplifier. Thus it improves the stability of operation as an amplifier and reduces the tendency of oscillations even at extremely high frequencies.

2. There is no phase change between the output and the input voltages.

3. We know that triode is superior to tetrodes and pentodes at ultra high frequency since it introduces less random noise. In grounded grid amplifier, the control grid acts in a dual nature. The advantages mentioned in first point is obtained in screen grid tubes only but now it is also obtained in triode. Thus grounded grid triode becomes an ideal amplifier in earlier stages of ultra high frequency amplifiers and receivers.

11·8. ANODE FOLLOWER :

The anode follower has almost the same gain and impedance properties as that of cathode follower but it has a phase reversal between input and output voltages. It is a simple amplifier with a voltage feed back through an impedance. In cathode follower, the cathode potential follows the grid potential whereas in anode follower, the anode potential falls by the same amount as the grid potential rises. Thus, there is no justification for the name anode follower Because of the almost equal potential changes in the grid and plate potentials, but with reverse phase, the circuit is sometimes called as *see saw* circuit. Figure 26 shows the circuit of anode follower. The voltage feed back is provided through Z_f. Figure 27 shows the equivalent circuit of the amplifier.

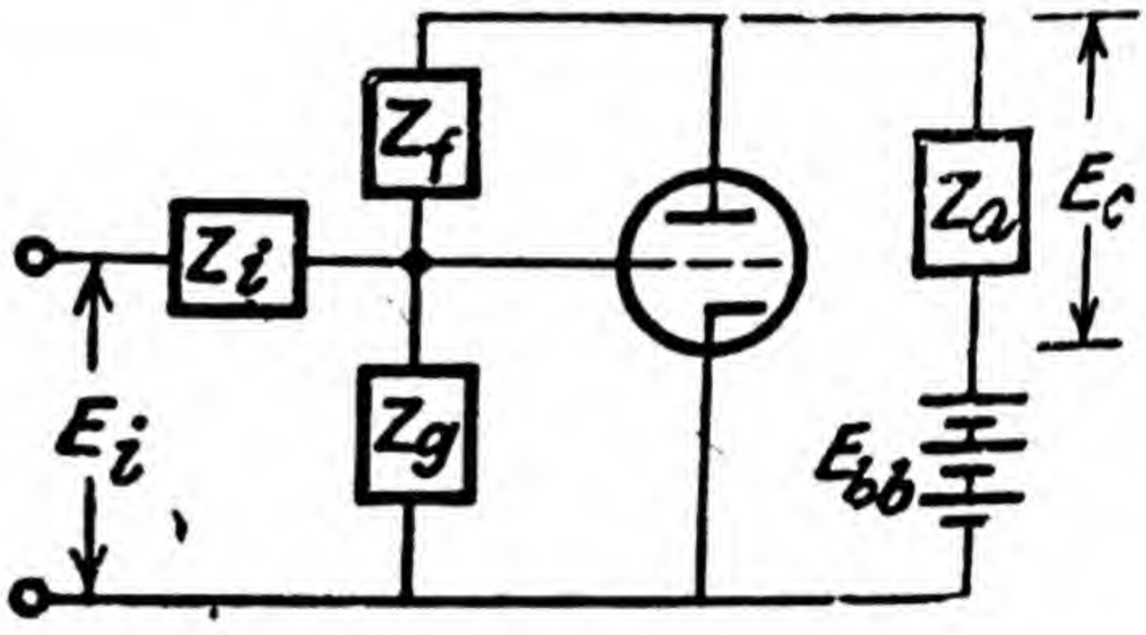


Fig. 26. Basic circuit of anode follower.

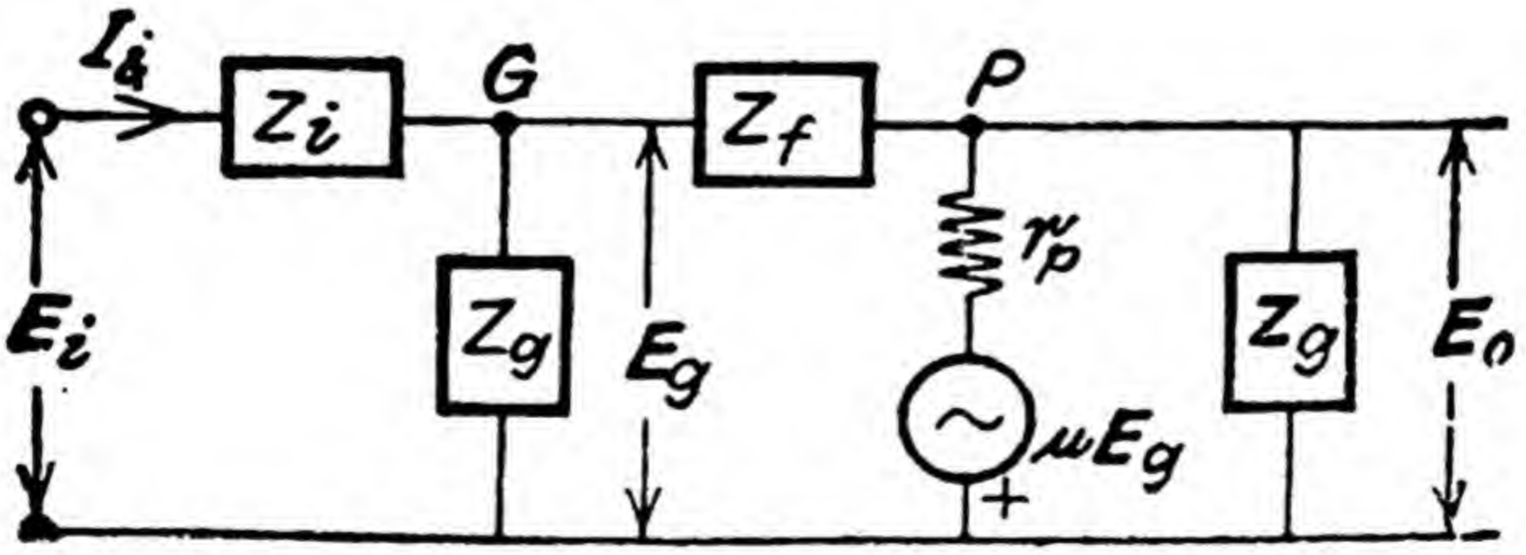


Fig. 27. a.c. equivalent circuit of figure 26.

The circuit of anode follower is simply one stage of voltage feed back amplifier with direct feed back from anode to grid through an impedance Z_f. In this case the nodal method of analysis will be easier. The node equation at G can be written as

$$\frac{E_i - E_g}{Z_i} = \frac{E_g}{Z_g} + \frac{E_g - E_0}{Z_f} \qquad \ldots(1)$$

Representing in terms of admittance, we have

$$E_g(Y_i+Y_f+Y_g)=E_iY_i+Y_fE_0$$

$$E_g=\frac{E_iY_i+Y_fE_0}{Y_i+Y_f+Y_g}, \qquad \ldots(2)$$

where $Y_i=\frac{1}{Z_i}$. $Y_f=\frac{1}{Z_f}$ and $Y_g=\frac{1}{Z_g}$.

The node equation at P is

$$\frac{E_g-E_0}{Z_f}=\frac{E_0-(-\mu E_g)}{r_p}+\frac{E_0}{Z_a} \qquad \ldots(3)$$

Representing in terms of admittances, we have

$$E_0=\frac{E_g(Y_f-g_m)}{Y_f+Y_p+Y_a}, \qquad \ldots(4)$$

where $Y_p=\frac{1}{r_p}$, $Y_a=\frac{1}{Z_a}$ and $g_m=\frac{\mu}{r_p}$

Combining equations (2) and (4), we get

$$E_0=\frac{(Y_f-g_m)}{(Y_f+Y_p+Y_a)}\times\frac{(E_iY_i+Y_fF_0)}{(Y_i+Y_f+Y_g)}$$

or $$E_0\left[1-\left(\frac{Y_f-g_m}{Y_f+Y_p+Y_a}\right)\left(\frac{Y_f}{Y_i+Y_f+Y_g}\right)\right]$$

$$=\frac{(Y_f-g_m)}{(Y_f+Y_p+Y_a)}\times\frac{E_iY_i}{(Y_i+Y_f+Y_g)}$$

Hence the voltage gain A_{fb} is given by

$$A_f=\frac{E_0}{E_i}=\frac{Y_i\,(Y_f-g_m)}{(Y_f+Y_p+Y_a)(Y_i+Y_f+Y_g)-Y_f\,(Y_f-g_m)}$$

or $$A_{fb}=\frac{Y_i(Y_f-g_m)}{(Y_i+Y_f+Y_g)(Y_p+Y_a)+Y_f(Y_i+Y_g+g_m)} \qquad \ldots(5)$$

In order to obtain the output impedance of the anode follower, the input voltage is replaced by a short circuit and a voltage generator E_0 is connected in the output circuit and the current drawn I_0 is noted. Then the output impedance is given by

$$Z_0=\frac{E_b}{I_0} \qquad \ldots(6)$$

The circuit for calculating the output impedance is shown in figure 28.

The output admittance Y_0 is

$$Y_0=\frac{I_0}{E_0}=\frac{I_p+I_g}{E_g} \qquad \ldots(7)$$

From the figure

$$I_p=\frac{F_0-(-\mu E_g)}{r_p}$$

and

$$I_g=I_1+I_2=\frac{E_g}{Z_g}+\frac{E_g}{Z_i}$$
$$=E_g\,(Y_g+Y_i)$$

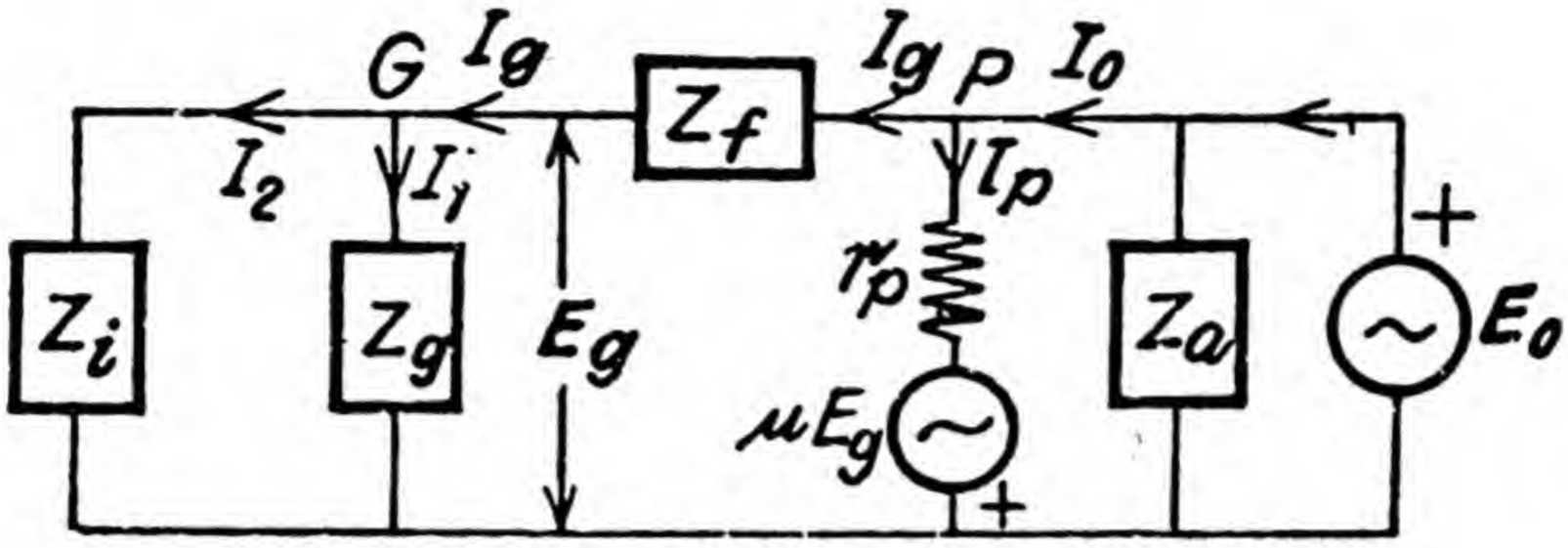


Fig. 28. Equivalent circuit for calculating the output impedance of anode follower.

$$\therefore\ Y_0=\frac{1}{E_0}\left[\frac{E_0-(-\mu E_g)}{r_p}+E_g\,(Y_g+Y_i)\right]$$
$$=\left[\,Y_p+\frac{E_g}{E_0}\,(\mu Y_p+Y_g+Y_i)\,\right] \qquad ...(8)$$

Again $$E_g=\frac{Y_f E_0}{Y_i+Y_f+Y_g} \qquad ...(9)$$

Substituting the value of E_g from equation (9) in equation (8), we get

$$Y_0=Y_p+\frac{Y_f\,(g_m+Y_g+Y_i)}{Y_i+Y_f+Y_g} \qquad ...(10)$$

NUMERICAL EXAMPLES

1. *An amplifier has a gain of* 400. *When the negative feed back is applied, the gain is reduced to* 300. *Find the feed back ratio.*

We know that $$A_f=\frac{A}{1+A\alpha}$$

$$\therefore\quad 300=\frac{400}{1+400\alpha}$$

gr
$$1+400\alpha=\frac{400}{300}=\frac{4}{3}$$
$$3+1200\alpha=4$$
$$\therefore\quad \alpha=\frac{1}{1200}$$

2. *The voltage gain of an amplifier is* 100 *when the plate supply voltage is* 200 *volts. When negative feed back is applied, the gain, with plate supply voltage of* 200 *volts, is reduced by a*

factor of 40. *Calculate voltage gain with feed back for the two plate supply voltages.*

We know that $A_f = \frac{A}{1+A\alpha}$

With plate supply voltage of 200 volts,

$$\frac{1000}{40} = \frac{1000}{1+A\alpha'}$$

i.e. $1+A\alpha = 40$

$A\alpha = 39$ (feed back factor),

$$\alpha = \frac{39}{1000}$$

$$= 0{\cdot}039$$

Then $A_f = \frac{1000}{40} = 25.$

With plate supply voltage of 150 volts,

$$A_f = \frac{800}{1+800\ (0{\cdot}039)}$$

$$= 24{\cdot}84.$$

3. *Referring to current feed back amplifier cathode resistance* $R_o = 1000$ *ohms, impedance* $Z_a = 50{,}000$ *ohms,* $\mu = 20$, $r_p = 20{,}000$. *Calculate*

(*a*) *voltage gain of amplifier,*
(*b*) *internal impedance,*
(*c*) *output terminal impedance,*

For the two cases :

(*i*) output developed across Z_a,
(*ii*) output devoloped across R_c.

Case (i) : (*a*) voltage gain of the amplifier is given by

$$A_f = \frac{\mu Z_a}{r_p + R_o\ (1+\mu) + Z_a}$$

$$= \frac{20 \times 50{,}000}{2 \times 10^4 + 21 \times 10^3 + 50000}$$

$$= 10{\cdot}98.$$

(*b*) Internal impedance $= r_p + R_c\ (1+\mu)$

$$= 2 \times 10^4 + 21 \times 1000$$

$$= 41 \times 10^3 \text{ ohms.}$$

(*c*) Output terminal impedance

$$= \frac{\text{internal impedance} \times Z_a}{\text{internal impedance} + Z_a}$$

$$= \frac{41 \times 10^3 \times 5 \times 10^4}{41 \times 10^3 + 5 \times 10^4}$$

$$= 22{\cdot}5.10^3 \text{ ohms.}$$

Case (ii) : (*a*) Output developed across R_c,

$$A_f = \frac{\mu R_c}{r_p + (\mu+1)\, R_c + Z_a}$$

$$= \frac{20 \times 1000}{2 \times 10^4 + 21 \times 1000 + 5 \times 10^4}$$

$$= 0{\cdot}219.$$

(*b*) Internal impedance $= \frac{r_p + Z_a}{1+\mu}$

$$= \frac{2 \times 10^4 + 5 \times 10^4}{21}$$

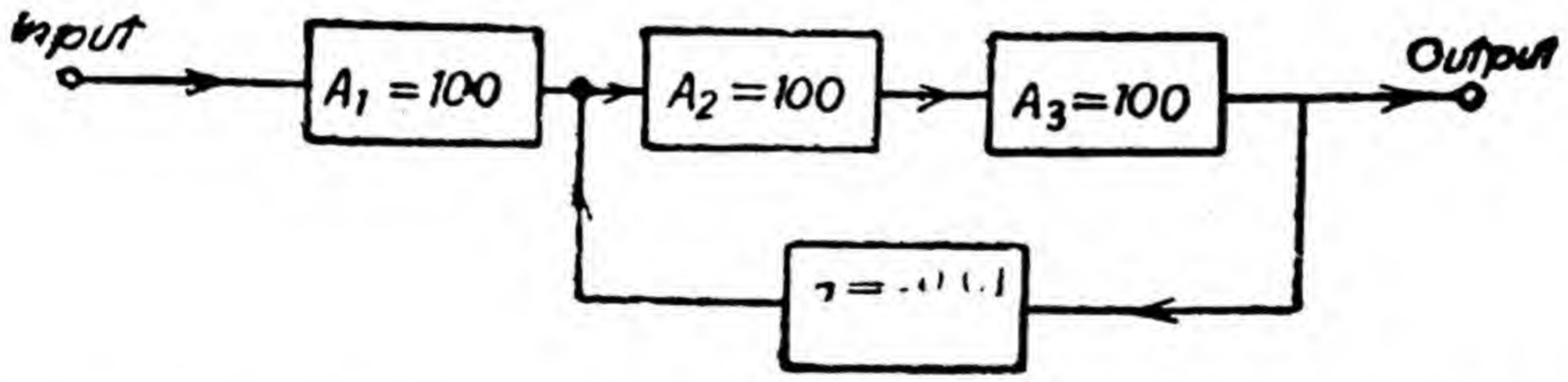
$$= 3333{\cdot}3\Omega.$$

(*c*) Output terminal impedance $= \frac{\text{Internal impedance} \times R_o}{\text{Internal impedance} + R_o}$

$$= \frac{3333{\cdot}3 \times 1000}{3333{\cdot}3 + 1000}$$

$$= 775{\cdot}1\Omega.$$

4. *An amplifier has three stages each with a gain of* 100. *Negative feed-back is now applied by feeding* $-\frac{1}{100}$ *of the output voltage to the input of the second stage. Find the overall voltage gain of the amplifier with feed-back. Again if the negative feed-back is applied to the input of the first stage, calculate the overall voltage gain with feed-back.*



$\mathbf{A}$ = voltage gain of the last two stages without feed-back

$= 100 \times 100 = 10{,}000.$

For feed-back to be negative, feed-back factor $A\alpha$ must be negative.

Since $\mathbf{A} = +10000$, α must be negative ; then

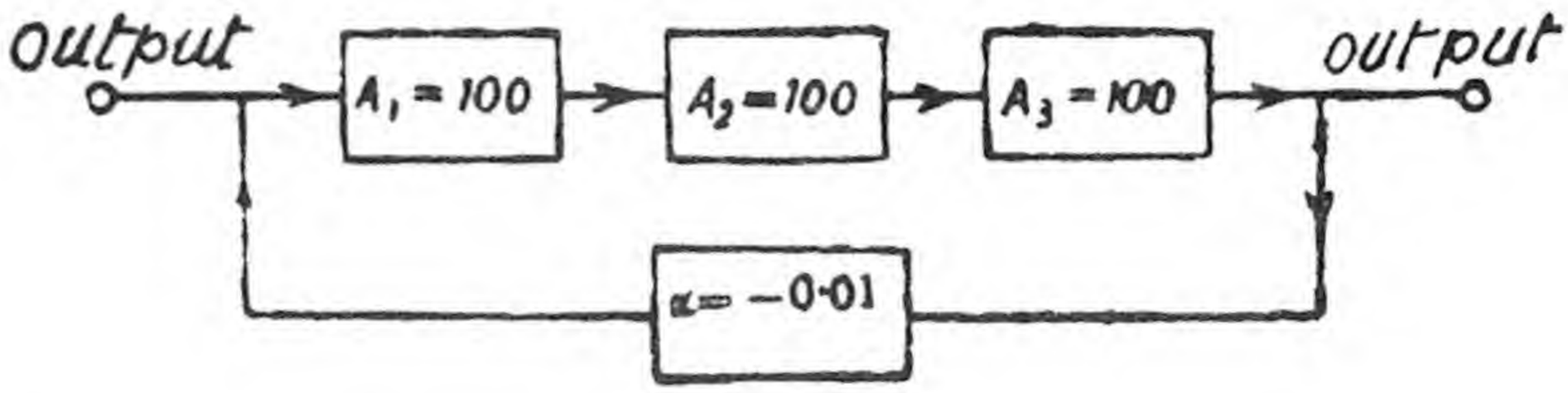
$$\alpha' = -0{\cdot}01.$$

Hence voltage gain of the last two stages with feed-back

$$= \frac{\mathbf{A}}{1 - \mathbf{A}\alpha} = \frac{10^4}{1 + 10^4 \times 0{\cdot}01} = 99.$$

Overall gain of all the three stages $= 99 \times 100$

$$= \mathbf{9900}.$$



Again in the second case,

$$\alpha = 0{\cdot}01,$$

$$\mathbf{A} = 100 \times 100 \times 100 = 10^6.$$

Hence overall gain in the three stages with negative feed-back

$$= \frac{\mathbf{A}}{1 - \mathbf{A}\alpha} = \frac{10^6}{1 + 10^6 \times 0{\cdot}01} = 99{\cdot}99.$$

5. *An amplifier has an amplification of* $(2000 + j\,0)$ *at* 2 *kc/s and* $(600 - j\,1000)$ *at* 20 *kc/s. Compare the values of these amplifications if* 5% *of the output voltage is fed back to the input as a negative feed back.*

In the given problem $\alpha = -5/100 = -0{\cdot}05$

(*a*) At 2*kc/s*. $\mathbf{A} = 2000 + j0$

$$\therefore \mathbf{A}_f = \frac{A}{1 - \alpha A} = \frac{2000}{1 + 2000 \times 0{\cdot}05} = 19{\cdot}8 \angle 0°$$

(*a*) At 20 *kc/s*. $\mathbf{A} = 600 - j\,1000$

$$= \sqrt{\{(600)^2 + (1000)^2\}}\ \angle \tan^{-1}\left(-\frac{1000}{600}\right)$$

$$= 1166 \angle -59°$$

$$\therefore \mathbf{A}_f = \frac{1166 \angle -59°}{1 + (600 - j\,1000)\ 0{\cdot}05} = \frac{116 \angle -59°}{31 - j\,50}$$

$$= \frac{1166 \angle -59°}{52 \angle -58{\cdot}2}$$

$$= 198 \angle -0{\cdot}8°$$

6. *An amplifier has a voltage gain of* 200, *distortion of* 10% *and a normal input of* 0 5 *volts, before negative feed back is applied. When negative feed back with* $\alpha = 0{\cdot}05$ *is applied, then find the new gain and distortion. If the amplifier is required to produce the same output voltage as before feed back, then find the new value of the input voltage.*

In case of negative feed back, the gain is given by

$$\mathbf{A}_f = \frac{\mathbf{A}}{1 + \mathbf{A}_f}$$

$$= \frac{200}{1 + {\cdot}05 \times 200} = \frac{200}{11} = 18{\cdot}8.$$

Distortion D_n' with negative feed back is

$$D_n' = \frac{D_n}{1+\alpha A} = \frac{10}{1+\cdot 05\times 200} = \frac{10}{11} = 0\cdot 909\%.$$

For the same output as before feedback, we have

$$A_f E_s = A E_g$$

$$E_s = \frac{A}{A_f} E_g = (1+\alpha A)\ E_g$$

$$= (1+\cdot 05\times 200)\ 0\cdot 5 = 5\cdot 5 \text{ volts.}$$

7. *In a cathode follower, the different values are , amplification factor* $\mu = 50$, *plate resistance* $r_p = 20,000$ *ohms, cathode resistance* $Z_a = 2000$ *ohms.*
Calculate

(*a*) *Voltage gain,*
(*b*) *Internal impedance,*
(*c*) *Output impedance.*

(*a*) Voltage gain $= \dfrac{\mu Z_a}{r_p + Z_a (1+\mu)}$

$$= \frac{50\times 2000}{20{,}000 + (50+1)\ 2000}$$

$$= 0\cdot 81.$$

(*b*) Plate current in cathode follower $I_p = \dfrac{\mu E_g}{r_p + (1+\mu) Z_a}$

$$= \frac{\frac{\mu}{1+\mu} E_g}{\frac{r_p}{1+\mu} + Z_a}$$

which shows that the internal impedance is

$$= \frac{r_p}{1+\mu} = \frac{20.000}{51} = 392\cdot 1 \text{ ohms.}$$

(*c*) Output impedance $= \dfrac{r_p Z_a}{r_p + Z_a (1+\mu)}$

$$= \frac{20{,}000\times 2{,}000}{20{,}000 + 2{,}000\ (1+50)}$$

$$= \frac{40{,}000}{20+102} = \frac{40{,}000}{122} = 327\ 1 \text{ ohms.}$$

8. *A triode employed as a cathode follower has amplification factor of* 40 *and mutual conductance of* 4 *m.A/volt. The load consists of a capacitance of* 1000 $\mu\mu f$ *Calculate the magnitude and phase angle of the voltage gain of the amplifier at a frequency of* 79·6 kH_z.

In the given problem $g_m = 4\times 10^{-3}$ *mho*; $\mu = 40$

$$\therefore\ r_p = \frac{\mu}{g_m} = \frac{40}{4\times 10^{-3}} = 10^4 \text{ ohms.}$$

Reactance of the condenser X_C is given by

$$X_C=\frac{1}{2\pi fC}=\frac{1}{2\times 3\cdot 14\times 79\cdot 6\times 10^3\times 1000\times 10^{-12}}$$

$=2000$ ohms.

Now $\mathbf{I}_p$ is given by

$$\mathbf{I}_p=\frac{\mu \mathbf{E}_g}{r_p+\mathbf{Z}_a\,(1+\mu)},$$

and the voltage gain is given by

$$\mathbf{A}_f=\frac{\mathbf{V}_0}{\mathbf{E}_g}=\frac{\mu \mathbf{Z}_a}{r_p+(1+\mu)\,\mathbf{Z}_a}=\frac{40\times(-j\,2000)}{10^4+(40+1)\,(-j\,2000)}$$

$$=\frac{-j\,80\times 10^3}{10-j\,8200}$$

$$\text{Magnitude } A_f=\frac{80\times 10^{-3}}{\sqrt{\{100+(82)^2\times 10^3\}}}=0\cdot 9685.$$

$$\text{Phase angle}=270+\tan^{-1}8\cdot 2=270^\circ+83\cdot 1^\circ$$

$$=353\cdot 1^\circ.$$

9. *The amplification factor of a valve is* 20 *and dynamic plate resistance is* 8,000 *ohms. Anode load* $\mathbf{Z}_a=20{,}000$ *ohms. Cathode bias resistance* $=1000$ *ohms. An audio signal of* 1 *mV r.m.s. is applied between grid and earth. Calculate*

(*i*) *a.c. output voltage when no cathode bypass condenser is used, and*

(*ii*) *voltage gain for* 1000 *c/s. when cathode bypass condenser is* 10 μF.

(*iii*) *voltage gain with fixed bias.*

In this case the feed back is through cathode resistance R_c.

(*i*) $$A_f=\frac{\mu \mathbf{Z}_a}{r_p+R_c\,(1+\mu)+\mathbf{Z}_a}$$

$$=\frac{20\times 20{,}000}{8\,000+1000\,(1+20)+20000}$$

$$=\frac{400000}{49000}=8\cdot 16$$

Now input voltage is 1mV

$\therefore$ a.c. output $=8\cdot 16\times 1=$ **8·16 mV.**

(*ii*) Let the cathode impedance be $\mathbf{Z}_c$, then

$$\frac{1}{\mathbf{Z}_c}=\frac{1}{R_c}+\frac{1}{-jX_c}=\frac{1}{R_c}+j\omega C$$

$$\mathbf{Z}_c=\frac{R_c}{1+j\omega C\,R_c}=\frac{1000}{1+j\;2\pi f.\;10\times 10^{-6}\times 1000}$$

$$=\frac{1000}{1+j.2\pi f.10^{-2}}=\frac{1000}{1+j.2\pi.\;1000+10^{-2}}$$

$$=\frac{1000}{1+j\,20\pi}=\frac{1000}{1+j\,62{\cdot}8}=0{\cdot}25\ (1-j\,62{\cdot}8)$$

$$A_f=\frac{20\times 20{,}000}{8000+0{\cdot}15\ (1-j\ 62{\cdot}8)\times 21+20{,}000}$$

$$=\frac{20\times 20{,}000}{8000+5{\cdot}25\ (1-j\ 62{\cdot}8)+20{,}000}=\frac{20\times 20{,}000}{28005{\cdot}25-j\ 329{\cdot}7}$$

$j\ 329{\cdot}7$ may be neglected when compared to $28005{\cdot}25$

$$\therefore\quad A_f=\frac{20\times 20{,}000}{28005{\cdot}25}=14{\cdot}3.$$

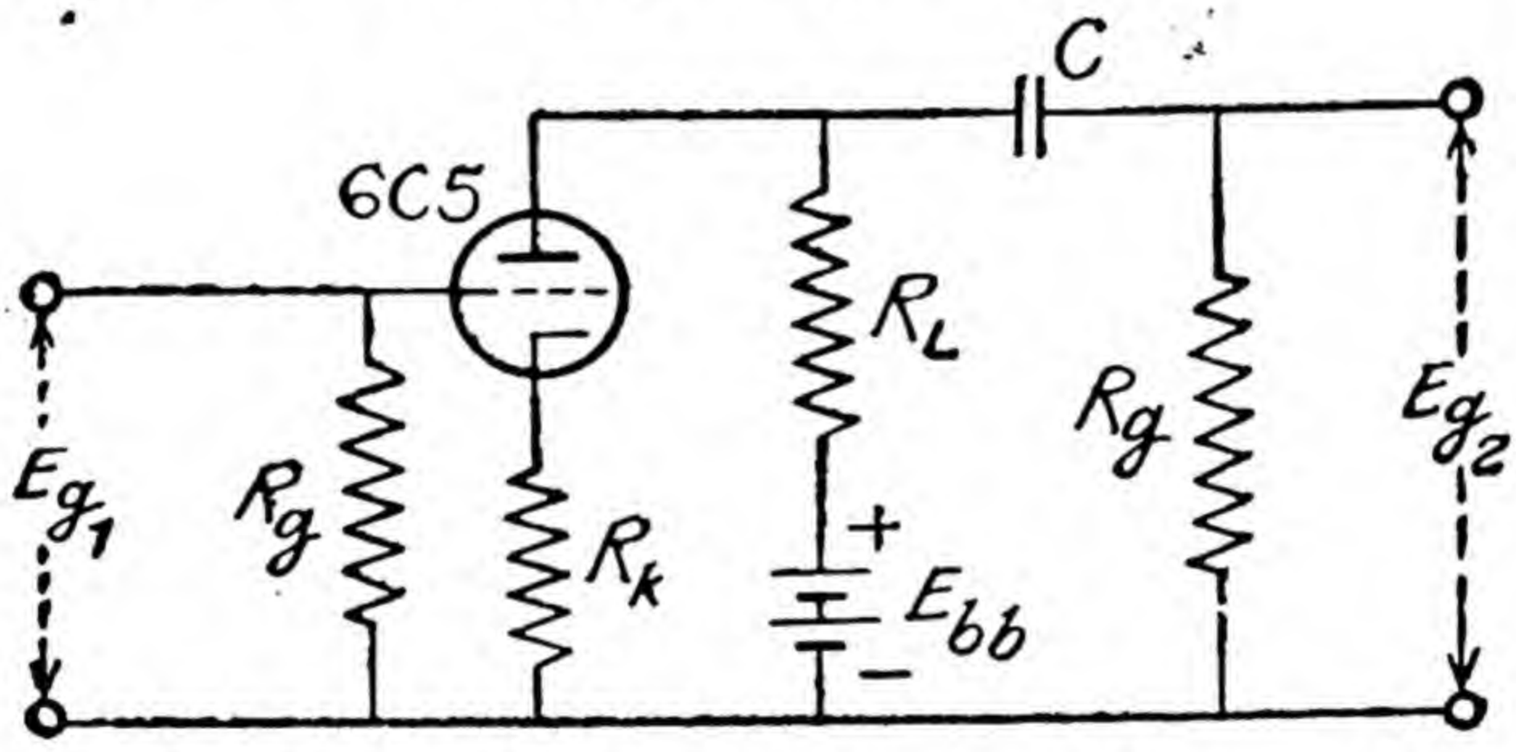
(iii) With fixed bias $A=\dfrac{\mu Z_a}{r_p+Z_a}=\dfrac{20\times 20000}{8000+20{,}000}$

$=14{\cdot}28.$

10. *Following values related to the single stage RC coupled amplifier with current inverse feed back are :*

$R_l=20$ *kilo ohms,* $R_g=1$ *mega-ohm,* $E_{bb}=250$ *volts,* $\mu=20$, $g_m=2000$ *μ-mhos.*

Reactance of coupling condenser C may be neglected. Shunt capacitance may also be neglected. Calculate the value of cathode resistor R_k *to produce voltage gain of* -10.



Let Z_L be the impedance of the parallel combination of R_g and R_l. Hence

$$Z_L=\frac{R_gR_l}{R_g+R_l}=\frac{(1000\times 10^3)\ (20\times 10^3)}{(1000\times 10^3)+(20\times 10^3)}$$

$=19{\cdot}6\times 10^3$ ohms.

$$r_p=\frac{\mu}{g_m}=\frac{20}{2000\times 10^{-6}}=10^4 \text{ ohms.}$$

$$\text{Voltage gain}=\frac{-\mu Z_L}{r_p+(\mu+1)\ R_k+Z_L}$$

$$-10=-\frac{20\times 19{\cdot}6\times 10^3}{10^4+21\ R_k+19{\cdot}6\times 10^3}$$

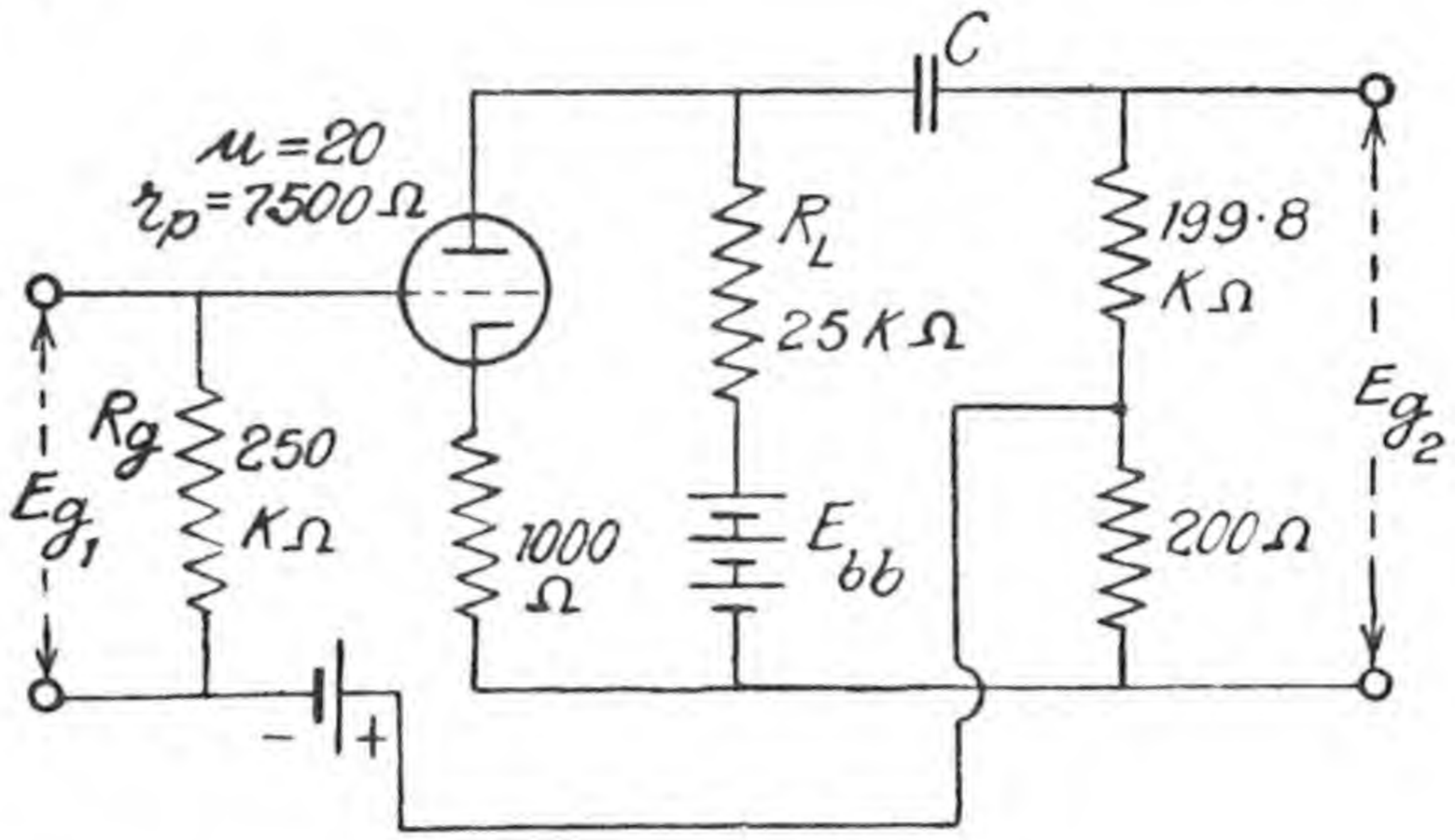
or $\quad 21R_k+(29{\cdot}6)\ 10^3=2\times 19{\cdot}6\times 10^3$

$$21\ R_k=39{\cdot}2\times 10^3-29{\cdot}6\times 10^3,$$

$$R_k=\frac{9{\cdot}6\times 10^3}{21}$$

$=0{\cdot}45\times 10^3$ ohms.

11. ***For the amplifier shown, calculate the voltage gain and effective internal impedance. Reactance of coupling condenser may be neglected.***



From the figure feed-back ratio,

$$\alpha=\frac{200}{200+(199\cdot8\times10^3)}=0\cdot001.$$

The load impedance

Z_L=parallel combination of 25 $K\Omega$ and 200 $K\Omega$

$$=\frac{(25\times10^3)\times(200\times10^3)}{(200+25)\times10^3}$$

$$=22\cdot22\times10^3 \text{ ohms.}$$

Overall gain in compound feed-back is given by

$$=\frac{\mu Z_L}{r_p+(1+\mu)\ R_c+(1+\mu\alpha)\ Z_L} \qquad \ldots(1)$$

$$=\frac{20\times22\cdot22\times10^3}{7600+(20+1)\ 10^3+(1+20\times0\cdot001)\ 22\cdot22\times10^3}$$

$$=8\cdot76.$$

Effective internal impedance can be obtained from equation (1) in the following manner :

$$=\frac{\frac{\mu}{1+\mu\alpha}Z_L}{\frac{r_p+(\mu+1)\ R_c}{1+\mu\alpha}+Z_L}.$$

Hence internal impedance

$$=\frac{r_p+(\mu+1)\ R_c}{1+\mu\alpha}=\frac{7500+(20+1)\times1000}{1+20+0\cdot001}$$

$$=28000\ \Omega.$$

12. ***In a RC coupled cathode follower as shown on page 516, R_c=1000 ohms, amplification factor μ=40, dynamic plate resistance r_p=10,000 ohms, grid bias resistance $R_g=25\times10^4$ ohms, coupling condenser C=·01 μf, grid input capacity of the following stage C_S=100 pf. Calculate the (a) voltage gain (b) output impedance***

of the the amplifier in the medium range of frequency, (c) voltage gain at 1·5 mc/s.

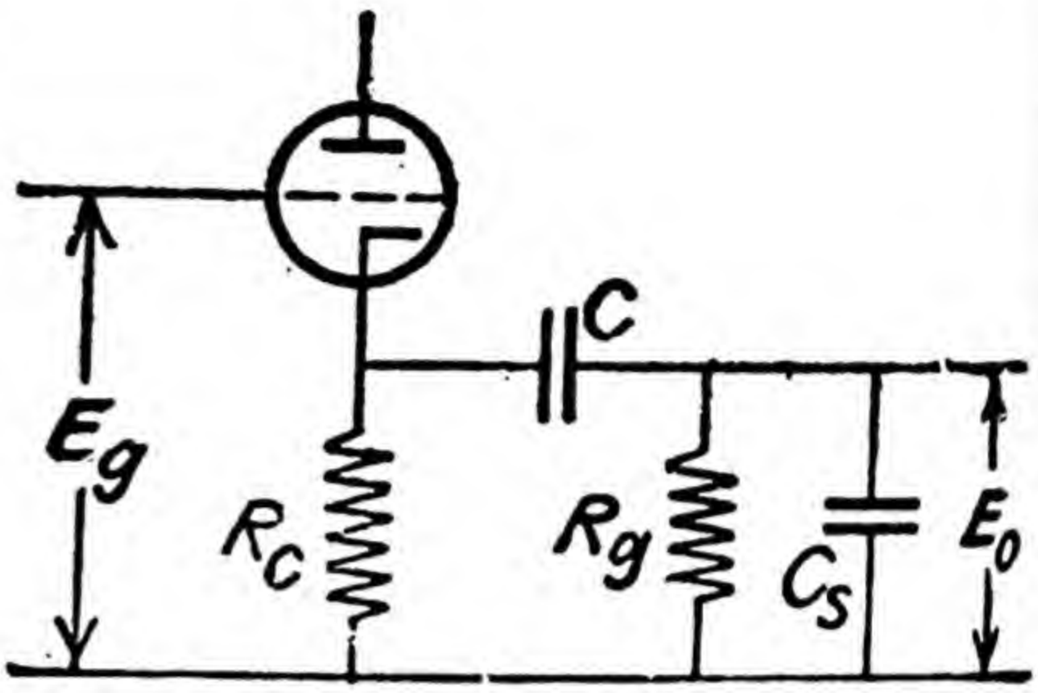
(a) In medium frequency range the effect of C and C_s may be neglected.

The load impedance in this case would be parallel combination of R_c and R_g. Let it be Z_L.

$$\therefore \quad \frac{1}{Z_L}=\frac{1}{R_c}+\frac{1}{R_g}=\frac{R_g+R_c}{R_cR_g}$$

$$\text{or} \quad Z_L=\frac{R_cR_g}{R_g+R_c}=\frac{1000\times 25\times 10^4}{1000+25\times 10^4}$$

$$=996 \text{ ohms.}$$



RC coupled cathode follower.

The medium frequency gain A_{mf} in this case is given by

$$A_{mf}=\frac{\mu Z_L}{r_p+Z_L(1+\mu)}$$

$$=\frac{40\times 996}{10{,}000+996\times 41}=0\cdot 78$$

(b) The internal impedance $Z_t=\dfrac{r_p}{1+\mu}$

$$Z_t=\frac{10{,}000}{41}=243\cdot 9 \text{ ohms.}$$

The output impedance Z_0 would be parallel combination of Z_t and Z_L.

$$\therefore \quad \frac{1}{Z_0}=\frac{1}{Z_t}+\frac{1}{Z_L} \text{ or } Z_0=\frac{Z_tZ_L}{Z_t+Z_L}$$

$$=\frac{243\cdot 9\times 996}{243\cdot 9+996}=193\cdot 4 \text{ ohm.}$$

(c) In high frequency range C is neglected and let the load impedance be Z_L'.

$$\text{Now} \quad \frac{1}{Z_L'}=\frac{1}{R_c}+\frac{1}{R_g}+\frac{1}{-j/\omega C_S}=\frac{1}{R_c}+\frac{1}{R_g}+j\omega C_S$$

$$=\frac{1}{1000}+\frac{1}{25\times 10^4}+j\times 2\pi\times 1\cdot 5\times 10^6\times 100\times 10^{-12}$$

$$=(10\,04+j\,9\cdot 4)\,10^{-4}.$$

Voltage gain at 1·5 mc/s is given by

$$A_{hf}=\frac{\mu Z_L'}{r_p+Z_L'(1+\mu)}=\frac{\mu}{\frac{r_p}{Z_L'}+(1+\mu)}$$

$$=\frac{40}{10000\,(10\,04+j\,9\cdot 4)\times 10^{-4}+41}$$

$$= \frac{40}{51{\cdot}04 + j9{\cdot}4}$$

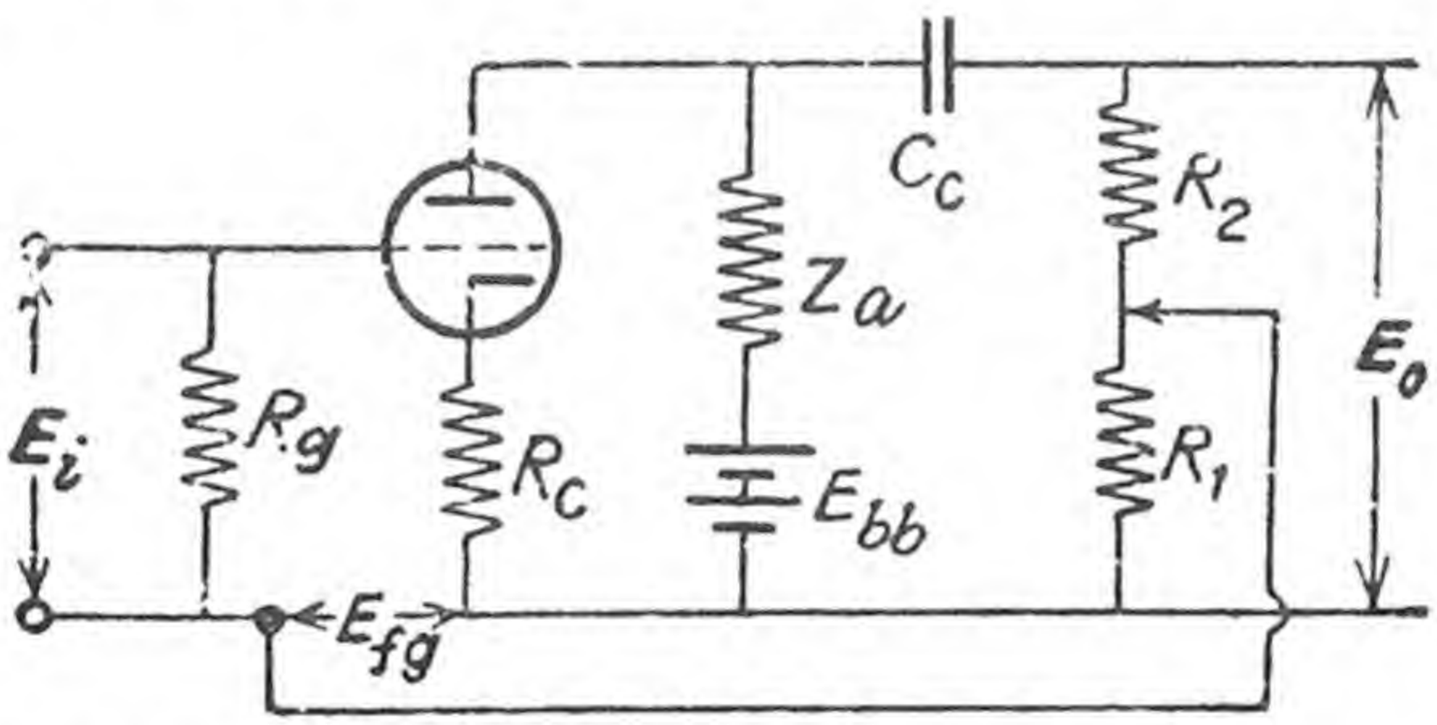
or $$|\mathbf{A}_{hf}| = \frac{40}{\sqrt{\{(51{\cdot}04)^2 + (9{\cdot}4)^2\}}} = \frac{40}{51{\cdot}8}$$

$$= \mathbf{0{\cdot}77}$$

13. *Following particulars relate to the composite feed back amplifier shown in the following diagram ;* $\mu = 50$, $r_p = 20$ *K*Ω. $Z_a = 50$ *K*Ω, $R_1 = 1$ *K*Ω *and* $R_2 = 49$ *K*Ω.

Reactance of coupling condenser C_c *is negligble at operating frequency. Calculate the voltage gain and effective internal impedance of the amplifier for the following conditions :*

(a) $R_c = 0$, *no voltage feed back, (b)* $R_c = 0$ *with voltage feed back, (c)* $R_c = 1000$ *ohms, no voltage feed back, (d)* $R_c = 1000$ *ohms, with voltage feed back.*



(a) $R_c = 0$; No voltage feed back.

The output load impedance would be a parallel combination Z_a and $(R_1 + R_2)$.

$$\therefore \mathbf{Z_L} = \frac{Z_a (R_1 + R_2)}{Z_a + (R_1 + R_2)} = \frac{50 \times 10^3 \times (1 + 49) \times 10^3}{50 \times 10^3 + 50 \times 10^3}$$

$$= \frac{2500 \times 10^6}{100 \times 10^3} = 25 \times 10^3.$$

$$\mathbf{A} = \frac{-\mu \mathbf{Z_L}}{r_p + \mathbf{Z_L}} = -\frac{50 \times 25 \times 10^3}{20 \times 10^3 + 25 \times 10^3} = -22{\cdot}22.$$

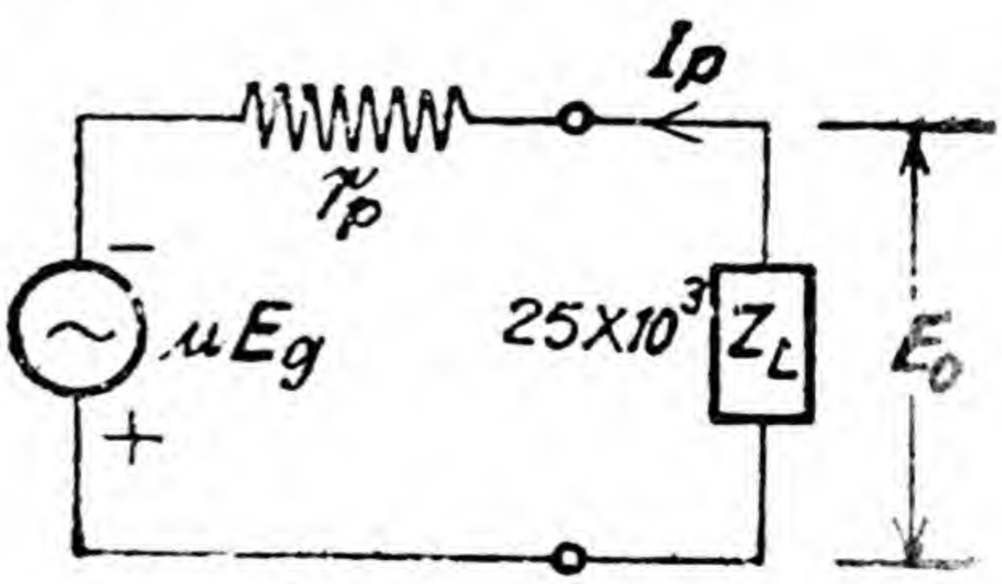
Internal impedance $= r_p = 20$ *K*Ω.

Output impedance would be a parallel combination of r_p and Z_L.

(b) $R_c = 0$; voltage feed back is used.

The a.c. equivalent circuit in this case is given below :

$$\alpha = \frac{1000}{50000} = {\cdot}02.$$



From the figure

$$\mathbf{I}_p (r_p + \mathbf{Z_L}) = \mu \mathbf{E}_g = \mu [\mathbf{E}_i + \alpha \mathbf{E}_0]$$

$$= \mu [\mathbf{E}_i + \alpha (-\mathbf{I}_p . \mathbf{Z_L})]$$

or $$\mathbf{I}_p = \frac{\mu \mathbf{E}_i}{r_p + \mathbf{Z_L} (1 + \mu\alpha)}$$

$$\text{Voltage gain } A_{fb}=\frac{-I_p Z_L}{E_i}=\frac{-\mu Z_L}{r_p+Z_L(1+\mu\alpha)}$$

$$=\frac{-50\times25\times10^3}{20\times10^3+25\times10^3(1+50\times\cdot02)}$$

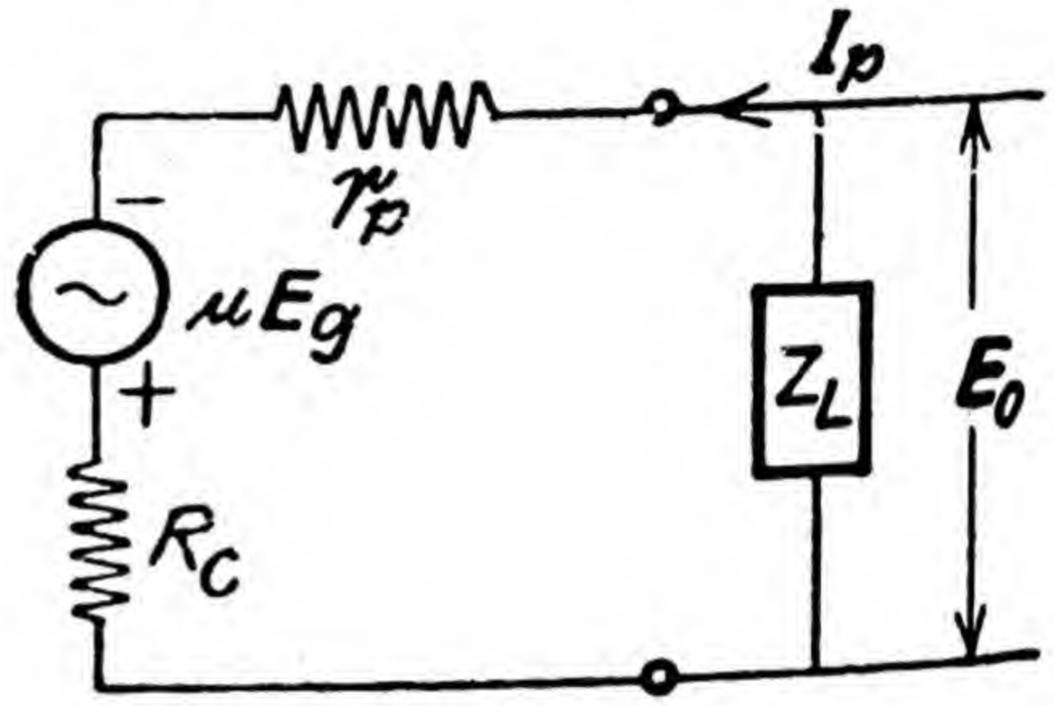
$$=-28\cdot83$$

$$\text{Effective internal impedance}=\frac{r_p}{1+\mu\alpha}=\frac{20\times10^3}{1+50\times\cdot02}=10\ K\Omega.$$

Output terminal impedance would be a parallel combination of effective internal impedance and Z_L.

(c) $R_c=1000\Omega$, no voltage feed back.

The a.c. equivalent circuit is shown in figure.



The current I_p flowing in the circuit is given by

$$I_p\ (R_c+r_p+Z_L)=\mu E_g$$

$$=\mu\ (E_i-I_p R_c)$$

$$\text{or}\quad I_p=\frac{\mu E_i}{r_p+R_c\ (1+\mu)+Z_L}$$

$$\therefore\quad A_{fb}=\frac{-I_p\ Z_L}{E_i}=\frac{-\mu Z_L}{r_p+R_c\ (1+\mu)+Z_L}$$

$$=\frac{-50\times25\times10^3}{20\times10^3+1000\ (1+10)+25\times10^3}$$

$$=-20\cdot68.$$

$$\text{Effective internal impedance}=r_p+R_c\ (1+\mu)$$

$$=20\times10^3+1000\ (51)$$

$$=71\times10^3 \text{ ohms.}$$

Output impedance would be parallel combination of effective internal impedance and load impedance Z_L.

(d) $R_c=1000$ ohms ; voltage feed back is used.

In this case, we have

$$I_p\ (R_c+r_p+Z_L)=\mu E_g$$

$$=\mu\ [E_i+\alpha E_g-I_p R_c]$$

$$=\mu\ [E_i-\alpha I_p Z_L-I_p R_c].$$

$$\therefore\quad I_p=\frac{\mu E_i}{r_p+R_c\ (1+\mu)+Z_L\ (1+\alpha\mu)}$$

$$=\frac{\frac{\mu}{1+\alpha\mu}E_i}{r_p+\frac{R_c\ (1+\mu)}{\alpha\mu+1}+Z_L.}$$

$$\therefore\quad \text{Internal impedance}=\frac{r_p+R_c\ (1+\mu)}{1+\alpha\mu}$$

$$=\frac{20\times10^3+1000\ (1+50)}{(1+\cdot02\times50)}$$

$$=45\cdot5K\Omega.$$

$$\mathbf{A}_{fb}=\frac{-\mu\mathbf{Z}_L}{r_p+R_o(1+\mu)+\mathbf{Z}_L(1+\alpha\mu)}$$

$$=\frac{-50\times25\times10^3}{20\times10^3+1000\ (1+50)+25\times10^3\ (1+\cdot02\times50)}$$

$$=-13\cdot9.$$

14. *A grounded grid amplifier uses a triode having* $\mu=20$, $r_p=$ 10,000 *ohms and load resistance* 12000 *ohms. The impedance of the input voltage source is* 1000 *ohms. Calculate the voltage gain and output terminal impedance.*

$$A_f=\frac{(1+\mu)\ R_L}{r_p+R_L+\mathbf{Z}_i\ (1+\mu)}$$

$$=\frac{(1+20)\ 12\times10^3}{10^4+12\times10^3+10^3\ (1+20)}$$

$$=5\cdot86.$$

$$\text{Input impedance } \mathbf{Z}_{in}=\frac{R_L+r_p+\mathbf{Z}_i\ (1+\mu)}{1+\mu}$$

$$=\frac{12\times10^3+10^4+10^3\times21}{21}$$

$$=2\cdot04\times10^3 \text{ ohms}$$

$$\text{Output impedance}=\frac{\mathbf{Z}_{in}\times R_L}{\mathbf{Z}_{in}+R_L}$$

$$=\frac{2\times10^3\times12\times10^3}{2\times10^3+12\times10^3}=\mathbf{1\cdot71\times10^3\Omega}$$

EXERCISES AND PROBLEMS

1. Enumerate the various types of feed back used in amplifiers. What are the advantages of using negative feed back ?
2. Discuss the advantages and disadvantages of applying negative feed back to an amplifier.
3. Obtain an expression for the gain of an amplifier using negative feed back. Explain why the application of the negative feed back to an amplifier improves its frequency response curve and reduces the distortion in the output.
4. Describe the effect of negative feed back upon the stability of an amplifier.
5. What do you mean by :
   (a) current feed-back,
   (b) voltage feed-back,
   (c) compound feed-back ?

Derive the expressions for the gain in all the three cases.

6. Show that the cathode follower is a negative feed back amplifier and prove that its voltage gain cannot exceed unity.
What are the main advantages of cathode follower ?
7. Derive the expression for the gain of a RC cathode follower for
(a) low frequency range,
(b) mid frequency range, and
(c) high frequency range.
8. What are the special characteristics of the cathode follower circuit ? Derive an expression for the voltage gain of a triode used as a cathode follower.
9. Draw a circuit diagram of cathode follower using a triode valve. Obtain expression for input and output impedances.
10. Describe grounded grid or cathode input amplifier. How it reduces the tendency of oscillations ? Show that it is an ideal amplifier in U.H.F. circuits.
11. An amplifier has gain of 100 $\underline{|180}$. When a negative feed back is applied the gain is reduced to 25 $\underline{|180}$. Determine the feed back network.
12. There are two stages in an amplifier. Each has a gain of 50. Negative feed back is applied by feeding 1/100 of the output voltage to the input. Determine the overall voltage gain of the amplifier with feed back.
13. An amplifier has voltage gain of 5000 with normal plate supply voltage and a voltage gain of 3000 with reduced plate supply voltage. A negative feed back ratio $\beta = -0{\cdot}005$ is applied to the amplifier. Calculate the voltage gain with feed back under the two conditions of plate supply voltage and find the percentage reduction in voltage gain caused by reduction in plate supply voltage under no feed back and under feed back conditions.
14. A cathode follower circuit employs a pentode having $r_p = 250K$ ohms., $g_m = 4mA/V$, $C_{gk} = 5pf$ and its screen grid is decoupled to the cathode by a capacitor of negligible reactance. An impedance $Z = (400 + j300)$ ohms is placed in cathode lead. Obtain expressions for, and hence evaluate (*a*) stage gain, (*b*) output impedance and (*c*) the input impdeance at 4 mc/s. All other capacities may be neglected.
17. A triode used in a cathode follower stage has an anode slope resistance of $8K$ ohm, and $\mu = 25$. The load resistance in the cathode circuit is $12K$ ohm, and the resistance between grid and cathode is $2Meg$ ohm. Deduce expressions for and hence evaluate the input and output resistances of the stage.
16. In a current feed back amplifier, the amplifier factor $\mu = 40$, mutual conductance $g_m = 2mA/$volt, load resistance $R_a = 40 \times 10^3$ ohms, grid leak resistance $R_g = 2Meg$ ohm., cathode resistance $R_k = 1000$ ohm. Calculate :
(*a*) voltage gain of the amplifier.
(*b*) internal impedance.
(*c*) output impedance.
Assume reactance of condenser to be negligible.

17. Following values relate to the single stage RC coupled amplifier with current inverse feed-back.

$R_L = 20K$ ohm. $R_g = 1Meg$ ohm, $E_{bb} = 250$ volts, $\mu = 20$, $g_m = 2000$ μ mhos. Reactance of coupling condenser may be neglected. Shunt capacitance may also be neglected. Calculate the value of cathode resistance to produce voltage gain of —6.

A grounded grid amplifier uses a triode having $\mu = 20$ and $r_p = 8900$ ohms and has load resistance of 13,000 ohms. The impedance of the input voltage source is 1000 ohms. Calculate the voltage gain and output terminal impedance.

12 AUDIO POWER AMPLIFIERS

The primary function of voltage amplifier is to increase the signal voltage from a low level to a reasonably high level without distortion *i.e.*. keeping the signal waveform unaltered. To achieve this object, load resistance is kept high and grid current is avoided. If the grid current is allowed to flow during a part of the signal, the grid to cathode circuit would offer non-linear resistance. This will produce a distortion. Thus we can say that when an a.f. amplifier works under class *A* conditions, it works as a very good voltage amplifier and amplitude distortion is very much reduced.

On the other hand, the function of power amplifier is to increase the signal power from a low level to such a high level as may be required by the load with very small grid driving power demand. To get the high power output, the tube operates beyond the linear range of its characteristics *i.e.* the operating point is allowed to make excursion beyond the linear range of the tube characteristics. By doing so, the ratio of average value of plate current I_b to the a.c. plate current $I_{\sim}$ is reduced. Its effect is to convert a large fraction of d.c. power from the supply into a.c. output. Thus the plate efficiency is increased and the wastage of power is decreased. The increase of plate efficiency is at the cost of distortion. The harmonic components increase in magnitude as the signal amplitude is increased to larger and larger values. It is essential to reduce distortion in power amplifiers. Two techniques are used (*i*) push-pull operation and (*ii*) tuned circuit as the plate load. Power amplifiers may be operated under class *A*, class *B* and class *C* or any point between these limits. If the operation is more or less linear, the equivalent circuit method is applied otherwise the graphical method gives accurate results.

The plate circuit efficiencies under classes *A*, *B* and *C* operations are in increasing order, but at the same time harmonic distortion increases in similar order. Radio frequency power amplifiers can conveniently utilise *B* or *C* operation with high plate circuit efficiency because the tuned circuit eliminates all distortions. In audio frequency power amplifiers class *A* and *B* operations alone can be used. To reduce distortion, push-pull operation is always used in class *B* audio frequency power amplifiers. Class *C* audio frequency operation introduces such a large distortion that even push pull operation cannot handle it.

12·1. CLASS A_1 UNTUNED POWER AMPLIFIER :

These amplifiers are known as low frequency power amplifiers. The circuit of these amplifiers are the same as that of low frequency voltage amplifiers except for the output circuit, which may be series fed or shunt fed. In series fed circuit both a.c. and d.c. of the valve flow through the same load while in shunt fed the a.c. flows through the actual load and d.c. flows through the subsidiary circuit. Here we shall discuss series fed A_1 power amplifiers.

Series fed Power amplifier :

Figure 1 (*a*) shows the basic circuit of an untuned class A_1 series fed power amplifier using a triode valve. Its equivalent circuit is shown in figure 1 (*b*).

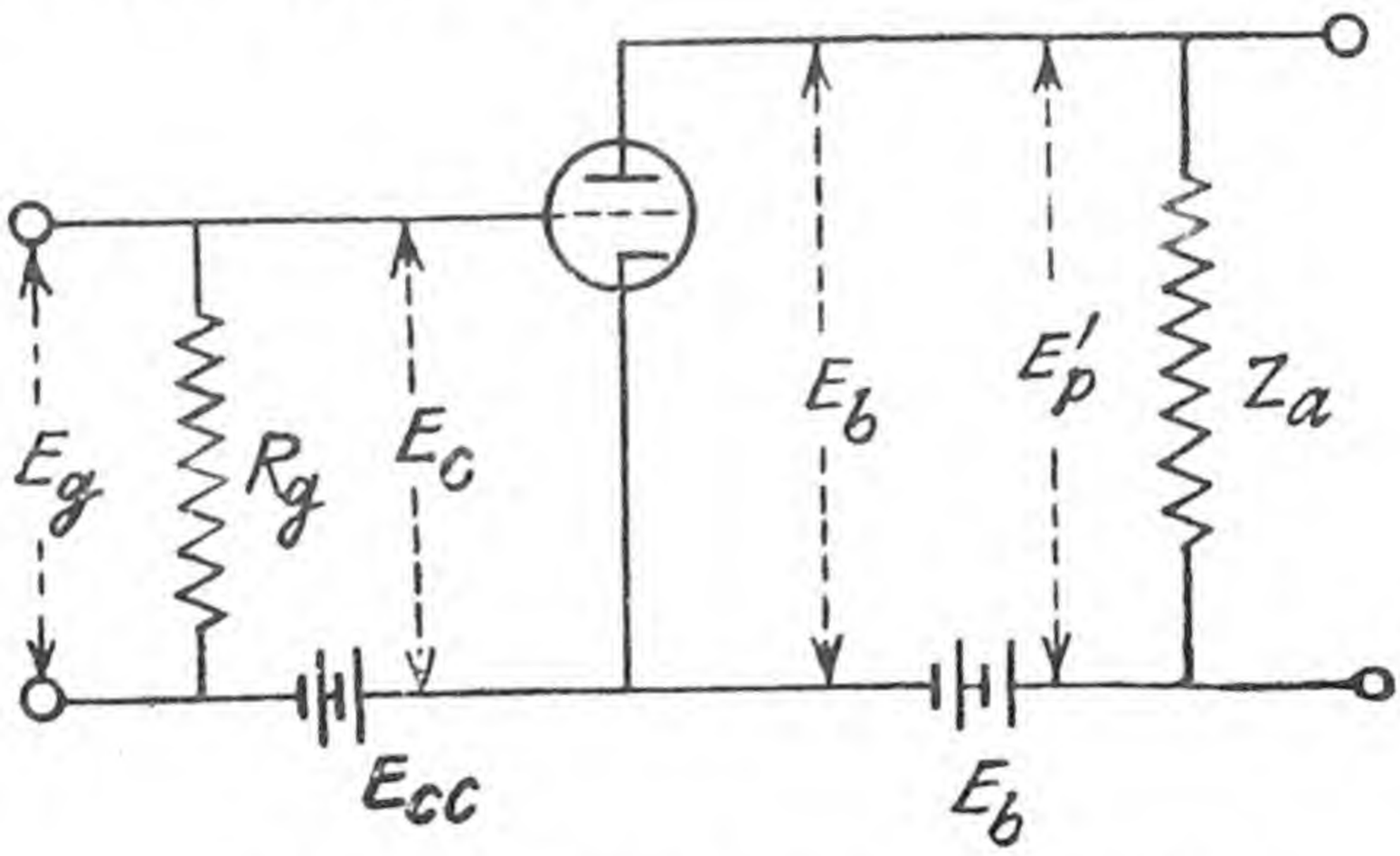


Fig. 1 (*a*) Triode as a power ampiifier.

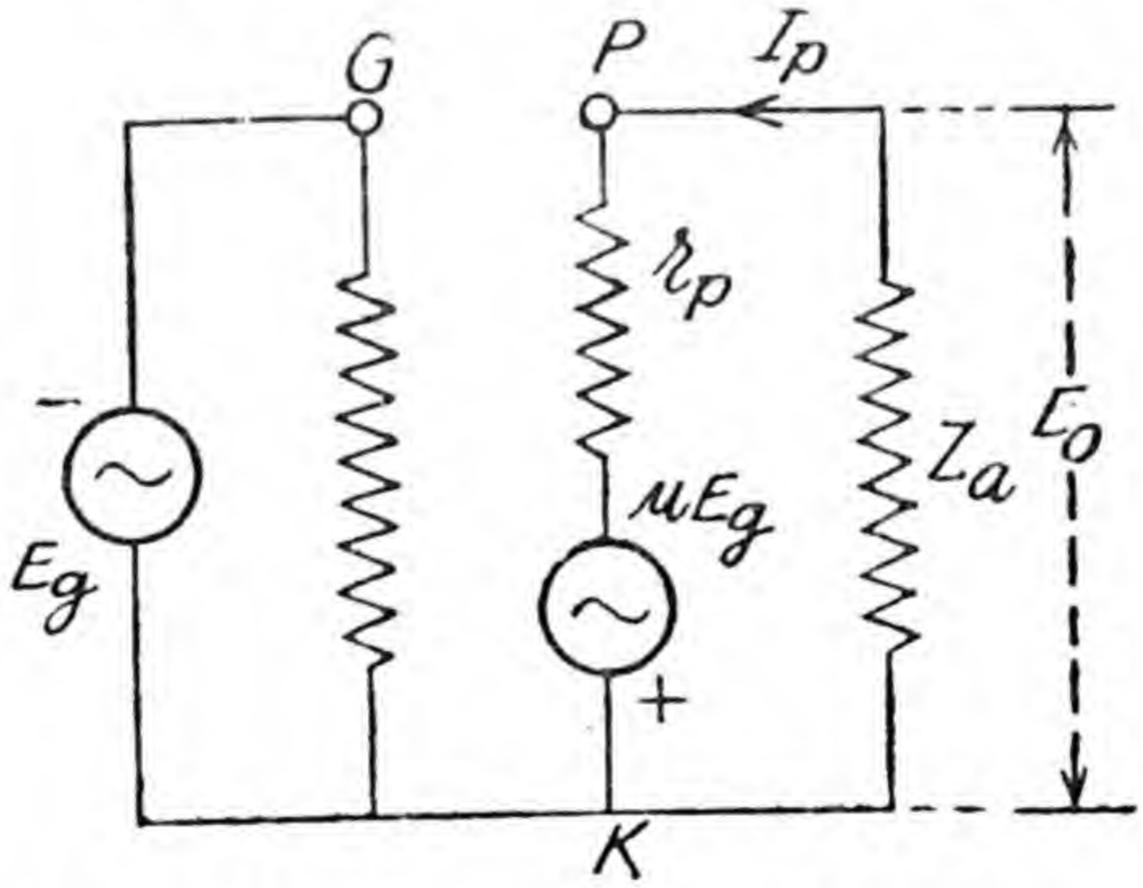


Fig. 1 (*b*). A.c. equivalent circuit.

We know that in any amplifier, the power is fed to the vacuum tube from the three sources; (*i*) from grid circuit, (*ii*) from plate circuit and (*iii*) from cathode heating circuit. Under class A_1 operation, the power from grid circuit is negligible. The grid voltage controls the plate current or we can say it controls the a.c. power output. When grid singal voltage increases, the amplitude of a.c. plate current increases and thus the a.c. power output increases. The power supply from the plate power supply is important and

constitutes the major part which require careful study. Here we assume that : (*a*) the input signal is purely sinusoidal with no d.c. component and (*b*) the load is purely resistive which is R_L which means that load line is straight line rather than eliptical as obtained with reactive load. The load line over plate characteristics is shown in fig. 2. For the complete analysis we shall consider the circuit under 'zero signal' and 'with signal' conditions.

Under zero signal (or quiescent) condition :

Let I_q and E_q be zero signal current and plate voltage respectively. The total power supplied by the plate power supply is given by

$$P_{bb}=E_{bb}\, I_q. \qquad ...(1)$$

Total total power P_{bb} is used in two parts : a part is used as plate dissipation *i.e.* a part is used by the tube which appears in the form of heat at plate and the rest in the plate load. The plate dissipation is equal to the product of the voltage across the tube and current flowing through it *i.e.*

$$P_{d0}=E_q.I_q, \qquad ...(2)$$

where P_{d0} is plate disspated under zero signal condition.

The power absorbed in plate load is given by

$$P_{d.c.}=I_q^2\, R_L. \qquad ...(3)$$

For series fed amplifier

$$E_{bb}=E_q+I_q\, R_L$$

or $$I_qE_{bb}=I_q\, E_q+I_q^2\, R_L$$

or $$P_{bb}=P_{d.0.}+P_{d.c.} \qquad ..(4)$$

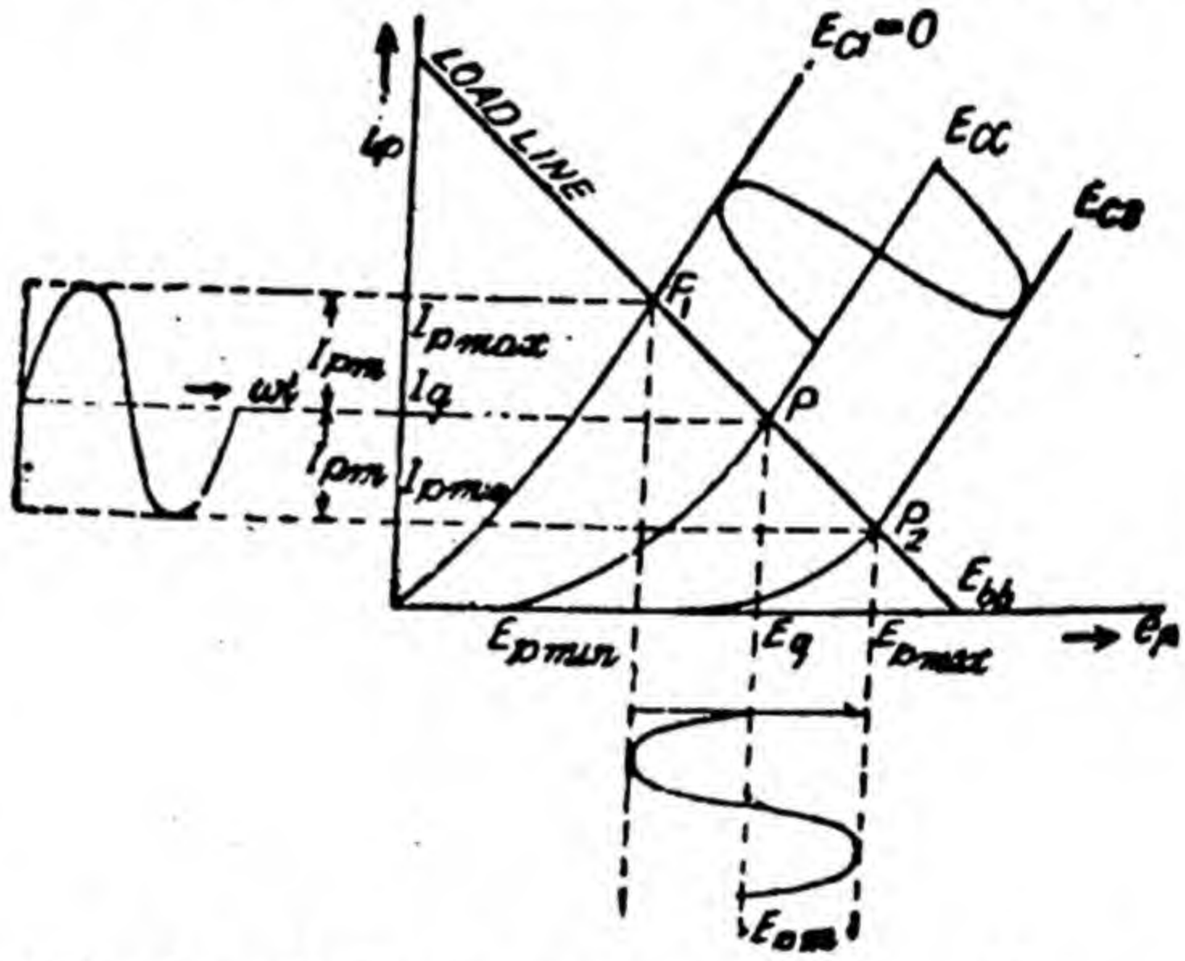


Fig. 2. Output current and voltage waveforms in a series fed A_1 triode power amplifier.

Operation with signal:

Let us consider that a grid signal is applied, then the plate current has d.c. component as well as a c. component. Let the grid signal be a cosine wave represented by

$$e_g=E_{gm}\cos\,\omega t. \qquad ...(5)$$

The average power input from the plate power supply E_{bb} is

$$P_{bs}=\frac{1}{2\pi}\int_0^{2\pi} E_{bb}\, i_b\, d\,(\omega t)$$

$$=\frac{1}{2\pi}\int_0^{2\pi} E_{bb}\,(I_q+i_p)\; d\,(\omega t). \qquad ...(6)$$

The a.c component of the plate current is usually not a pure sine wave for the sinusoidal input because of the non-linearity in the characteristics and it can be expressed as

$$i_p=I_{p0}+I_{p1m}\cos\,\omega t+I_{p2m}\cos\,2\omega t+\ldots$$

As the operation is linear A_1, so the a c. component of the output current would be pure cosine wave and hence

$$I_{p0}=I_{p2m}=I_{p3m}=0$$

$$\therefore \quad P_{bs}=E_{bb}\, i_q+\frac{1}{2\pi}\int_0^{2\pi} E_{bb}\,(I_{p1m}\cos\omega t)\, d\,(\omega t)$$

$$=E_{bb}\, I_q+E_{bb}\, I_{p1m}.\ \frac{1}{2\pi}\Big[\sin\omega t\ \Big]_0^{2\pi}$$

$$=E_{bb}\, I_q. \qquad \ldots(7)$$

This is same as equation (1). This shows that for a linear class A_1 operation the power input to the circuit from plate supply source is constant whether the signal is present or not.

The power delivered to the load is given by

$$P_L=\frac{1}{2\pi}\int_0^{2\pi} i_b^2\, R_L\, d\,(\omega t)=\frac{R_L}{2\pi}\int_0^{2\pi}(I_q+i_p)^2\, d(\omega t)$$

$$=\frac{R_L}{2\pi}\left[\int_0^{2\pi} I_q^2\, d(\omega t)+\int_0^{2\pi} i_p^2\, d(\omega t)+2\int_0^{2\pi} I_q\, i_p\, d\,(\omega t)\right].$$

But $i_p=I_{p1m}\cos\omega t$ and I_q is constant d.c. current

$$\therefore \quad P_L=R_L\,[I_q^2+I^2{}_{r\ m\cdot s}+I_q.0], \qquad (8)$$

where $I_{r.m.s.}$ is the *r.m.s.* value of the fundamental component of the current.

$$\therefore \quad P_L=P_{d.c.}+P_{a.c.} \text{ (where } P_{a.c.} \text{ is a.c. power output)}$$

$$P_{a.c.}=I^2{}_{p\,m}\,\frac{R_L}{2\pi}\int_0^{2\pi}\left(\frac{\cos\ \omega t+1}{2}\right) d\,(\omega t)$$

$$=\frac{I^2{}_{p1m}\ R_L}{4\pi}\left[\frac{\sin\ \omega t}{2}+\omega t\ \right]_0^{2\pi}=\frac{I^2{}_{p\,m}\ R_L}{2}=I^2{}_{r.m\ s.}\ R_L \ldots(9)$$

Now according to the law of conservation of energy

$$P_{bs}=P_L+P_{d.s.}=P_{d.c.}+P_{a.c.}+P_{d.s.},$$

where P_{bs} is the total power supplied from the source, $P_{d.c.}$ is the plate dissipation in the tube with signal.

From equation (1) and 7) we have $P_{bs}=P_{bb}=P_{d.c.}+P_{d.0}$

$$\therefore \quad P_{d.c.}+P_{d.0}=P_{d.c.}+P_{a.c.}+P_{d.s.}$$

$$P_{d.s.}=P_{a.0.}-P_{a.c.} \qquad \ldots(10)$$

Equation (10) shows that plate dissipitation with signal is less than that without signal. The difference of the two is equal to the a.c. power output across the load. This fact is very helpful in designing the power amplifiers. The amplifiers are designed for maximum plate dissipation under signal condition.

Out of the total power P_{bb} received from plate supply source, the part $P_{a.c.}$ is available as the useful a.c. output power. The percentage plate circuit efficiency is given by

$$\eta_p=\frac{P_{a\ c.}}{P_{bb}}\times 100=\frac{P_{a.c.}}{E_{bb}\cdot I_q}\times 100 \text{ percent.}$$

The d.c. power in the load does not provide useful power output. In order to increase the plate circuit efficiency I_q should be as small as possible. High value of I_q is responsible for low plate circuit efficiency of class *A* amplifier. To increase it the operations *B* and *C* are used.

12·2. MAXIMUM POWER SENSITIVITY OF SERIES FED A_1 POWER AMPLIFIERS :

It is clear from the previous article that to have a high plate circuit efficiency, the useful a.c. power output must be maximized. The output depends upon (i) tube, (ii) plate supply voltage E_{bb},

(*iii*) magnitude of input signal and (*iv*) load impedance. For the purpose of present analysis we assume that (*i*) the tube, (*ii*) linear class A_1 operation, (*iii*) grid signal voltage and (*iv*) zero signal operating point are prescribed and only variable left is load impedance which is assumed complex for the generality. Hence

$$\mathbf{Z}_a = R_L + jX_L. \qquad \ldots(1)$$

Considering figure 1 (*b*), the a.c. plate current $\mathbf{I}_p$ is given by

$$\mathbf{I}_p = \frac{\mu \mathbf{E}_g}{r_p + \mathbf{Z}_a} = \frac{\mu \mathbf{E}_g}{(r_p + R_L) + jX_L}$$

Magnitude $$I_p = \frac{\mu E_g}{[(r_p + R_L)^2 + X_L^2]^{1/2}} \qquad \ldots(2)$$

The a.c. power output is given by

$$P_{a.c.} = I_p^2 . R_L = \frac{\mu^2 E_g^2 \, R_L}{(r_p + R_L)^2 + X_L^2}. \qquad \ldots(3)$$

The power sensitivity is defined as the ratio of useful power output to the square of grid signal voltage

$$\text{Power sensitivity} = \frac{P_{a.c.}}{E_g^2} = \frac{\mu^2 R_L}{(r_p + R_L)^2 + X_L^2}. \qquad \ldots(4)$$

Equation (3) indicates that for large power output, the input signal must be large and if the signal is weak it should be amplified before feeding to the power amplifier. The power output also depends upon R_L and X_L for a given tube and input signal. To maximize the power output both R_L and X_L may be variable and hence three conditions are possible (i) R_L alone is variable, (ii) X_L alone is variable and (iii) both are variable. If both R_L and X_L vary independently, the condition becomes very complex. Hence we assume that R_L and X_L vary in such a manner that X_L/R_L is constant.

(*i*) *When R_L alone is variable* : In order to get the condition of maximum power output, we differentiate equation (3) with respect to R_L and equate it to zero *i.e.*

$$\frac{dP_{a.c.}}{dR_L} = \mu^2 E_g^2 \left[\frac{(r_p + R_L)^2 + X_L^2 - 2(r_p + R_L) R_L}{\{(r_p + R_L)^2 + X_L^2\}^2} \right] = 0$$

or $$(r_p + R_L)^2 + X_L^2 = 2(r_p + R_L) R_L$$

or $$R_L = \sqrt{(r_p^2 + X_L^2)}. \qquad \ldots(5)$$

(*ii*) *When X_L alone is variable*: It is clear from equation (3) that when X_L alone is variable, $P_{a.c.}$ is maximum when $X_L = 0$.

(*iii*) *When Z_a is variable* : If Z_a is variable such that $X_L/R_L = Q$ = constant, then equation (3) can be written as

$$P_{a.c.} = \frac{\mu^2 E_g^2 R_L}{(r_p + R_L)^2 + (Q R_L)^2}$$

$$\frac{dP_{a.c.}}{dR_L} = \mu^2 E_g^2 \left[\frac{(r_p + R_L)^2 + (QR_L)^2 - R_L\{2(r_p + R_L) + 2Q^2 R_L\}}{\{(r_p + R_L)Q^2 + (R_L)^2\}^2} \right] = 0$$

or $$(r_p+R_L)^2+(Q\ R_L)^2=R_L\{2(r_p+R_L)+2Q^2R_L\}$$

or $$r_p=\sqrt{(R_L^2+X_L^2)}=Z_a. \quad \ldots(6)$$

In most cases Z_a is pure resistance R_L *i e.* $X_L=0$. In that case, the power output is given by

$$P_{a\cdot c\cdot}=\mu^2E_g^2\ \frac{R_L}{(r_p+R_L)^2} \quad \ldots(7)$$

From equation (5), it is maximum when $R_L=r_p$ *i.e.* load resistance is equal to the source resistance r_p. The power output so obtained (when $r_p=R_L$) is the maximum power available from any energy source and is defined as available power (P_{av}).

Hence $$P_{av}=(P_{a\cdot c\cdot})_{max\cdot}=\frac{(\mu E_g)^2r_p}{(2r_p)^2}=\frac{\mu^2E_g^2}{4r_p}=-\frac{(g_mE_p)^2r_p}{4}. \quad \ldots(8)$$

Now maximum power sensitivity is given by

$$\frac{P_{a\cdot c\cdot}}{E_g^2}=\frac{\mu^2}{4r_p} \quad \ldots(9)$$

Again we have

$$\frac{P_{a\cdot c\cdot}}{P_{av}}=\frac{\mu^2E_g^2R_L}{(r_p+R_L)^2}\times\frac{4r_p}{\mu^2E_g^2}=\frac{R_L}{(r_p+R_L)^2}\times 4r_p=\frac{4}{(r_p/R_L-2+R_L/r_p} \quad \ldots(10)$$

Figure 3 shows the variation of ratio $P_{a\cdot c\cdot}/P_{av}$ with ratio R_L/r_p. It is observed from the figure that the curve is symmetrical on either side of value $R_L/r_p=1$. This is due to the symmetry of the denominator of equation 10. It is also observed that the output is maximum when $R_L/r_p=1$.

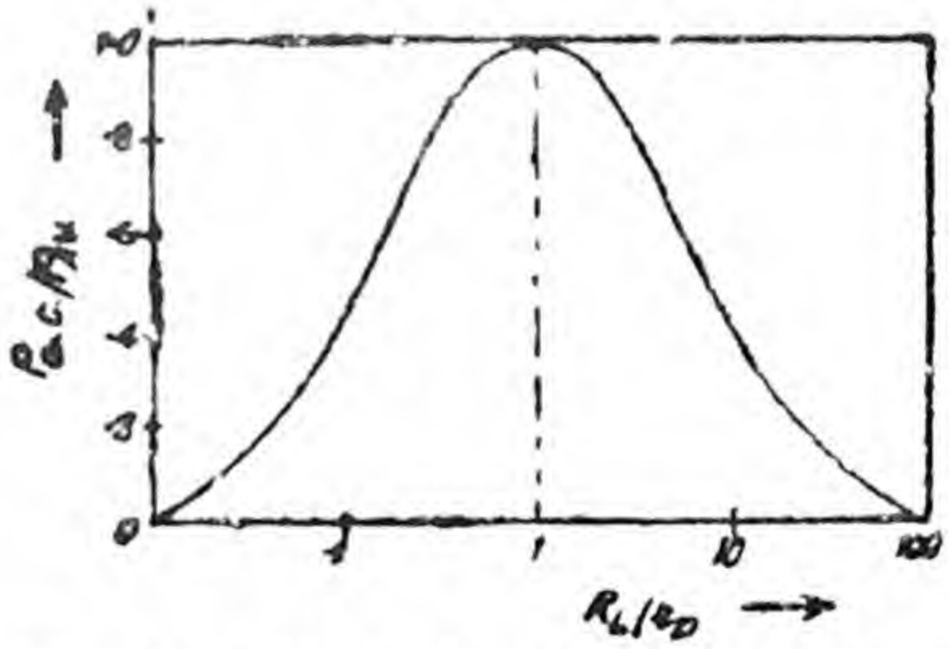


Fig. 3. Variation of a.c. power output as a function of R_L/r_p.

In equation (9), the quantity $(\mu^2/4r_p)$ is known as figure of merit of the tube and depends on the tube construction. Since r_p is required to be equal to load resistance, it is preferred to use a low r_p tube to have a reasonable amount of power output with usual value of plate supply voltage.

12·3. SHUNT-FED AND TRANSFORMER COUPLED LINEAR A_1 POWER AMPLIFIERS :

In series fed A_1 power amplifiers the load is directly connected to the plate but sometimes it is not desirable. One of the reasons being that the considerable amount of quiescent direct current flows through the load resistance causing a large wastage of d.c. power. The d.c. power contributes nothing to the output a.c. power. Secondly, the load resistance of an audio power amplifier

may be the resistance of voice coil of a loudspeaker and it is undesirable to pass a large d.c. through it otherwise it may cause polarisation of the coil. Hence transformer coupled load is generally preferred in audio power amplifiers, whereas on rare occassions, shunt fed (parallel fed system) is used.

Shunt fed circuit : The circuit of a shunt fed system is shown in figure 4. The system is often called as impedance capacitance coupled system. Generally, the inductance possesses a negligible small d.c. resistance.

In this system the plate supply is connected to the plate of the tube through a high inductance L, the load resistance being connected across the output through a capacitor C. The inductance L must be so chosen that $\omega L \gg R_L$ and C must be chosen that $1/\omega C \ll R_L$ over the operating range of frequencies.

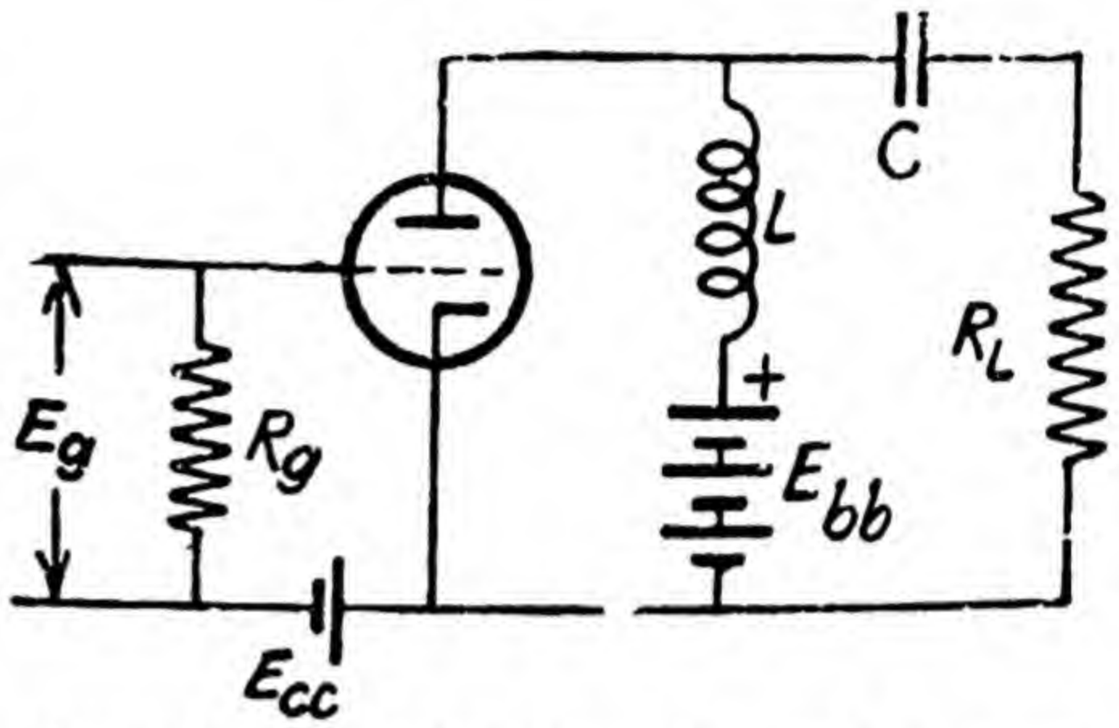


Fig 4. A parallel or shunt fed power amplifier.

Efficiency of shunt fed power amplifier : Figure 5 shows the d.c and a.c. load lines for this amplifier.

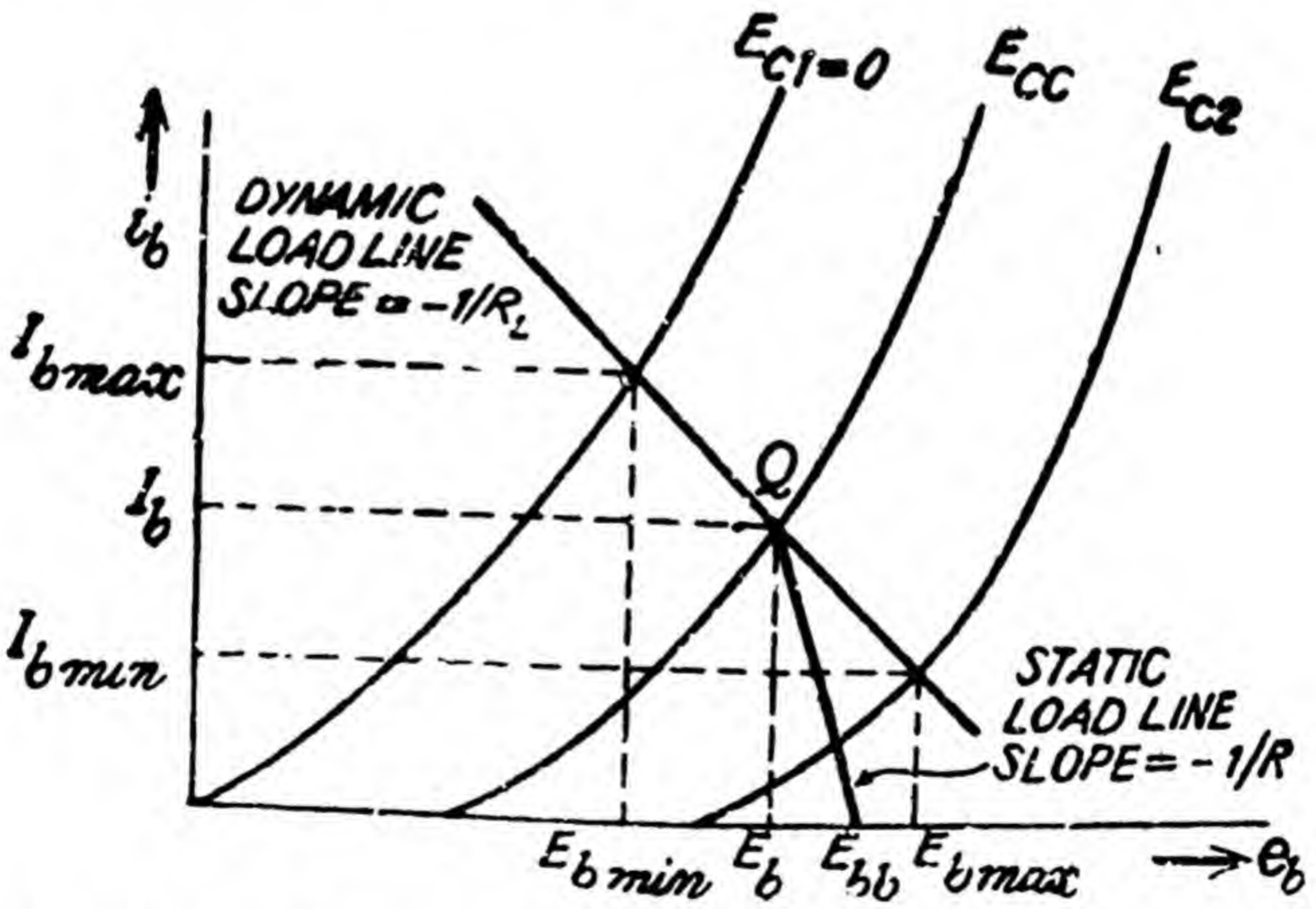


Fig. 5. Dynamic and static load lines of shunt fed amplifier.

First of all the operating (quiescent) point Q is obtained from the intersection of the d.c. load line and the characteristic for E_{cc} In this case the plate current through the tube is determined by the small resistance R of the inductor L. As the d.c. resistance of the inductor coil L is negligibly small, for simplicity it is assumed that $R_{d.c.}$ is equal to zero. The load line will be perfectly vertical and $E_b = E_{bb}$. Now the next step is to draw the a.c. line passing through Q and having a slope corresponding the a.c. impedance of the circuit. The a.c. impedance is equal to R_L since reactance

of inductor L at input frequency is large and may be neglected. The reactance of C in comparison to R_L is very small and is neglected.

$$P_{a.c.}=\frac{(\mathrm{I}_{b\ max}-\mathrm{I}_{b\ min})\ (\mathrm{E}_{b\ max}-\mathrm{E}_{b\ min})}{8} \qquad \ldots(1)$$

The total power drawn from the supply voltage E_{bb} is

$$P_{bb}=E_{bb}\times \mathrm{I}_b$$
$$=E_{bb}\left[\frac{\mathrm{I}_{b\ max}+\mathrm{I}_{b\ min}}{2}\right]$$
$$=\mathrm{E}_b\left[\frac{\mathrm{I}_{b\ max}+\mathrm{I}_{b\ min}}{2}\right] \qquad \text{since } R_{d.c.}=0 \qquad \ldots(2)$$

The efficiency η is given by

$$\eta=\frac{(\mathrm{I}_{b\ max}-\mathrm{I}_{b\ min})\ (\mathrm{E}_{b\ max}-\mathrm{E}_{b\ min})/8}{\mathrm{E}_b\ (\mathrm{I}_{b\ max}+\mathrm{I}_{b\ min})/2}$$
$$=\frac{(\mathrm{I}_{b\ max}-\mathrm{I}_{b\ min})(\mathrm{E}_{b\ max}-\mathrm{E}_{b\ min})}{4\ \mathrm{E}_b\ (\mathrm{I}_{b\ max}+\mathrm{I}_{b\ min} \qquad \ldots(3)$$

When large signal is applied, $\mathrm{I}_{b\ max}=2\ \mathrm{I}_b$ and $\mathrm{E}_{b\ max}=2E_b$, $\mathrm{I}_{b\ min}=0$ and $\mathrm{E}_{b\ min}=0$. Hence maximum power efficiency is given by

$$\eta=\frac{(2\mathrm{I}_b-0)(2\mathrm{E}_b-0)}{4\mathrm{E}_b\ (2\mathrm{I}_b)}\times 100\%.$$
$$=50\%. \qquad \ldots(4)$$

Thus using an inductively load a better efficiency may be achieved.

Transformer coupled amplifier :

The transformer coupling is useful when very low impedance such as the voice coil of loudspeaker having resistance of a few ohms is to be fed from a power amplifier. The transformer used as a coupling device is commonly a step down transformer. The circuit arrangement of a transformer coup'ed amplifier is shown in figure 6.

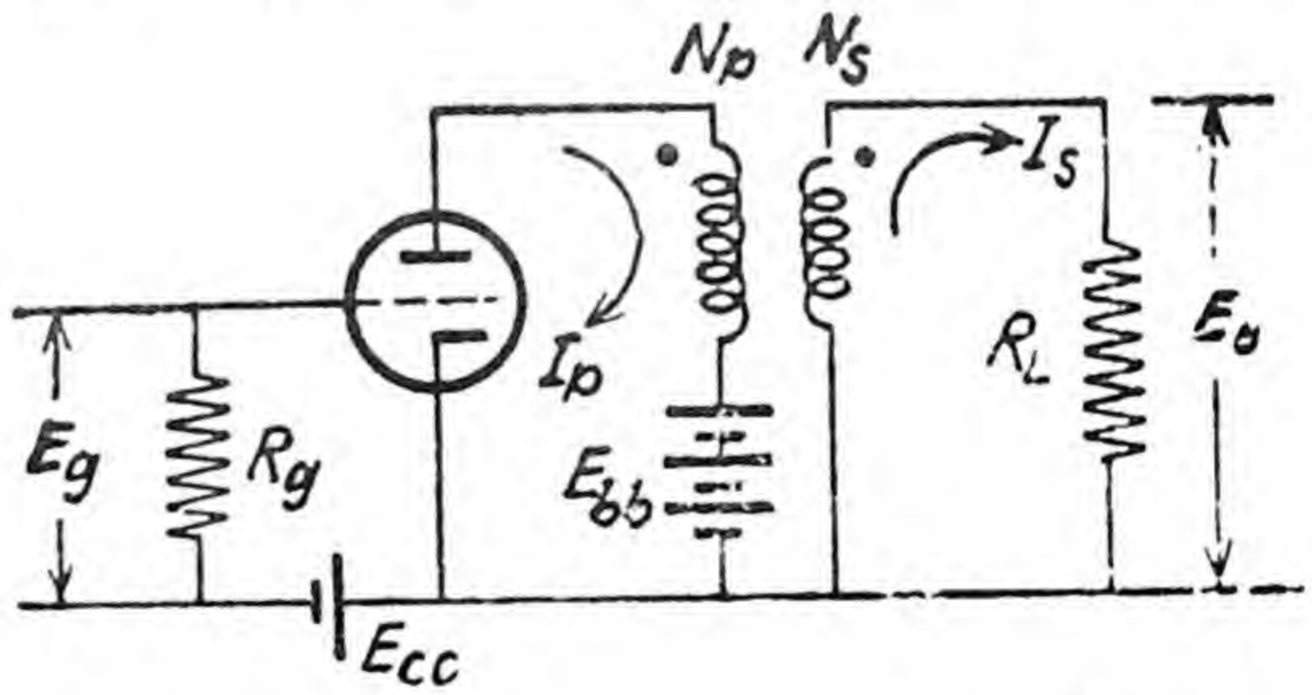


Fig. 6. Transformer coupled amplifier.

From ideal transformer, we have

$$E_p=\frac{N_p}{N_s}E_0,$$

and

$$\mathrm{I}_p=\frac{N_s}{N_p}\mathrm{I}_s,\ ,$$

where E_x and E_0 are the primary and secondary potentials respec-

tively, and $\mathbf{I}_p$ and $\mathbf{I}_s$ are the primary and secondary currents. Dividing the primary voltage $\mathbf{E}_p$ by primary current $\mathbf{I}_p$, we get

$$\frac{\mathbf{E}_p}{\mathbf{I}_p}=\left(\frac{N_p}{N_s}\right)^2\frac{\mathbf{E}_0}{\mathbf{I}_s}. \qquad ...(5)$$

Denoting $\mathbf{E}_p/\mathbf{I}_p$ by R_L' and $\mathbf{E}_0/\mathbf{I}_s$ by R_L, equation (5) takes the form

$$R_L'=\left(\frac{N_p}{N_s}\right)^2 R_L, \qquad ...(6)$$

where R_L' and R_L represent effective input and output impedance of ideal transformer. Equation (6) suggests that a certain amount of impedance is reflected from the secondary to primary which is given by

$$R_{ref}=R_L'=\left(\frac{N_p}{N_s}\right)^2 R_L. \qquad ...(7)$$

The reflected impedance depends upon the turns ratio and impedance of secondary.

Equation (7) is true for an ideal transformer. In general, however, the secondary and primary resistances can not be neglected and due to imperfect coupling, the leakage inductances should also be considered. By taking these factors into account, the input impedance is given by

$$Z_p=R_L\left(\frac{N_p}{N_s}\right)^2+R_p+R_s\left(\frac{N_p}{N_s}\right)+j\omega\left[L_p+L_s\left(\frac{N_p}{N_s}\right)^2\right], \qquad ...(8)$$

where R_p and R_s are the d.c. resistances of the primary and secondary respectively, and L_p and L_s are leakage inductances of primary and secondary respectively.

In transformer coupled amplifier the d.c. load line is drawn for R_p and a.c. load line is drawn for $(R_L+R_s))(N_p/N_s)^2$, neglecting the leakage inductances. Fig. 5 also applies for this coupling.

12·4. MAXIMUM UNDISTORTED POWER :

In the region of small plate current, considerable curvature of the plate characteristic exists which causes a distortion. Hence in distortion free class *A* operation, in order to obtain the maximum possible output, without making the instantaneous plate current too small during the most negative part of applied signal, and without driving the grid positive at the positive peaks of the applied signal, it is necessary to maintain a careful balance among the grid bias, load impedance plate supply potential and plate resistance.

Figure 7 shows the graphical construction for the determination of operating conditions for maximum undistorted output power. $I_{p\ min}$ indicates the minimum plate current below which operation is not to be allowed in order to avoid distortion. Secondly point of operation may be allowed between zero grid bias and a negative grid bias which yields plate current equal to

$I_{P\ min}$. Thirdly, we have assumed that above the horizontal line corresponding to plate current of $I_{P\ min}$. plate characteristics are almost linear with a slope corresponding to Z_a and hence amplification is distortion free, *i.e.* $Z_a = \tan\theta$.

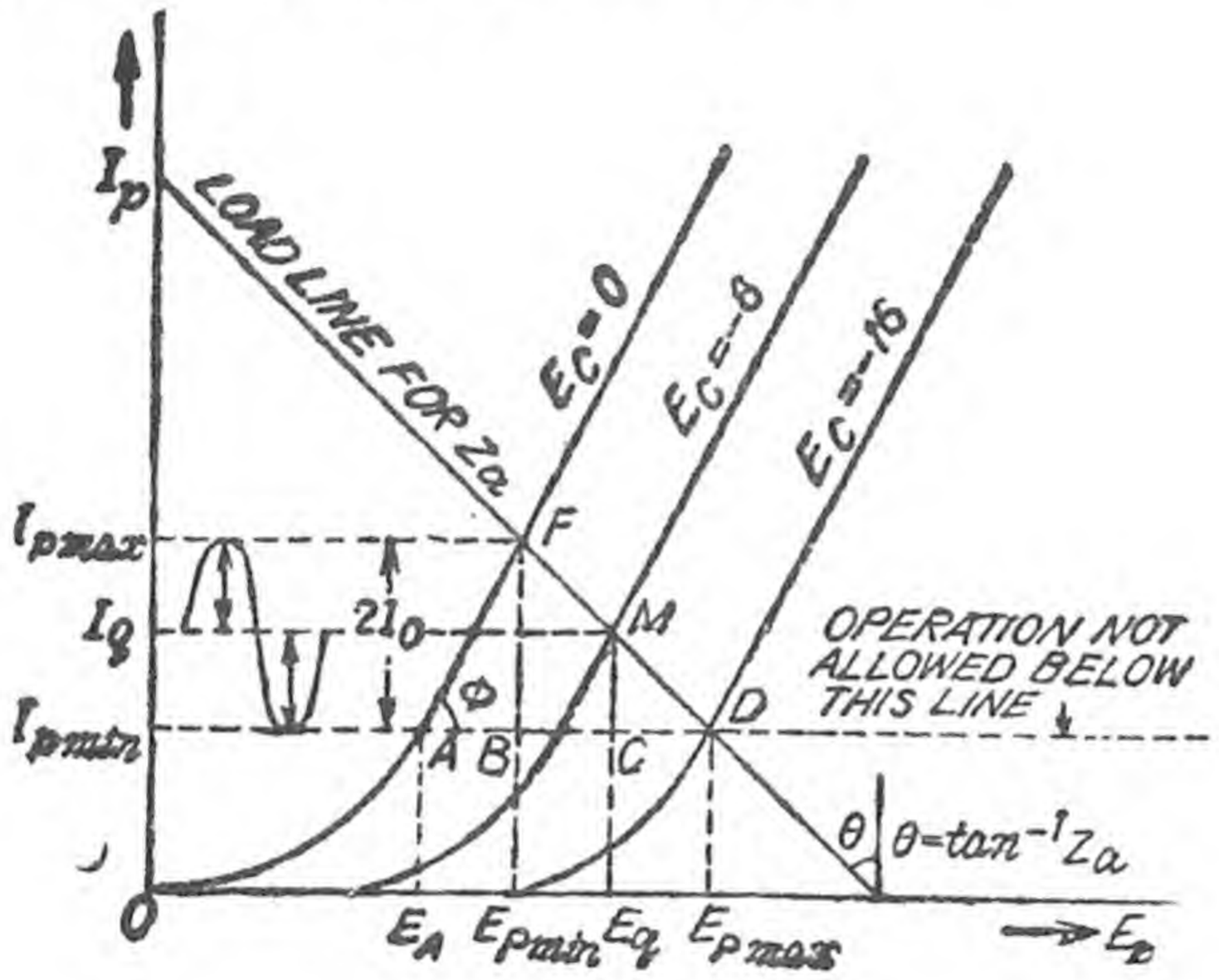


Fig. 7. Output current and voltage waveforms in triode amplifier.

From the fig. 7, ϕ is the angle which the straight portion of the characteristic makes with the voltage axis, *i.e.*,

$$\tan\phi = \frac{2I_0}{AB}$$

and $$\frac{\text{change in } E_p}{\text{change in } I_p} = r_p = \frac{AB}{2I_0}.$$

So that $AB = 2I_0 r_p$.

Again $E_q = E_A + 2I_0 r_p + E_0$

but since $E_A = Z_a I_0$,

we find $E_q = E_A + 2I_0 r_p + Z_a I_0$

$$= E_A + I_0 (Z_a + 2r_p$$

Solving this expression for I_0,

$$I_0 = \frac{E_q - E_A}{Z_a + 2r_p}. \qquad \ldots(1)$$

The a.c. power to the load is given by

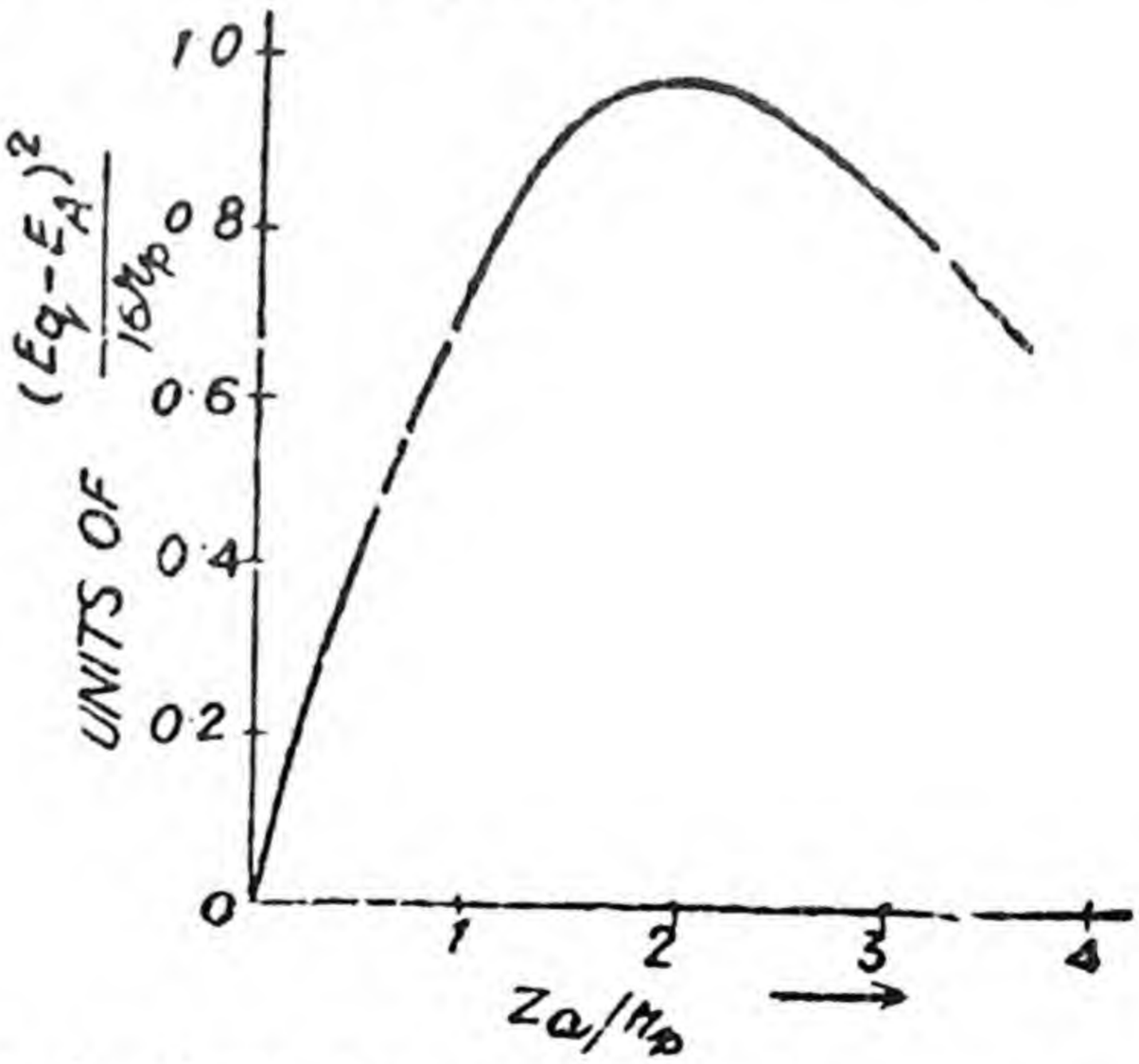


Fig. 8. Output current and voltage waveforms in triode power amplifier.

$$P_{a\cdot c.} = \frac{I_0^2 Z_a}{2} = \frac{(E_q - E_A)^2}{16 r_p} \frac{4Z_a/2r_p}{\left(1 + \frac{Z_a}{2r_p}\right)^2}. \qquad \ldots(2)$$

A sketch showing a variation of the power output as a function of Z_a/r_p is shown in figure 8.

In this case output reaches a maximum when

$$\frac{\partial P_{a.c.}}{\partial Z_a}=0.$$

Equation (2) in a simplified form can be written as

$$P_{a.c.}=\tfrac{1}{2}\frac{(E_g-E_A)^2.Z_a}{(Z_a+2r_p)^2}$$

$$\frac{\partial P_{a.c.}}{\partial Z_a}=\tfrac{1}{2}\frac{(E_g-E_A)^2\{(Z_p+2r_p)^2.1-Z_a\,2\,(Z_a+2r_p)\}}{(Z_a+2r_p)^4}=0$$

or $$(Z_a+2r_p)-2Z_a=0$$

or $$Z_p=2r_p. \qquad ...(3)$$

Therefore, in order to obtain the maximum undistorted power output, the load Z_a should be equal t› the twice the plate resistance r_p. Let the maximum value be denoted by $(P_{a.c.})_{max}$, then

$$(P_{a.c.})_{max}=\frac{(E_g-E_A)^2}{16r_p} \qquad ...(4)$$

and $$\frac{P_{a.c.}}{(P_{a.c.})_{max}}=\frac{4\,(Z_a/2r_p)}{(1+Z_a/2r_p)^2}. \qquad ...(5)$$

The optimum grid bias for these conditions may be obtained in the following way. When the point of operation moves from point M to F, the grid voltage cha nges by an amount equal to E_c and the plate current changes by I_0. Hence

$$I_0=\frac{\mu E_c}{Z_a+r_p}. \qquad ...(6)$$

From the equations (1) and (6), we have

$$\frac{(E_g-E_A)}{Z_p+2r_p}=\frac{\mu E_c}{Z_a+r_p}.$$

For load $Z_a=2r_p$, we have

$$(E_g-E_A).3.r_p=\mu.E_c.4.r_p$$

or $$E_c=\frac{3.(E_g-E_A)}{4\mu}. \qquad ...(7)$$

The value of E_A can be obtained directly from plate characteristics. With the help of equation (7), we can obtain the value of $(P_{a.c})_{max.}$ in terms of E_c. Substituting the value of (E_g-E_A) from equation (7) in equation (4), we have

$$(P_{a.c.})_{max.}=\left(\frac{4}{3}\ \mu E_c\ \right)^2\frac{1}{16r_p}$$

$$=\frac{\mu^2E_c^2}{9r_p}. \qquad (8)$$

When the grid is driven to zero volt, the peak value of the sinusoidal input signal is equal to E_c and the r.m.s. value of input signal. E_q is equal to $E_c/\sqrt{2}$. From equation (8),

$$\therefore \quad (P_{a\cdot o\cdot})_{max\cdot} = \tfrac{2}{9}\mu g_m E_g^2 \qquad \ldots(9)$$

$$\text{and power sensitivity} = \frac{(P_{a\cdot o\cdot})_{max\cdot}}{E_q^2} = \frac{2}{9}\,\mu g_{m\cdot} \qquad \ldots(10)$$

12·5. THEORETICAL PLATE CIRCUIT EFFICIENCY OF TRIODE POWER AMPLIFIER :

An approximate expression for theoretical value of plate circuit efficiency η_p may be found by considering the tube to have ideal characteristics, *i.e.*, perfectly linear, parallel and equispaced for equal increment of parameter.

Figure 9 shows the ideal plate characteristic of a triode. Operation is linear class *A* so that grid voltage E_c never exceeds zero and at the same time plate current is never cut-off.

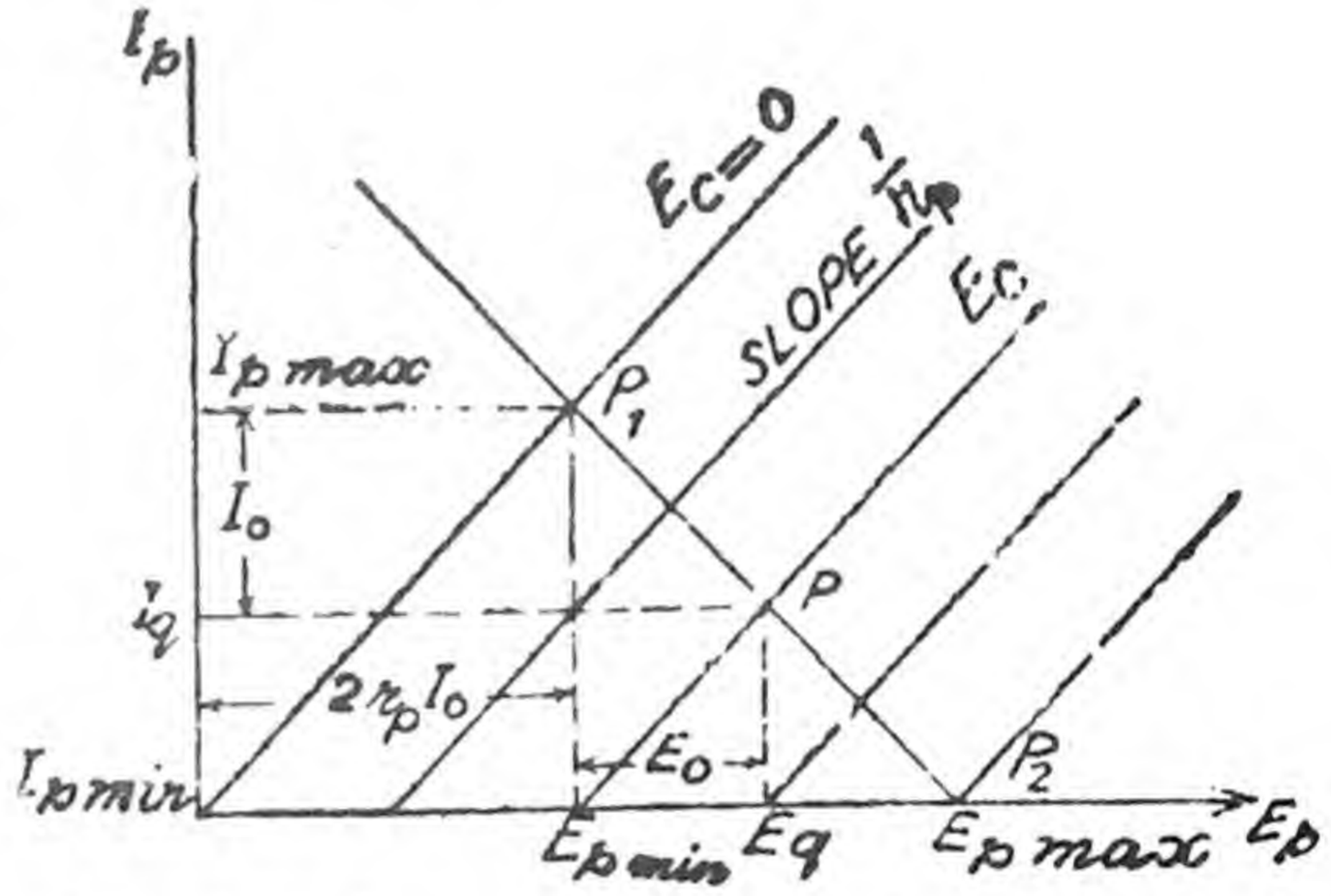


Fig. 9. Plate characteristic of a triode.

The plate circuit efficiency of the amplifier is given by

$$\eta_p = \frac{\text{a c. power output to the load}}{\text{d.c. power input to the plate circuit}} \times 100\%. \qquad \ldots(1)$$

Let E_{bb} and I_q be the battery voltage and plate current respectively when no grid signal is applied. The total power supplied from the plate power supply will be $E_{bb}\,I_q$. Let this be denoted by P_{bb}. A part of this power is absorbed in the plate load given by $P_{d\cdot c\cdot} = I_q^2 Z_a$. The other part is used in increasing the kinetic energy of electrons in passing from cathode to plate and on striking the plate and they dissipate this energy in the form of heat. Heat produced at the plate per second is called dissipation. If P_{p0} denotes the average power dissipated by the plate, then $P_{p0} = E_q . I_q$, where E_q is the plate voltage for zero signal. Hence

$$P_{bb} = P_{d\cdot c\cdot} + P_{p0\cdot} \qquad \ldots(2)$$

When a signal voltage is applied, the average kinetic energy of the electrons reaching the plate is decreased. Thus the application of input signal decreases the plate dissipation. The complete analysis shows that dissipation $P_p = P_{p0} - P_{a\cdot c}$. $\ldots(3)$

Since the power P_{bb} from plate supply source remains the same for linear *A* operation, the function of the tube is simply to remove

an amount of power $P_{a.c.}$ from plate dissipation and make it available as the useful a.c. output power across the load resistance. In this way the heating of the plate is replaced by the amount of the a.c. power that the tube supplies to the load and hence the tube is not so hot when delivering power to a load as when there is no such a.c. power transfer.

From equation (1) the efficiency may be written as

$$\eta_p = \frac{E_p I_p}{E_{bb} I_q} \times 100\%. \qquad ...(4)$$

Suppose the grid does not swing beyond $E_c=0$ corresponding to the point P_1 and may swing to the point P_2 corresponding to zero plate current. This gives maximum excursion of grid voltage and hence gives maximum power output. Then by the upper choice of E_c

$$I_q = I_0$$

and

$$\eta_p = \frac{E_p I_p}{E_{bb} I_0} = \frac{E_0 I_0}{2E_{bb} I_0} = 50 \frac{E_0}{E_{bb}}\%. \qquad ...(5)$$

For series fed circuit, the point marked $(E_p)_{max} = E_{bb}$ and from the figure

$$(E_p)_{max} = E_{bb} = 2E_0 + 2I_0 r_p.$$

Now from equation (5)

$$\eta_p = 50 \times \frac{E_0}{2(E_0 + I_0 r_p)} = \frac{25}{1 + r_p \frac{I_0}{E_0}}\%.$$

Since $E_0 = I_0 Z_a$,

then

$$\eta_p = \frac{25}{1 + r_p / Z_a}\%.$$

The theoretical maximum plate circuit efficiency for the series fed amplifier is 25%.

12·6. CLASS A PUSH-PULL AMPLIFIER :

The fundamental aim of the amplifiers is to obtain an exact and enlarged reproduction of the applied input signal. For this achievement, different techniques are employed. One method to obtain greater output is the push-pull arrangement. Two tubes of identical characteristics are connected in push-pull arrangement as shown in fig. 10. The output of the push pull amplifier is twice as large as in a single tube. It has the additional advantage of eliminating even harmonic distortion from the output of the amplifier. Hence for the same maximum permitted harmonic distortion output power much greater than twice that of a single tube may be obtained.

A circuit with two triode in push-pull for audio frequency power amplification is shown in fig. 10.

The input grid cathode voltages are required to have 180°

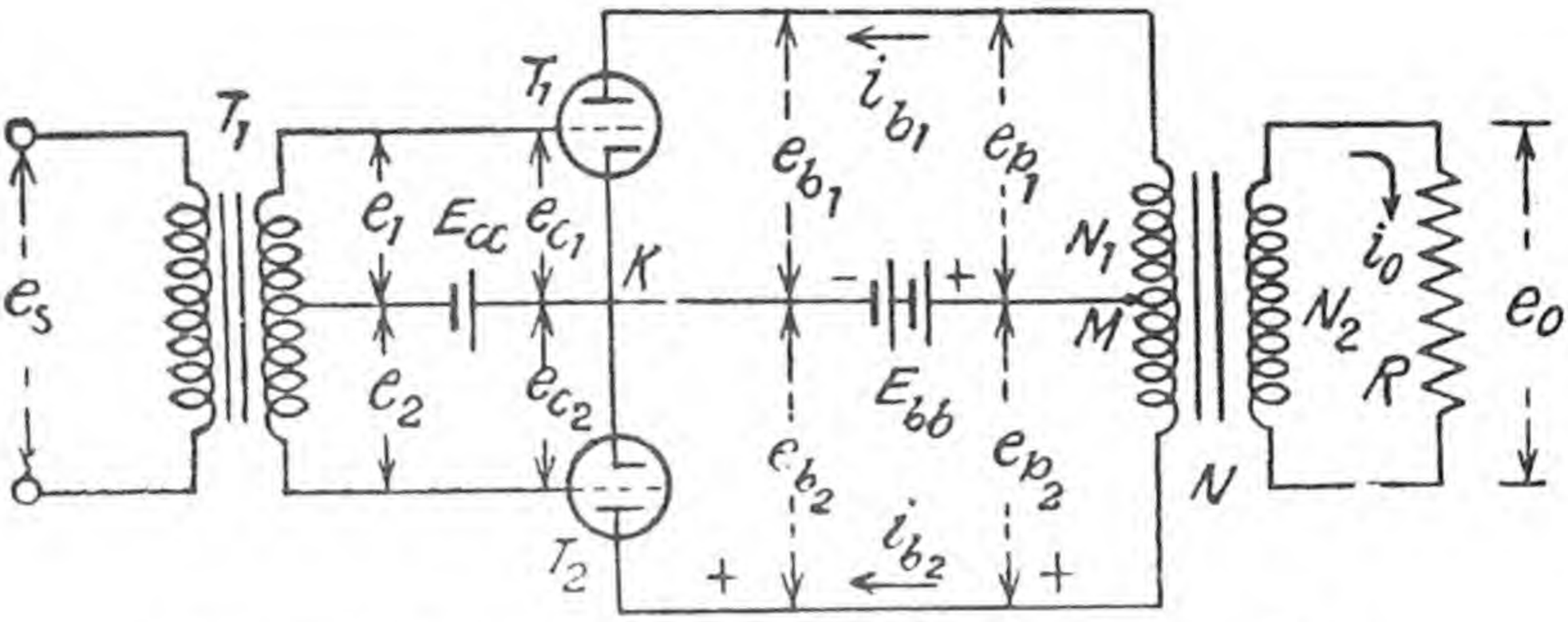


Fig. 10. Circuit arrangement of push-pull amplifier.

phase relation, which is provided by a centre-tapped transformer. Since the grid cathode voltages are out of phase, the plate current of one tube will reach its maximum at the instant the other tube reaches its minimum. This is illustrated in the figure 11, in which

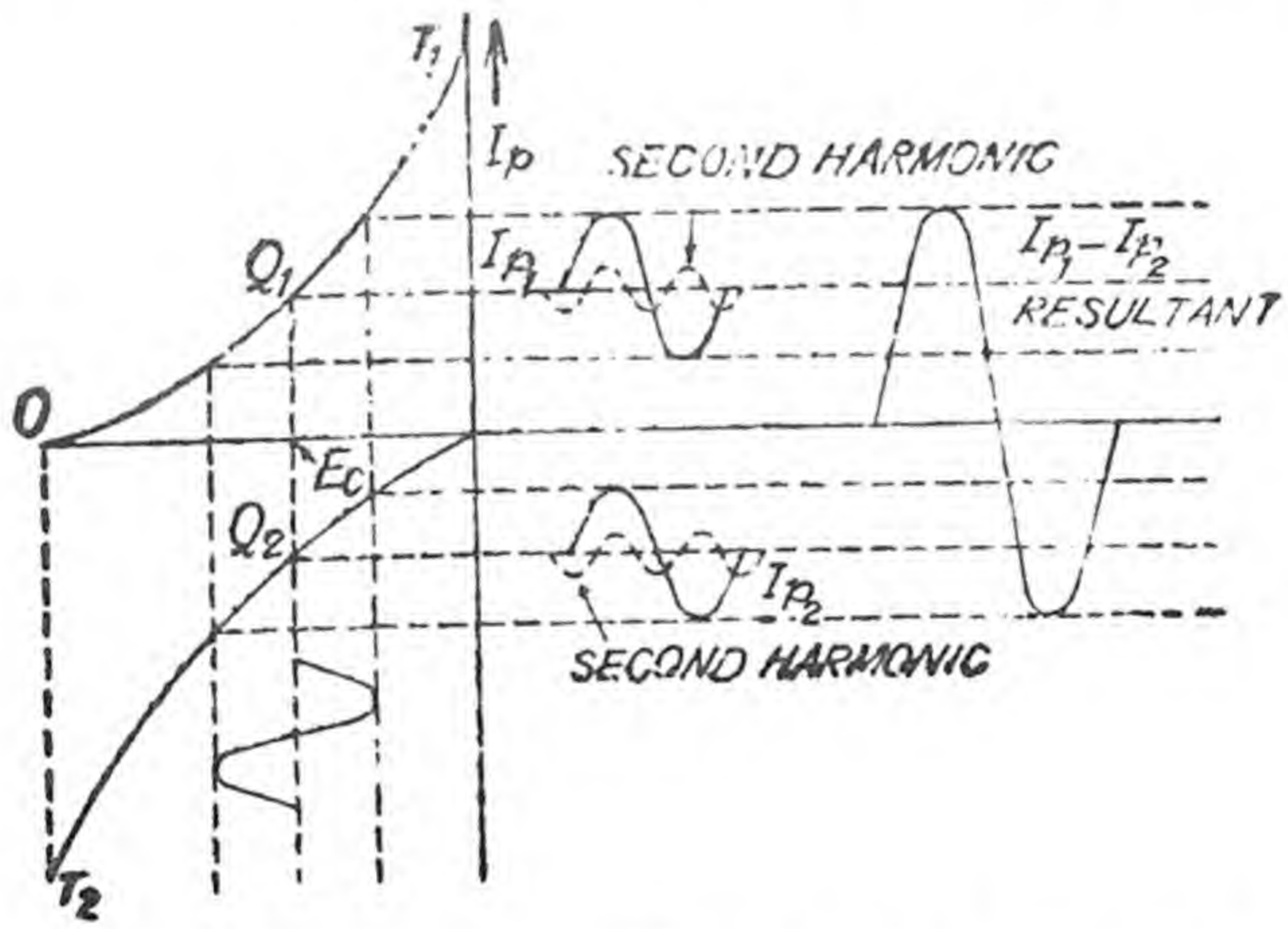


Fig. 11, Output of push-pull amplifier with even harmonics cancelled.

the dynamic characteristic of second tube has been reversed to express the grid voltage phase relation and aligned on the E_c value of the first tube Both of the dynamic characteristics are curved. Harmonic (even) generated in I_{p1} and I_{p2} are reversed and so cancelled. Figure 12 shows the output with even harmonics cancelled.

In the analysis of push-pull amplifier, following assumptions are made :

(*i*) Grid bias and grid drive are so adjusted that plate current flows for the entire a.c. cycle.

(*ii*) Load is a pure resistance.

(*iii*) Transformers are ideal.

(*iv*) Tubes have identical characteristics.

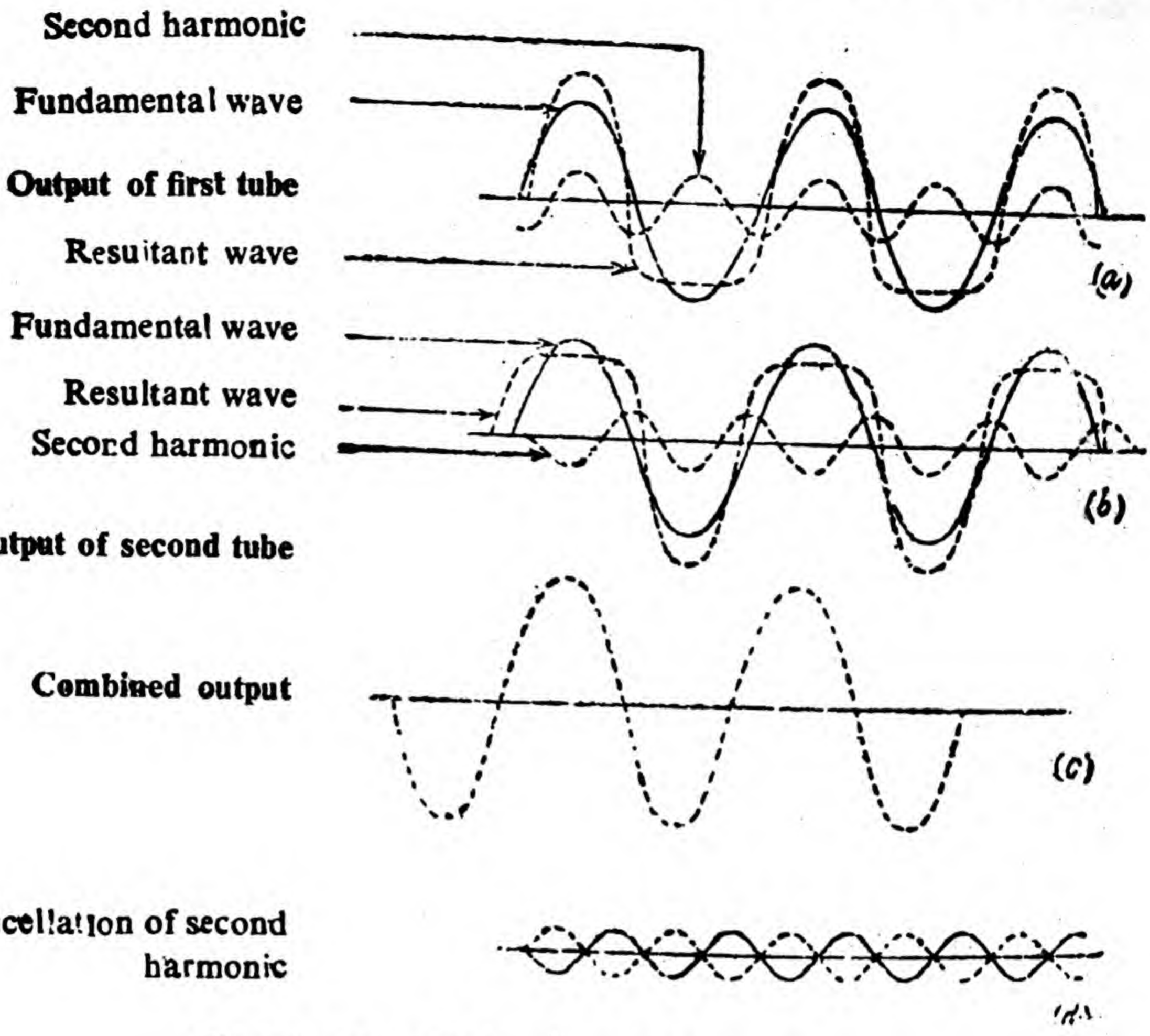


Fig. 12 Cancellation of second harmonic.

To examine certain features of such an amplifier, we consider the following cases :

Zero signal operation of push pull amplifier :

When no signal is applied, the various relations are

$$e_s=0,\ e_1=e_2=0,\ e_{c1}=e_{c2}=E_{cc},\ e_{p1}=e_{p2}=0,$$
$$e_{b1}=e_{b2},\ i_{b1}=i_{b2}=I_b.$$

It is seen that zero signal plate currents in the two tubes flows the same value I_b. Under zero signal condition, the total plate current through the plate supply source is $2I_b$.

Under zero signal condition, plate current of value I_b each flows in opposite directions through the two halves of the primary of output transformer. Consequently, the magnetomotive force in the upper half winding of N_1 turns tends to send the magnetic flux through the core in one direction while the lower half winding of N_1 turns tends to send the magnetic flux in the reverse direction. Thus the net magnetisation of the core resulting from plate current is zero as the number of turns and currents in two sections are equal. This cancellation of magnetisation is an important feature of the push pull connection over the single tube.

Small signal operation of push pull amplifier :

Now let us consider the effect of a small grid signal voltage that the operating point travels over the linear region of the tube

characteristics. Obviously

$$e_1 = -e_2. \qquad \ldots(1)$$

Total instantaneous grid voltages are given by

$$e_{c_1} = E_{cc} + e_1 \qquad \ldots(2)$$

and

$$e_{c_2} = E_{cc} + e_2 = E_{cc} - e_1. \qquad \ldots(3)$$

$$i_{p_1} = i_{p_2}. \qquad \ldots(4)$$

Instantaneous total plate currents are given by

$$i_{b_1} = I_{b_1} + i_{p_1} = I_b + i_{p_1} \qquad \ldots(5)$$

and

$$i_{b_2} = I_{b_2} + i_{p_2} = I_b - i_{p_1} \text{ where } I_{b_1} = I_{b_2} = I_b \qquad \ldots(6)$$

Total current through plate power source

$$= i_{b_1} + i_{b_2} = (I_b + i_{p_1}) + (I_b - i_{p_1}) = 2I_b. \qquad \ldots(7)$$

Equation (7) shows that the current through plate power source has no a.c. component.

Considering the output transformer to be ideal *i.e.* having no leakage inductance and zero exciting current, the sum of the magnetomotive forces around the core due to the currents in these winding must be zero. Thus

$$N_1 i_{b_1} - N_1 i_{b_2} = N_2 \, i_0$$

or

$$i_0 = \frac{N_1}{N_2} (i_{b_1} - i_{b_2}). \qquad \ldots(8)$$

Substituting the values of i_{b_1} and i_{b_2} from equations (5) and (6) we get

$$i_0 = \frac{N_1}{N_2} \cdot 2i_{p_1}. \qquad \ldots(9)$$

This shows that in the output two a c. plate currents i_{p1} are additive in their effect. The output voltage e_0 is due to the flow of current i_0 in the load impedance R which is given by $e_0 = i_0 R$.

Large signal operation of push pull amplifier :

When the signal is of large amplitude, the operation is not confined to linear portion of the tube characteristics and so generates harmonics in plate current. Let the grid signal voltage be given by

$$e_1 = E_m \cos \omega t. \qquad \ldots(10)$$

The harmonics present in the plate current may be obtained from the following expression

$$i_{b_1} = I_b + I_{p0} + I_{p1m} \cos \omega t + I_{p2m} \cos 2\omega t + I_{p3m} \cos 3\omega t + \ldots \qquad \ldots(11)$$

Similarly,

$$i_{b_2} = I_b + I_{p0} + I_{p1m} \cos (\omega t + \pi) + I_{p2m} \cos (2\omega t + 2\pi) + I_{p3m} \cos (3\omega t + 3\pi) + \ldots$$

$$= I_b + I_{p0} - I_{p1m} \cos 2\omega t + I_{p2m} \cos 2\omega t - I_{p2m} \cos 3\omega t + \ldots \qquad \ldots(12)$$

Now $i_0 = \frac{N_1}{N_2}(i_{b_1} - i_{b_2}) = \frac{2N_1}{N_2}(I_{p1m}\cos\omega t + I_{p3m}\cos 3\omega t + \ldots)$ ...(13)

Equation (13) shows that all even harmonics are absent in the output of push pull amplifier. This property is superior to a single tube amplifier.

The circuit through the plate supply source is

$$i_{b_1} + i_{b_2} = 2I_b + 2I_{p_0} + 2(I_{p2m}\cos 2\omega t + I_{p4m}\cos 4\omega t + \ldots). \quad .. (14)$$

Thus with signal voltage the current through the plate supply source contains d.c. current $2I_b$, even harmonic terms and an additional d.c. component $2I_{p0}$.

Equivalent circuit of push-pull amplifier : If a push-pull class *A* amplifier is assumed as operating in a linear mode, the equivalent circuit may be drawn as shown in fig. 13 (*a*).

Further the centre connection carries no a.c. component of plate current due to cancellation that occurs, so that a.c equivalent circuit of figure 13 (*a*) reduces to one as shown in figure 13 (*b*) and the load when referred to the transformer has the value

$$R_{p_1} = \left(\frac{2N_1}{N_2}\right)^2 R.$$

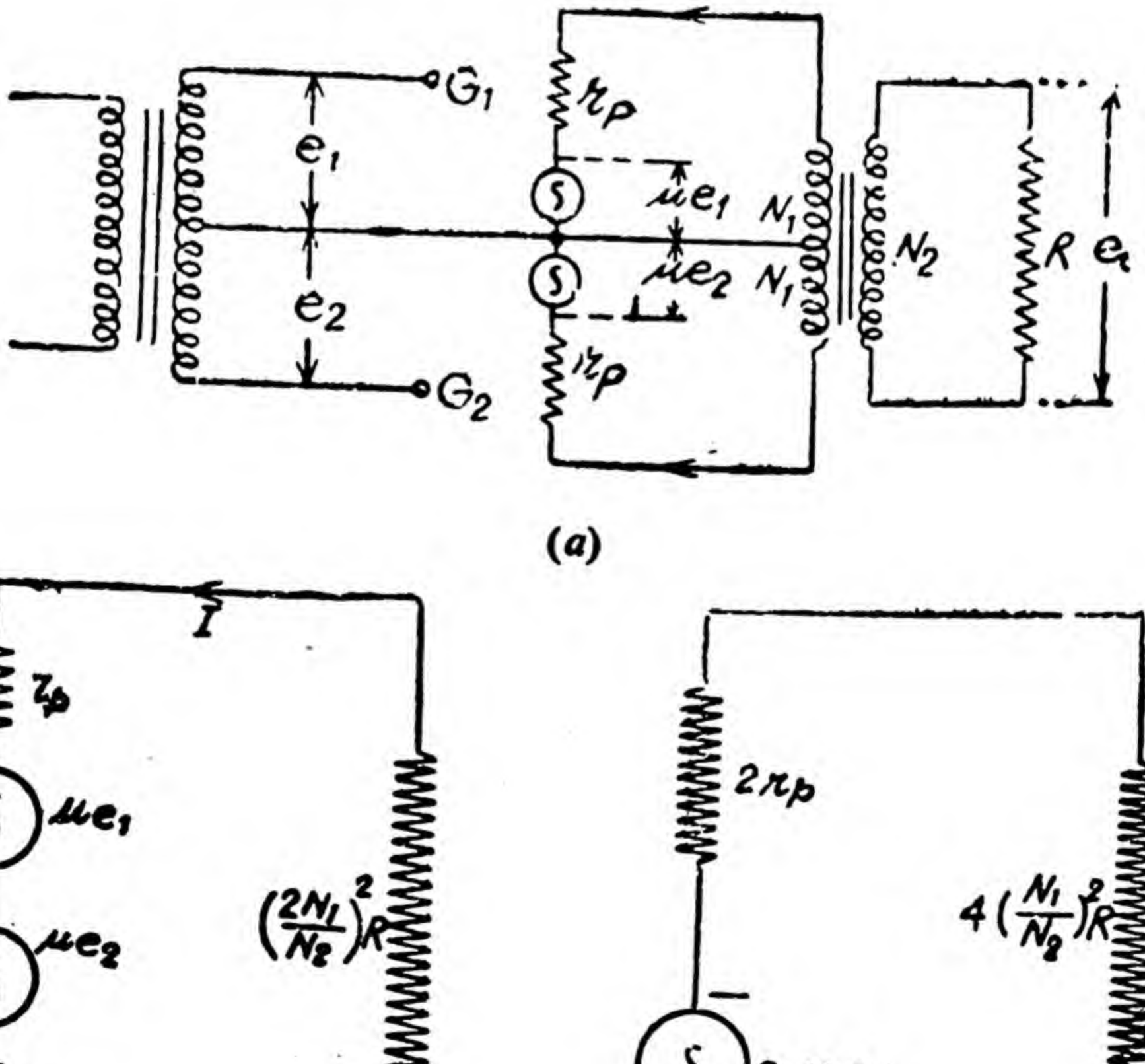


Fig. 13. Equivalent circuit for push-pull amplifier.

Equivalent circuit of figure 13 (*b*) may further be put in the form shown in the figure 13 (*c*) in which generator voltages and internal impedances have been grounded together. From these equivalent circuits, we calculate a.c. plate current, a.c. output voltages, a.c. output power, etc.

The plate current is

$$I=\frac{2\mu e_1}{2r_p+\left(\frac{2N_1}{N_2}\right)^2 K}$$

$$=\frac{\mu e_1}{r_p+2\left(\frac{N_1}{N_2}\right)^2 R}$$

The total power delivered to the load

$$P_{a.c.}=I^2\left(\frac{2N_1}{N_2}\right)^2 R=\left[\frac{\mu e_1}{r_p+2\left(\frac{N_1}{N_2}\right)^2.R}\right]^2\left(\frac{2N_1}{N_2}\right)^2.R.$$

It can easily be determined that maximum power output will be obtained when

$$2r_p=\left(\frac{2N_1}{N_2}\right)^2.R,$$

which states that the plate to plate load impedance should be equal to twice the plate resistance of one tube. Thus

$$P_{max.}=\left(\frac{\mu e_1}{2r_p}\right)^2\left(\frac{2N_1}{N_2}\right)^2 R.$$

12·7. CLASS B AUDIO FREQUENCY POWER AMPLIFIER :

Under Class *B* operation, the grid bias supply potential is made negative by an amount sufficient to reduce the plate current to zero for zero signal potential. The grid bias is so chosen that, when the dynamic characteristics of the two valves are fitted together, their nearly straight parts lie on the same line, as shown in figure 14 (*b*).

This adjustment restricts the flow of anode current to one valve or the other during the greater part of each cycle of the signal voltage. Near the begining of each half cycle, however, anode current flows in both valves, and therefore in both sections of the output transformer primary. The anode circuits are then mutually coupled and the distortion due to curvature of the characteristics near cut-off's reduced. If the two valves are identical, even harmonics will be eliminated from the output by the push-pull action, but odd harmonics remain.

Power Output and Efficiency : A theoretical analysis of class *B* amplifier performance is possible if we assume the composite dynamic characteristics to be linear and the tubes biased to actual cut off. The output current waveform will be sinusoidal input.

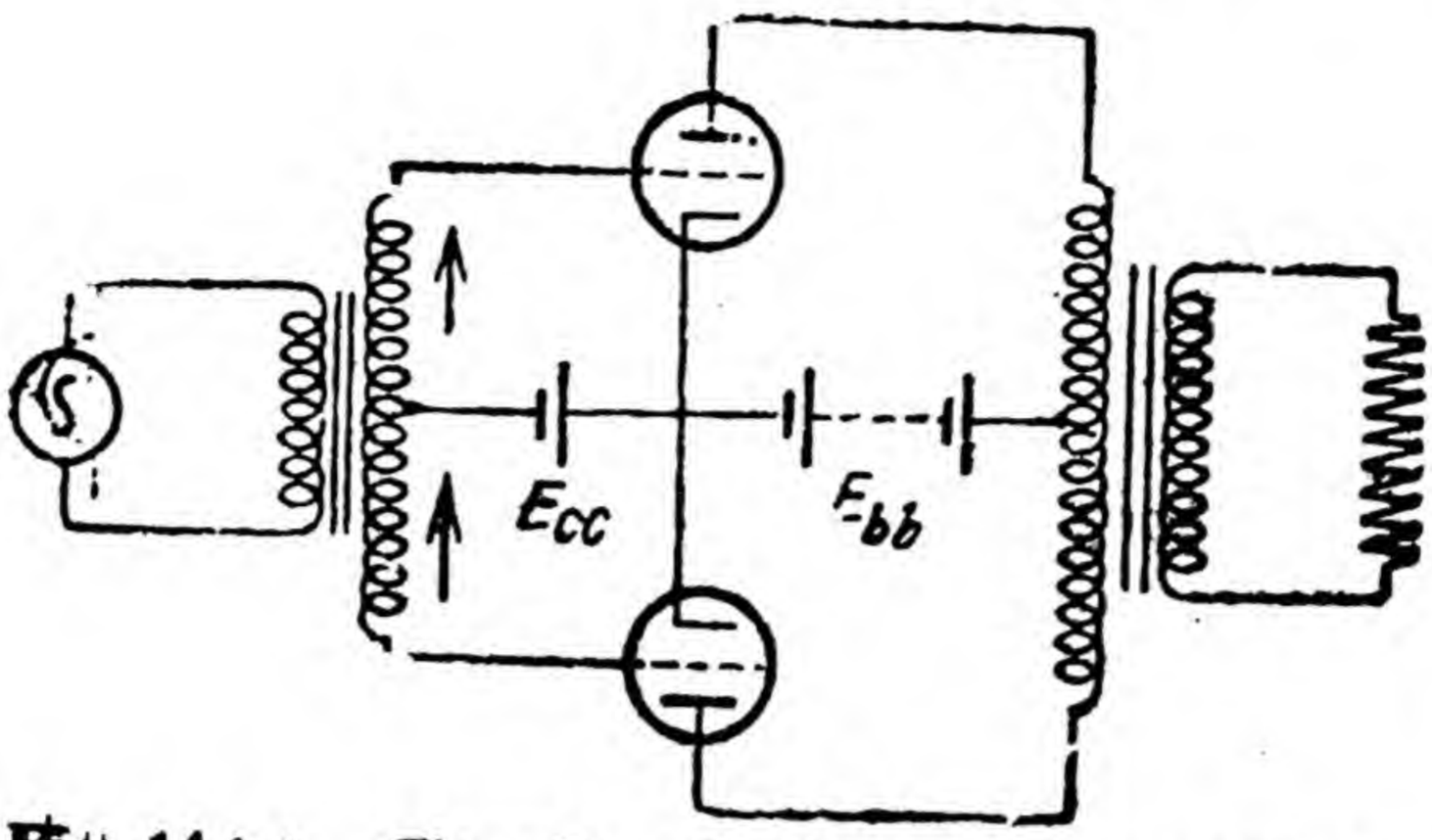


Fig. 14 (*a*). Class B audio frequency amplifier.

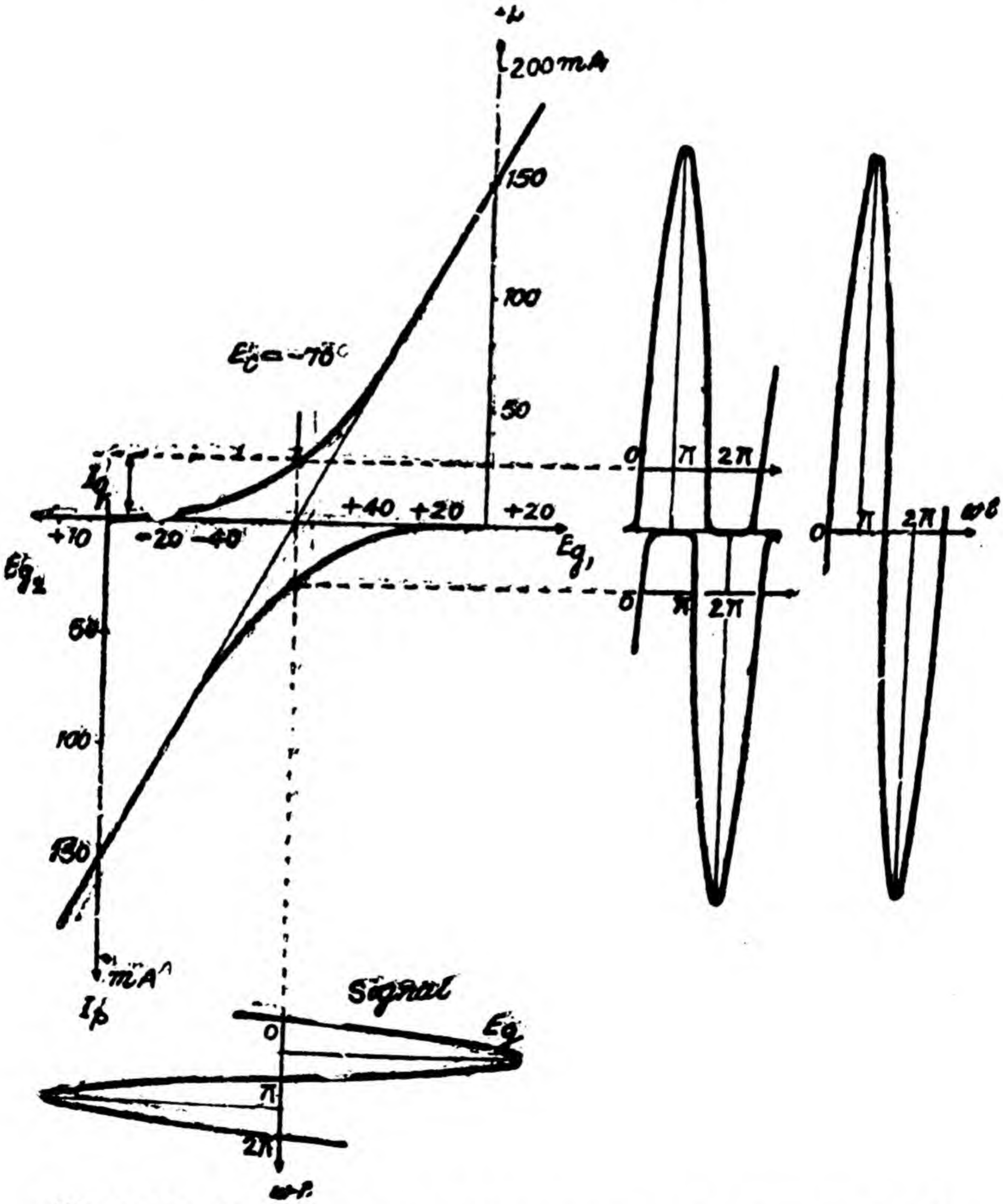


Fig. 14 (*b*). Dynamic transfer characteristics and current waveforms.

By use of properties of sine waves the average value of a half sine wave gives the d.c. current per tube as

$$I_{d.c.}=\frac{I_m}{\pi} \text{ amp./tube,}$$

and the total d.c. plate current is twice this value, where I_m is the peak current value in either tube.

As the output current is full sine wave, the effective a.c. load current is

$$I_{r.m.s.}=\frac{I_m}{\sqrt{2}} \text{ amp.}$$

and the total a.c. power output is

$$P_{a\ c.}=\frac{I_m^2 R}{2} \text{ watts,}$$

where R is the per tube load of the amplifier (one-fourth the plate to plate value). The d.c. power input is easily found as

$$P_{d.c.}=\frac{2I_m E_b}{\pi}.$$

Therefore the plate efficiency is given by the expression

$$\eta_p=\frac{P_{a.c.}}{P_{d.c.}}\times 100\%=\frac{I_m^2\ R/2}{2I_m\ E_b/\pi}\times 100\%.$$

Since $I_m.R$ represents the peak departure of plate voltage from E_b, so $I_mR=E_{min.}$ Thus efficiency will be

$$\eta_p=\frac{\pi}{4}\left(1-\frac{E_{min.}}{E_b}\right)\times 100\%.$$

The minimum value of $E_{min.}=0$.

∴ Maximum theoretical efficiency

$$\eta_p=\frac{\pi}{4}\times 100\%=78{\cdot}5\%.$$

NUMERICAL EXAMPLES

1. *The following data refer to a triode power amplifier using transformer coupled load :*

Plate supply voltage $E_{bb}=350V$
Grid bias $E_{cc}=-20V$
Grid signal $E_g=20\ \sin \omega t$
Zero signal plate current $I_b=30\ mA$
Load resistance across the secondary, $R=3\ \Omega$
Turns ratio of the trnnsformer, $\frac{N_1}{N_2}=30$
Amplification factor of the tube $\mu=10$
Plate resistance $r_p=1000\ \Omega.$

Assuming negligible distortion and ideal output transformer, calculate :

(*i*) *A. C. output voltage,* (*iii*) *A.C. power output,* (*iii*) *Zero signal plate dissipation,* (*iv*) *Plate dissipation with signal and* (*v*) *Plate circuit efficiency*

Since $e_g = 20 \sin \omega t$, the r.m.s. value of the grid voltage

$$E_g = \frac{F_{gm}}{\sqrt{2}} = \frac{20}{\sqrt{2}} = 14{\cdot}14V.$$

Reflected load on primary side $R_{a.c.} = R_L = \left(\frac{N_1}{N_2}\right)^2 R = (30)^2 \times 3$
$= 2700\Omega.$

(*i*) A. C. output voltage $= \frac{-\mu E_g}{r_p + R_L} R_L = \frac{10 \times 14{\cdot}14 \times 2700}{1000 + 2700}$
$= 103{\cdot}1$ V

(*ii*) A. C. power output $= \left(\frac{-\mu E_g}{r_p + R_L}\right)^2 R_L = \left(\frac{10 \times 14{\cdot}14}{3700}\right)^2 2700$
$= 3{\cdot}95$ W.

(*iii*) Zero signal plate dissipation $= P_{d0} = E_{bb}\ I_b$
$= 350 \times 30 \times 10^{-3}$
$= 10{\cdot}5$W.

(*iv*) Plate dissipation with signal $= P_{d.s.} = P_{d0} - P_{ac}$
$= 10{\cdot}5 - 3{\cdot}95 = 6{\cdot}55$ W.

(*v*) Power input from the supply voltage $P_{bb} = E_{bb}\ I_b$
$= 350 \times 30 \times 10^{-3} = 10{\cdot}5$ W.

Plate circuit efficiency $\eta = \frac{P_{ac}}{P_{bb}} \times 100$

$$= \frac{3{\cdot}95 \times 100}{10{\cdot}5} = 37{\cdot}50\%.$$

2. *A linear A_1 amplifier uses plate supply voltage $E_{bb} = 300$ volts grid bias $E_{cc} = -20$ volts, grid signal voltage $= 10$ volts r.m.s. at an angular frequency of 400 radians per second. Load resistance $R_L = 3000$ ohms. At zero signal operating point, plate current $I_q = 50$ mA. amplification factor $\mu = 8$ and dynamic plate resistance $r_p = 1000$ ohms, calculate (i) voltage gain, (ii) a.c. output voltage (iii) power supplied from the plate supply source (iv) zero signal plate dissipation (v) a.c. power output (vi) plate dissipation with signal applied (vii) plate circuit efficiency.*

(i) Voltage gain $= \frac{-\mu R_L}{r_p + R_L} = \frac{-8 \times 3000}{1000 + 3000} = -6$

(ii) a.c. output voltage $= \frac{-\mu E_g R_l}{r_p + R_L} = \frac{-8 \times 10 \times 3000}{1000 + 3000}$
$= -60$ volts.

(iii) Power supplied from the plate supply source is given by
$P_{bb} = E_{bb}\ I_q = 300 \times 50 \times 10^{-3}$
$= 15$ watts

(iv) Zero signal plate voltage
$E_b = E_{bb} - I_q R_L$
$= 300 - 50 \times 10^{-3} \times 3000 = 50$ volts.

Hence zero signal plate dissipation $P_{d0}=E_g.I_g$

$P_{d0}=150\times 50\times 10^{-3}=7{\cdot}5$ watts

(v) **A.C.** power output

$$P_{a\cdot c\cdot}=\left(\frac{\mu E_g}{r_p+R_L}\right)^2 R_L=\left(\frac{8\times 10}{1000+3000}\right)^2\times 3000$$

$$=\frac{3000}{2500}=1{\cdot}4 \text{ watt.}$$

(vi) Plate dissipation with signal applied is given by

$P_{a\cdot c\cdot}=P_{d\cdot o\cdot}-P_{a\cdot c\cdot}=7{\cdot}5-1{\cdot}4=6{\cdot}1$ watts

(vii) Percentage plate circuit efficiency $\eta_p=\frac{P_{a\cdot c\cdot}}{P_{bb}}\times 100.$

$$=\frac{1{\cdot}4}{15}\times 100=9{\cdot}33\%$$

3. *A linear A_1 power amplifier uses triode having amplification factor $\mu=8$ and dynamical plate resistance $=3200$ ohms. Amplitude of grid signal voltage is 15 volts. Calculate the maximum power output obtainable.*

The *A.C.* power output $P_{a\cdot c\cdot}=\left(\frac{\mu F_g}{r_p+R_L}\right)^2 R_L$

The power output is maximum when $r_p=R_L$

$$\therefore \quad (P_{a\cdot c\cdot})_{max}=\left(\frac{\mu E_g}{r_p+r_p}\right)^2 r_p$$

$$=\frac{\mu^2 E_g^2}{4r_p}=\frac{64\times(15/\sqrt{2})^2}{4\times 3200}$$

$$=\frac{64\times 225}{4\times 2\times 3200}=\frac{9}{16}$$

$=0{\cdot}56$ watt.

4. *A class A_1 amplifier uses triode having $\mu=20$ and $r_p=10{,}000$ ohms. A signal of 10 r.m.s. is applied at the input. Calculate the available power output of the amplifier. Express the a.c. output power in terms of the available power if the load impedance is (a) 8000 ohms and (b) $2000+j1000$ ohms.*

(*a*) The a.c. output power $P_{a\cdot c\cdot}=\left(\frac{\mu E_g}{r_p+R_L}\right)^2 R_L$

$$\therefore \quad P_{a\cdot c\cdot}=\left(\frac{20\times 10}{10{,}000+8000}\right)^2\times 8000=0{\cdot}987 \text{ watt.}$$

The available power $P_{av}=\frac{\mu^2 E_g^2}{4r_p^2}\times r_p=\frac{\mu^2 E_g^2}{4r_p}.$

$$\therefore \quad P_{av}=\frac{(20\times 10)^2}{4\times 10^4}=1 \text{ watt.}$$

$$\therefore \quad \frac{P_{a\cdot c\cdot}}{P_{av}}=\frac{0{\cdot}987}{1}=0{\cdot}987.$$

(*b*) In this case $Z_L = 2000 + j\ 1000$,

$$P_{av} = \frac{(\mu E_g)^2}{(r_p + R_L)^2 + (X_L)^2} \times R_L$$

$$= \frac{(20 \times 10)^2 \times 2000}{(10^4 + 2000)^2 + (1000)^2}$$

$$= 0{\cdot}551 \text{ watt.}$$

Hence $\dfrac{P_{a.c.}}{P_{av}} = \dfrac{{\cdot}987}{{\cdot}551} = 1{\cdot}7.$

5. *In a linear A_1 amplifier triode used has amplification factor $\mu = 10$ and dynamical plate resistance $r_p = 10$ kilo ohms, Variable load impedance is a complex impedance of constant Q and has impedance angle of 60 degrees. Calculate the value of resistive and inductive elements of load impedance required for producing maximum a.c. power output. If the applied signal is 12 volts r.m s., calculate the maximnm power output, maximum power sensitivity, plate dissipation and plate circuit efficiency given that plate supply voltage $E_{bb} = 300$ volts and zero signal plate voltage is 200 volts.*

For maximum power output $r_p = \sqrt{(R_L^2 + X_L^2)}$.

According to the problem $\dfrac{X_L}{R_L} = \tan 60° = \sqrt{3}$

or $$X_L = \sqrt{3} R_L.$$

Hence $r_p = \sqrt{\{R_L^2 + (\sqrt{3}\ R_L)^2\}} = \sqrt{(4 R_L^2\}} = 2R_L$

or $$R_L = \frac{r_p}{2} = \frac{10 \times 10^3}{2} = 5 \text{ kilo ohms.}$$

Also $X_L = \sqrt{3} \times 5 \times\ 0^3 = 8660$ ohms.

Now $Z_L = R_L + j\ X_L = 5000 + j\ 8660$ ohm.

The *A.C.* power output $P_{a.c.} = \dfrac{(\mu E_g)^2 \times R_L}{(r_p + R_L)^2 + (X_L)^4}$

$$\therefore\ P_{a.c.} = \frac{(10 \times 1\ )^2 \times 5000}{(10{,}000 + 5000)^2 + (8660)^2}$$

$$= 0{\cdot}24 \text{ watt.}$$

Maximum power sensitivity $= \dfrac{P_{a\ c.}}{E_g^2} = \dfrac{0{\cdot}24}{(12)^2}$

$$= 0{\cdot}00166.$$

Zero signal plate current $I_q = \dfrac{300 - 200}{5000} = 20$ mA.

D.C. input power from plate supply source is given by $P_{bb} = E_{bb} \times I_q = 300 \times 20 \times 10^{-3} = 6$ watt.

Hence plate circuit efficiency $= \dfrac{0{\cdot}94 \times 100}{6} = 4\%.$

Plate dissipation with zero signal applied is given by $P_{pj} = P_{bb} - I_q^2 R_L = 6 - (20 \times 10^{-3})^2 \times 5 \times 10^3 = 6 - 2 = 4$ watts.

Plate dissipation with signal $= P_{p0} - P_{a.c.} = 4 - \cdot 24 = 3 \cdot 76$ watts.

6. *A loudspeaker offers 5 ohm resistive impedance to a pentode output power amplifier. No signal plate current is 30 mA. Input signal to the grid is 15 volts r.m.s. Determine the a.c. power output and plate efficiency. Turns ratio $N_p/N_s = 20$. Amplification factor $\mu = 400$, plate resistance $r_p = 400{,}000$ ohms, plate supply source voltage $= 250$ V.*

Reflected load impedance $R_p = \left(\frac{N_p}{N_s}\right)^2 R_L$

$= (20)^2 \times 5 = 2000.$

Now $P_{ac.} = \left(\frac{\mu E_g}{r_p + R_p}\right)^2 R$

$= \left(\frac{400 \times 15}{4 \times 10^5 + 2000}\right)^2 \times 2000$

$= \cdot 445$ watt.

Power input $= E_{bb} I_q$

$= 250 + 30 \times 10^{-3} = 7 \cdot 5$ watt.

Plate efficiency $\eta = \frac{\cdot 445}{7 \cdot 5} \times 100$

$= 5 \cdot 9\%.$

7. *A transformer coupled power amplifier uses triode whose amplification factor $\mu = 15$, plate resistance $r_p = 2000$ ohms. Loudspeaker forms a load whose resistance is 2 ohms. Plate supply voltage is 300 volts. It is desired to have distortionless amplification by not allowing the plate current to fall below a certain minimum value. For maximum output power, a certain load resistance and grid bias is needed. Calculate the turns ratio for the above load resistance on the primary side. If the grid bias is -10 volts then what is the maximum power output and maximum power sensitivity?*

For undistorted amplification in triode with prescribed zero signal plate voltage the load resistance on primary side is

$$R_p = 2r_p.$$

Hence $2r_p = \left(\frac{N_p}{N_s}\right)^2 R_L$

or $\left(\frac{N_p}{N_s}\right)^2 = \frac{2r_p.}{R_L} = \frac{2 \times 2000}{2} = 2000$

$\therefore \quad \frac{N_p}{R_s} = \cdot 4 \cdot 7.$

For maximum power output

$E_{cc} = \sqrt{2} E_g = 10$ volts.

$E_g = \frac{10}{\sqrt{2}} = 7$ volts.

Maximum power output $(P_a)_{max} = \left(\frac{\mu E_g}{r_p + R_p}\right)^2 R_p$

$$=\frac{\mu^2 E_g^2}{9\, r_p^2}\times 2r_g$$

$$=\frac{2\mu^2 E_g^2}{r_p}$$

$$=\frac{2\times 225\times 49}{9\times 2000}$$

$$=1{\cdot}225 \text{ watt.}$$

Maximum power sensitivity is given by

$$=\frac{(P_{a\cdot c.})_{max}}{E_g^2}=\frac{2\mu^2 E_g^2}{9r_p\times E_g^2}$$

$$=\frac{2\mu^2}{9r_p}=\frac{2\times 225}{9\times 2000}$$

$$=0{\cdot}025 \text{ mho.}$$

8. *In a shunt fed pentode A_1 power amplifier, tube has $\mu=200$ and $r_p=400\ K\Omega$. Plate supply voltage $E_{bb}=360$ volts, Load resistance $R_L=40K\Omega$. Choke has series resistance of $R=25\Omega$. Zero signal plate current $I_q=40$ mA. Calculate the value of input signal voltage required to produce a.c. output power of 3 watts. Also calculate for this value of input signal voltage the following quantities (i) zero signal plate dissipation (ii) plate dissipation with signal (iii) plate circuit efficiency (iv) power sensitivity.*

We know that

$$P_{a\cdot c.}=\frac{\mu^2 E_g^2\times R_L}{(r_p+R_L)^2}$$

or
$$3=\frac{(200)^2\ E_g^2\times 40\times 10^3}{(400\times 10^3+40\times 10^3)^2}$$

or
$$E_g^2=\frac{3\times(440)^2\times 10^6}{(200)^2\times 40\times 10^3}$$

or
$$E_g=11\sqrt{3}=19{\cdot}03.$$

(i) Zero signal plate dissipation P_{p0} is given by

$$P_{p0}=E_q I_q=(E_{bb}-I_q R)\ I_q$$

$$=(360-40\times 10^{-3}\times 25)\ 40\times 10^{-3}$$

$$=(360-1)\ 40\times 10^{-3}$$

$$=140{\cdot}36 \text{ watts.}$$

(ii) Plate dissipation with signal P_p is given by

$$P_p=P_{p0}-P_{a\cdot c.}=14{\cdot}36-3$$

$$=11\ 36 \text{ watts,}$$

(iii) Plate circuit efficiency η_p is given by

$$\eta_p=\frac{P_{a\cdot c.}}{P_{bb}}\times 100=\frac{P_{a\cdot c.}\times 100}{E_{bb}.I_q}$$

$$=\frac{3\times 100}{360\times 40\times 10^{-3}}$$

$$=20{\cdot}83\%.$$

(iv) Power sensitivity $= \frac{P_{a.c.}}{E_0^2} = \frac{3}{(19{\cdot}03)^2}$

$= 8{\cdot}2 \times 10^{-3}$ mho.

9. *A push pull class A_1 power amplifier using two triodes, each having $\mu = 12$ and $r_p = 10K\Omega$ is fed from an anode supply of 500 V. The total d.c. current from the supply is 20 m.A. A load of 5 ohms is connected across the secondary winding of output transformer. If the ratio of primary to secondary winding of output transformer is 40, calculate (a) output voltage, (b) a.c. output power and (c) plate circuit efficiency for an input signal of 20 V r.m.s. applied to each valve.*

In this case $R_{pp} = \left(\frac{2N_1}{N_2}\right)^2 R = (40)^2 \times 5 = 8{,}000$ ohms.

The plate current $I_p = \frac{2\mu\, e_1}{2r_p + (2N_1/N_2)^2\, R}$

$= \frac{2 \times 12 \times 20}{(2 \times 10)^4 + 8000} = 8{\cdot}57 \times 10^{-3}$ Amp.

(*a*) Output voltage is given by

$$\left(\frac{2N_1}{N_2}\right) e_0 = I\, R_{pp}$$

$= 8{\cdot}57 \times 10^{-3} \times 8000 = 68{\cdot}56.$

$\therefore \quad e_0 = \frac{68{\cdot}56}{40} = 1{\cdot}714$ volts.

(*b*) The a.c. power output $P_{a.c.}$ is given by

$P_{a.c.} = I^2.R_{pp} = (8{\cdot}57 \times 10^{-3})^2 \times 8000$

$= 0{\cdot}587$ watt.

(*c*) The plate circuit efficiency η_p is given by

$$\eta_p = \frac{P_{a.c.}}{P_{bb}} \times 100,$$

where $P_{bb} = E_{bb}.2I = 400 \times 20 \times 10^{-3} = 8$ watts.

$\therefore \quad \eta_p = \frac{0{\cdot}587}{8} \times 100 = 7{\cdot}34\%.$

10. *A class A pushpull amplifier uses two triodes each having $\mu = 12$ and $r_p = 10K\Omega$. Plate supply voltage $E_{bb} = 250V$. Zero signal plate current of each tube = 20 m.A. Load impedance is $5 + j5$ ohms at the input signal frequency. The output transformer total primary to secondary turns ratio $(2N_1/N_2)$ is 20. Applied a.c. voltage to each tube is 25 volts. Calculate (i) d.c. input power from plate supply source, (ii) a.c. plate current, (iii) a.c. output voltage across the load impedance, (iv) a c. output power, (v) plate circuit efficiency and (vi) plate dissipation with and without signal.*

(*i*) d.c. input power P_{bb} from the plate supply source is given by

$$P_{bb}=E_{bb}\,2I_g=250\times2\times20\times10^{-3}$$
$$=10 \text{ watts}$$

(ii) Plate to plate impedance $\mathbf{Z}_{pp}$ is given by

$$\mathbf{Z}_{pp}=\left(\frac{2N_1}{N_2}\right)^2\mathbf{Z}_L=(20)^2\,(5+j5)=2000+j\,2000.$$

$\therefore$ a.c. plate current $\mathbf{I}_p=\dfrac{2\mu E_g}{2r_p+\mathbf{Z}_{pp}}=\dfrac{2\times12\times25}{(2\times10^4)+(2000+j\,2000)}$

or $$I_p=\frac{600\times10^{-3}}{\sqrt{\{(32)^2+(20)^2\}}}=20{\cdot}2\times10^{-3} \text{ amp.}$$

(iii) We know that $E_0\,(2N_1/N_2)=I_p\mathbf{Z}_{pp}$.

$\therefore$ a.c. output voltage $E_0=\dfrac{I_p\mathbf{Z}_{pp}}{(2N_1/N_2)}$

$$=\frac{20{\cdot}2\times10^{-3}\times(2000+j\,2000)}{20}$$

$$\therefore\quad E_0=\frac{57{\cdot}13}{20}=2{\cdot}85 \text{ volts.}$$

(iv) a.c. output power $P_{ac}=I_p^2.R_{pp}$

$$=(20{\cdot}2\times10^{-3})^2\times2000$$
$$=0{\cdot}816 \text{ watts.}$$

(v) Plate circuit efficiency $=\dfrac{P_{a.c.}}{P_{bb}}\times100$

$$=\frac{{\cdot}816\times100}{10}=8{\cdot}16\%$$

(vi) Without signal, plate dissipation P_{p0} in each tube is given by

$$2P_{p0}=P_{bb}.$$

$$\therefore\quad P_{p0}=\frac{P_{bb}}{2}=\frac{10}{2}=5 \text{ watts.}$$

With signal, plate dissipation P_p is given by

$$=10-{\cdot}816=9{\cdot}184.$$

$$P_p=\frac{9{\cdot}184}{2}=4{\cdot}592 \text{ watts.}$$

EXERCISES AND PROBLEMS

1. (a) State typical difference between a voltage amplifier and a power amplifier.
   (b) Explain the purpose of a power amplifier.
2. Calculate the power output of class *A* amplifier. What is the approximate overall efficiency you would expect under ideal conditions assuming that the signal input to the grid was the maximum allowable without distortion ?

3. Obtain an expression of plate efficiency used in connection with a radio frequency power amplifier. On what factors does the efficiency depend ?
4. Prove that for maximum output power with undistorted amplification in a triode, with a prescribed zero signal plate voltage, load resistance must be equal to $2r_p$.
5. Describe the action of a transformer coupled amplifier. In case of an ideal transformer coupled A.F. amplifier, draw the equivalent circuit and derive an expression for its gain at medium frequency range.
6. Derive the expression for the power output of class 'A' push pull amplifier.
7. Give the circuit diagram of a class '*B*' audio frequency amplifier stage and explain the principle of operation. What are the advantages and disadvantages of class *C* compared with class *A* operation of audio frequency amplifiers ?
8. Determine the efficiency of class *B* amplifiers. State the reasons for not attaining the maximum efficiency in all cases of amplifiers.
9. A triode power amplifier is operated with plate supply voltage=250 volts, grid bias=−15 volts, grid signal voltage=6 volts r.m s. load resistance =1500 ohms, plate current with no signal=40 mA, amplification factor μ=5, plate resistance=5000 ohms. Calculate
   (a) voltage gain,
   (b) a.c output,
   (c) a c. power output, and
   (d) plate efficiency.
10. Determine the maximum power output of a triode valve used as a power amplifier with amplification factor=5, plate resistance r_p=4000 ohms, max. value of signal input voltage= 10 volts.
11. In a linear A_1 shunt fed triode amplifier load resistance R_L is 50*K* ohm, plate supply voltage E_{bb}=300 volts and choke has resistance R=20 ohm. Zero signal plate current I_q=100 m A. Applied input voltage is $\sqrt{2}.30 \sin 2\pi \times 10^6 t$. Calculate (i) zero signal plate voltage E_b. (ii) d.c input power from plate supply source, (iii) a.c power output $P_{a.c.}$, (iv) plate circuit efficiency η_p, (v) zero signal plate dissipation P_{p0}, (vi) plate dissipation with signal, (vii) availabie power output P_{av} and (viii) power sensitivity.
12. A class A_1 amplifier uses triode which has amplification factor μ=12 and dynamic plate resistance r_p=10,000 ohm. It is operated for distortionless amplification with fixed value of zero signal plat voltage of 200 volts *i.e.* the plate current is never allowed to fall below a certain minimum. Load resistance R_L and grid bias E_{cc} are so chosen as to provide minimum undistorted power output while maintaining *A* operation E_{cc} so used is −18 volts Calculate this value of load resistance R_L, maximum a c. power output and maximum power sensitivity. If grid signal E_g is reduced to 8 volts r.m.s. calculate the power output.
13. A triode is used in the power amplifier circuit. The transformer primary winding resistance is 162 ohms. The total resistance in the secondary circuit is 11 ohms. The transformer turns ratio N_1/N_2=10. If E_{bb}=250 volts and cathode resistor=400 ohms. Draw the d.c. load line and th :

bais line and determine the quiescent point. Find the value of E_{b0}, E_{c0} and I_{b0}. The sinusoidal signal derives the grid to 0 volts relative to its cathode on its positive swing, i_b will not vary exactly sinusoidally because of non-linear plate characteristics but because i_b waveforms repeat periodically we can write, by Fourier analysis, the most important component as

$$i_b = I_{b0} + I_{p1m} \cos \omega t + I_{p2m} \cos 2\omega t,$$

(a) Determine the values of I_{b0}, I_{p1m} and I_{p2m},

(b) Determine the percentage of second harmonic distortion

$$(100\, I_{p2m}/I_{p1m}) \times 100.$$

(c) Determine the average a.c. power, P_{ao} delivered to the transformer primary.

(d) Determine the average plate dissipation P_b of the tube with the signal applied. If this exceeds 15 watts, the tube is over heated.

4. In a class A_1 audio frequency push-pull amplifier, each tube has $\mu = 15$ and $r_p = 10$ k ohms. Grid to cathode voltage E_g of the upper tube is 10 volts. Turn ratio $N_1/N_2 = 20$, where N_1 is the number of turns on half of primary winding in push-pull transformer and N_2 is the number of turns on the secondary. Load resistance $R_L = 5$ ohms. Plate supply voltage $E_{bb} = 200$ volts and zero signal plate current of each tube is 15 m.A. Calculate the a.c. output power, output voltage and plate circuit efficiency. Assume output transformer to be ideal.

15. A push-pull triode circuit is operating under class *B* with a plate to plate load of 500 ohm. The d.c. milliammeter in the common anode circuit reads 122 m.A. for a steady sinusoidal signal. What is the a.c. power output ?

13

RADIO FREQUENCY AMPLIFIERS
(TUNED VOLTAGE AND POWER AMPLIFIERS)

Amplifiers, designed to operate above the audio frequency range, are said to be *radio frequency amplifiers*. They are designed to amplify either a single high frequency or a relatively band of frequencies centred about a high frequency. Similar to audio frequency amplifiers they can be divided into two categories : one r. f. voltage amplifiers and the other r. f. power amplifiers. Their mode of operation can be classified as class *A*, class *B* or class *C*. Class *A* operation is employed when the main purpose is to secure voltage output. This mode is generally employed in electronic circuits and receiving sets. The power output of this mode of operation is relatively low but the distortion is very much reduced. Class *B* and Class *C* amplifiers are used when power output is an important factor. The plate efficiency is greater in both these operations in comparison to class *A* operation. They are used in radio transmitters with the difference that class *B* amplifiers are used after the amplitude modulation of the impressed signal, while class *C* amplifiers are used before the modulation stage.

The basic difference in the circuit of a low frequency and a high frequency amplifier is that of a plate load. The load for low frequency amplifier is resistive or partly inductive in some cases while for high frequency amplifiers it is a tuned circuit.

Radio frequency amplifiers can be further classified according to the width of frequency band that is to be amplified; the one class being the narrow band tuned amplifiers and the other wide band tuned amplifiers. When an active circuit is required to amplify a specific band of frequencies and reject signals of higher and lower frequencies, it is commonly known as a tuned or *narrow band amplifier*. Here the word 'tuned' refers to the circuit load that is normally a parallel *LC* resonant circuit. The parallel tuned circuit has a high impedance at resonant frequency and the impedance falls off sharply at frequencies on either side. The voltage gain of a tuned amplifier is thus maximum at resonant frequency and falls off sharply as the frequency is increased or decreased. In case of a pentode amplifier the voltage gain is almost equal to $-g_m Z_L$ because r_p is very large. Here Z_L is the tuned circuit impedance. Radio sets for the reception of radio frequency wave modulated by

vioce-frequency currents use narrow band amplifiers because the frequency band of such a wave, including the side band frequencies, is relatively small. If the side band frequencies in a modulated wave extend over a wide range, then amplifiers used to amplify them are referred as *wide band tuned amplifiers.* They deal with television signals.

13·1. VALVE PERFORMANCE AT HIGH FREQUENCIES :

At high frequencies above 50 mc./s., the performance of tubes falls off due to the following reasons :

(*i*) *due to stray inductance and capacitance.*
(*ii*) *due to transit time effect, and*
(*iii*) *due to skin effect.*

Here we shall consider only the first factor. At low frequencies the effects of stray inductance and capacitance may be neglected, while at high frequencies they create serious problems. The plate grid capacity in triodes can be eliminated by using the pentodes in frequency range of 100 kc/s to 100 mc/s. At high frequency end of this range the inductance of cathode lead becomes important and feed back takes place through this. To analyse the effect, we consider the circuit as shown in fig. 1.

From figure

$$E_g = E_1 - j\omega L_k I_k \qquad \ldots(1)$$

$$E_g = I_g \cdot \frac{1}{j\omega C_{gk}} \qquad \ldots(2)$$

$$I_k = g_m E_g$$

$$I_k = g_m \cdot \frac{I_g}{j\omega C_{gk}} \qquad \ldots(3)$$

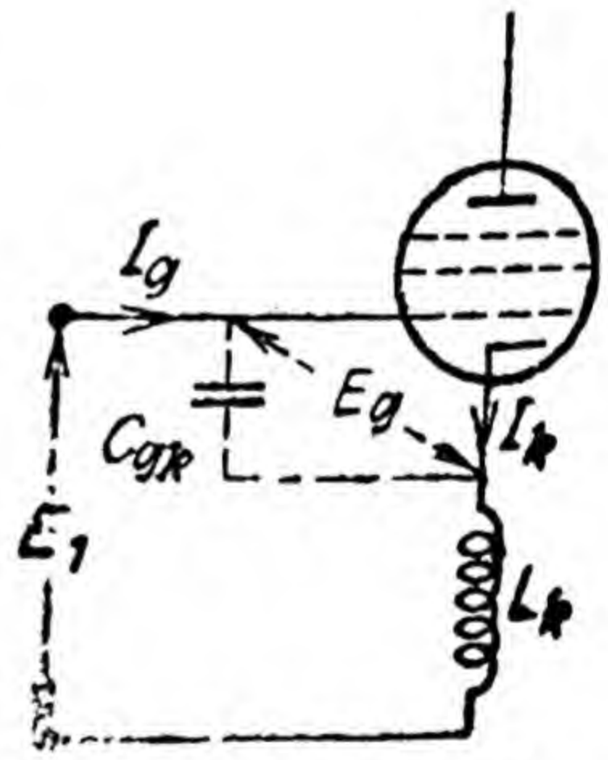


Fig. 1. Cathode inductance at high frequencies.

Substituting the value of E_g and I_k from equations (2) and (3) in equation (1), we have

$$\frac{I_g}{j\omega C_{gk}} = E_1 - j\omega L_k \cdot \frac{g_m I_g}{j\omega C_{gk}}.$$

$$\therefore \quad E_1 = \frac{I_g}{j\omega C_{gk}} [1 + j g_m \omega L_k].$$

The input admittance Y is

$$Y = \frac{I_g}{E_1} = \frac{j\omega C_{gk}}{(1 + j g_m \omega L_k)}. \qquad \ldots(4)$$

Multiplying equation (4) by $(1 - j g_m \omega L_k)$ above and below

$$Y = \frac{j\omega C_{gk}}{1 + \omega^2 g_m^2 L_k^2} + \frac{\omega^2 C_{gk} g_m L_k}{1 + \omega^2 L_k^2 g_m^2}$$

since $\omega^2 g_m^2 L_k^2 \ll 1$

$$Y = j\omega C_{gk} + \omega^2 C_{gk} g_m L_k.$$

The input capacity of the tube is therefore shunted by the conductance which is proportional to the square of frequency, input capacity and cathode lead inductance.

13·2. CONSIDERATION OF CIRCUIT DESIGN AND METHODS OF COUPLING :

R. F. amplifiers differ from audio frequency amplifiers in the choice of values and circuit elements. Triodes are usually not suitable for r. f. amplification on account of the feed back of energy from the output circuit to the input circuit through the interelec trode capacitance (C_{gp}) This feed back action, which at higher frequencies becomes prominent because of effectiveness of plate grid capacitance at these frequencies, makes the amplifier unstable and self oscillations may occur if the plate circuit is sufficiently inductive. If the impedance in the plate circuit is pure resistance or a capacitive reactance, no oscillations can take place although in the latter case anti regenerative feed back of sufficient magnitude may occur greatly to reduce the resultant gain. For this reason in r. f. amplifier tetrodes and pentodes are preferred.

R F. Amplifiers can be further classified according to methods used to couple the output circuit of the amplifiers to the input circuit of the tube in the following stages. The usual methods of coupling are :

(*i*) resistance-capacitance coupled circuits,

(*ii*) impedance coupled circuits, and

(*iii*) transformer coupled circuits.

The last two coupling methods are commoly employed in r f. amplifiers.

13·3. RC COUPLING :

It is used when untuned amplifier circuits are desired. RC coupling suffers from the following drawbacks :

(*a*) The resistances used have self-capacities in parallel with themselves, the value of the shunt impedance which they provide being lower, the higher the frequency.

(*b*) The anode cathode capacitance (C_{pk}) of the valve and the grid cathode capacitance (C_{gk}) of the succeeding valve together form a low impedance in parallel with the resistance; the effective value of the external impedance becomes lower the higher the frequency and a smaller voltage amplification factor is obtained.

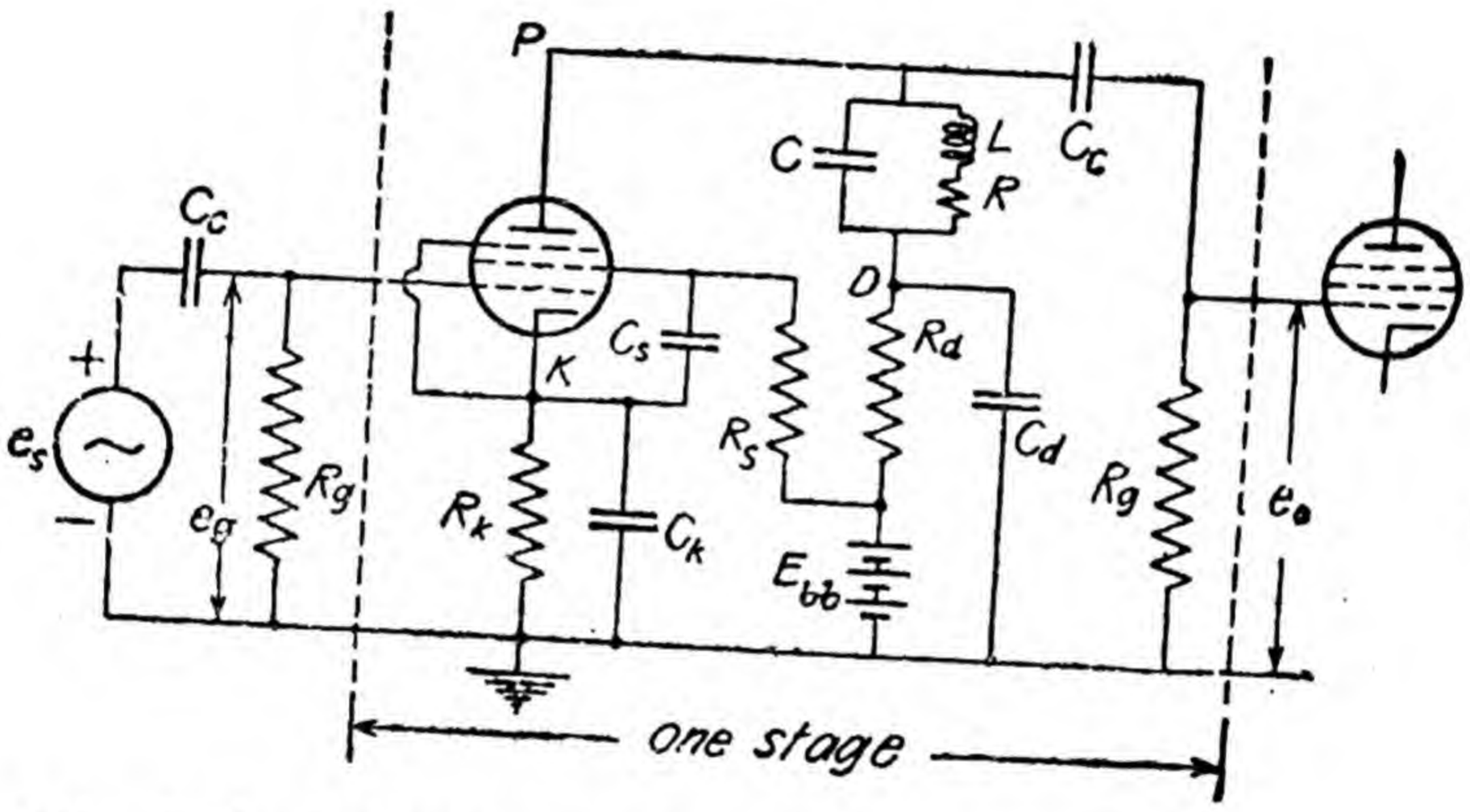
(*c*) The most serious drawback of this coupling in radio frequency stage is that the Miller effect becomes predominant.

13·4. IMPEDANCE COUPLING :

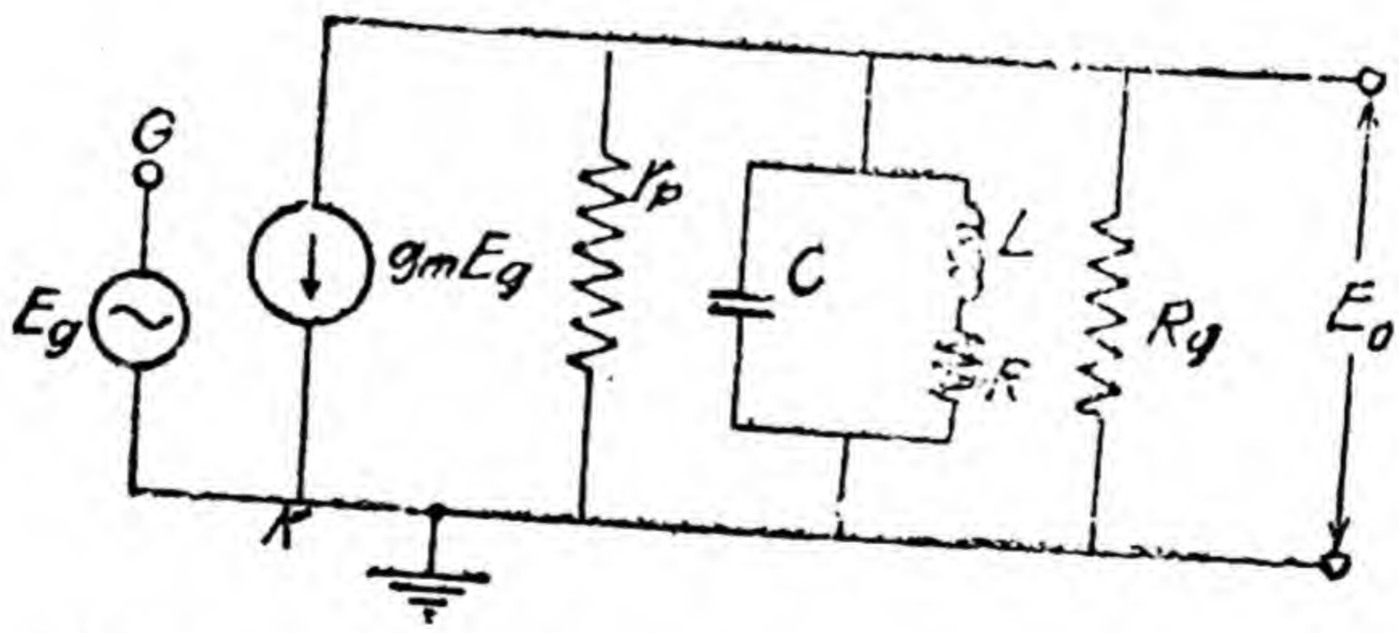
Impedance coupling is quite satisfactory for certain frequency bandwidth which is to be amplified uniformly. The input and

distributed capacitance tune the coil to parallel resonance, so that uniform amplification can be obtained only for the range of frequencies for which the parallel resonant impedance of the plate load is large compared to the plate resistance of the tube.

The simplest form of a tuned amplifier which employs a parallel resonant circuit as the plate load impedance is shown in fig. 2(*a*). It is similar to an ordinary RC coupled amplifier except for the tube plate load which is tuned to a single frequency, R_d and R_s are the plate and screen voltage dropping resistors which enable d.c. voltages to be applied to plate and screen according to the tube specifications. The combinations R_sC_s and R_dC_d act as decoupling circuits. The combination R_kC_k provides the proper cathode bias whereas R_gC_c provides coupling of signals to the grid of the next stage.



(*a*) A single tuned direct coupled amplifier circuit.



(*b*) A. c. equivalent circuit.

Fig. 2

The a.c. equivalent circuit is shown in fig. 2(*b*) where the tube s replaced by an equivalent current source, g_mE_g in parallel with e. Since C_k and C_d virtually act as short circuits at signal frequency he cathode and point *D* are at ground potential. The combinaons R_k C_k and R_d C_d are eliminated from the equivalent circuit

because they offer negligible reactance at the same signal frequency. As the reactance of C_g is negligibly small, the screen grid is also at the cathode potential and hence R_g is also excluded from the equivalent circuit. The coupling capacitor C_c is also neglected as it is supposed to be large enough to offer negligible reactance as compared to R_g at the same signal frequency. Thus the whole a.c. voltage develops across the plate load and appears across R_g. The inter-electrode capacitance between grid and cathode is very small for pentode tubes and thus offers a very high reactance making it virtually an open circuit. Hence it is not shown in the equivalent circuit.

From the equivalent circuit, the output voltage is

$$\mathbf{E}_0 = -g_m \mathbf{E}_g \mathbf{Z},$$

and voltage amplification of such a circuit at any frequency is given by

$$\mathbf{A} = \frac{\mathbf{E}_0}{\mathbf{E}_g} = -g_m \mathbf{Z}, \qquad \ldots (1)$$

where $\mathbf{Z}$ is the total impedance. The total load impedance is given by

$$\frac{1}{\mathbf{Z}} = \frac{1}{r_p} + \frac{1}{R_g} + \frac{1}{\mathbf{Z}_t}, \qquad \ldots (2)$$

where $\mathbf{Z}_t$ is the impedance of the anti-resonant circuit.

The impedance $\mathbf{Z}_t$ of anti-resonant circuit consisting of R, L and C is given by

$$\mathbf{Z}_t = \frac{(R + j\omega L)\left(\frac{1}{j\omega C}\right)}{R + j\omega L + \frac{1}{j\omega C}} = \frac{\frac{L}{C} - \frac{jR}{\omega C}}{R\left(1 + \frac{j\omega L}{R} + \frac{1}{j\omega CR}\right)}.$$

$$= \frac{R\left(\frac{L}{CR} - \frac{j}{\omega C}\right)}{R\left(1 + \frac{j\omega L}{R} + \frac{1}{j\omega CR}\right)} = \frac{\frac{L}{CR} - \frac{j}{\omega C}}{1 + \frac{j\omega L}{R}\left(1 - \frac{1}{\omega^2 LC}\right)} \qquad \ldots (3)$$

The resonant angular frequency of the circuit with $R = 0$ is given by

$$\omega_0 = \frac{1}{\sqrt{(LC)}}. \qquad \ldots (4)$$

The magnification factor or figure of merit Q_0 for the circuit at resonant angular frequency ω_0 is given by

$$Q_0 = \frac{\omega_0 L}{R}. \qquad \ldots (5)$$

Equation (3) can be written as

$$Z_t = \frac{\frac{L}{CR} - \frac{j}{\omega C}}{1 + j\frac{L}{R}\left(\omega - \frac{1}{\omega LC}\right)} = \frac{\frac{L}{CR} - \frac{j}{\omega C}}{1 + j\frac{L\omega_0}{R}\left(\frac{\omega}{\omega_0} - \frac{1}{\omega_0 \omega LC}\right)}$$

$$= \frac{\frac{L}{CR} - \frac{j}{\omega C}}{1 + jQ_0\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right)} \quad \ldots(6)$$

Let δ indicate the variation in frequency expressed as a fraction of the resonant frequency, then

$$\delta = \frac{\omega - \omega_0}{\omega_0} = \frac{\omega}{\omega_0} - 1$$

or
$$\frac{\omega}{\omega_0} = 1 + \delta \text{ and } \frac{\omega_0}{\omega} = \frac{1}{1+\delta} \quad \ldots(7)$$

Substituting the values from equation (7) in equation (6), we get

$$Z_t = \frac{\frac{L}{CR} - \frac{j}{\omega C}}{1 + jQ_0\left[(1+\delta) - \frac{1}{(1+\delta)}\right)} = \frac{\frac{L}{CR} - \frac{j}{\omega C}}{1 + j Q_0 \delta \left[\frac{\delta/2}{1+\delta}\right]} \quad \ldots(8)$$

When ωL is large compared to R. *i e.* $\omega L/R \gg 1$, as is true in most practical circuits, the imaginary term in equation (8) may be neglected. Moreover, when $\delta \ll 1$, the equation (8) has the form

$$Z_t = \frac{\frac{L}{CR}}{1 + j\,2Q_0\delta} \text{ when } \frac{\omega L}{R} \gg 1 \text{ and } \delta \ll 1 \quad \ldots(9)$$

At resonance $\delta = 0$ and impedance is a real quantity

i.e.

$$Z_t = R_t$$

$$Z_t = R_t = \frac{L}{CR} = \frac{\omega_0 L}{\omega_0 CR} = \frac{Q_0}{\omega_0 C} = \omega_0 L Q_0$$

Now
$$Z_t \simeq \frac{R_t}{1 + 2jQ_0\delta} \quad (10)$$

(i) Voltage gain at resonance :

$$A_{res.} = \frac{-g_m}{\frac{1}{r_p} + \frac{1}{R_g} + \frac{1}{R_t}} = \frac{-g_m R_t}{1 + \frac{R_t}{R_g} + \frac{R_t}{r_p}}$$

$$= \frac{-g_m\, \omega_0 L Q_0}{1 + \frac{\omega_0 L Q_0}{R_g} + \frac{\omega_0 L Q_0}{r_p}}$$

$$= -g_m \omega_0 L Q_e, \qquad \ldots(11)$$

where $$Q_e = \frac{Q_0}{1 + \frac{\omega_0 L Q_0}{R_g} + \frac{\omega_0 L Q_0}{r_p}}.$$

This equation suggests that for the voltage gain to be large not only Q_0 should be large but at the same time r_p and R_g should also be large. This latter condition can be met by a selection of pentode instead of a triode since pentode has higher plate resistance.

(ii) Voltage gain at any frequency ω near resonance :

$$\mathbf{A} = \frac{-g_m}{\frac{1}{r_p} + \frac{1}{R_g} + \frac{1 + j2Q_0\delta}{\omega_0 L Q_0}}$$

$$= \frac{-g_m \omega_0 L Q_0}{1 + j2Q_0\delta + \frac{\omega_0 L Q_0}{r_p} + \frac{\omega_0 L Q_0}{R_g}}$$

$$= \frac{-g_m \omega_0 L Q_0}{j2Q_0\delta + (\omega_0 L/Q_e)} = \frac{-g_m \omega_0 L Q_e}{1 + j2\delta Q_e} \qquad \ldots(12)$$

Hence

$$\frac{\mathbf{A}}{\mathbf{A}_{res.}} = \frac{1}{1 + j2\delta Q_e} \qquad \ldots(13)$$

$$\text{Magnitude } \left|\frac{A}{A_{res.}}\right| = \frac{1}{\sqrt{[1 + (2\delta Q_e)^2]}}. \qquad \ldots(14)$$

The phase angle of $\mathbf{A}/\mathbf{A}_{res.}$ is given by

$$\phi = \tan^{-1} 2\delta Q_e. \qquad \ldots(15)$$

Figure (3) shows the plot of the relative gain. If

$$(2\delta Q_e)^2 = 1$$

in equation (14), then

$$\frac{A}{A_{res.}} = \frac{1}{\sqrt{2}} = 0{\cdot}707. \qquad \ldots(16)$$

Then the gain is 70% of the gain at resonance.

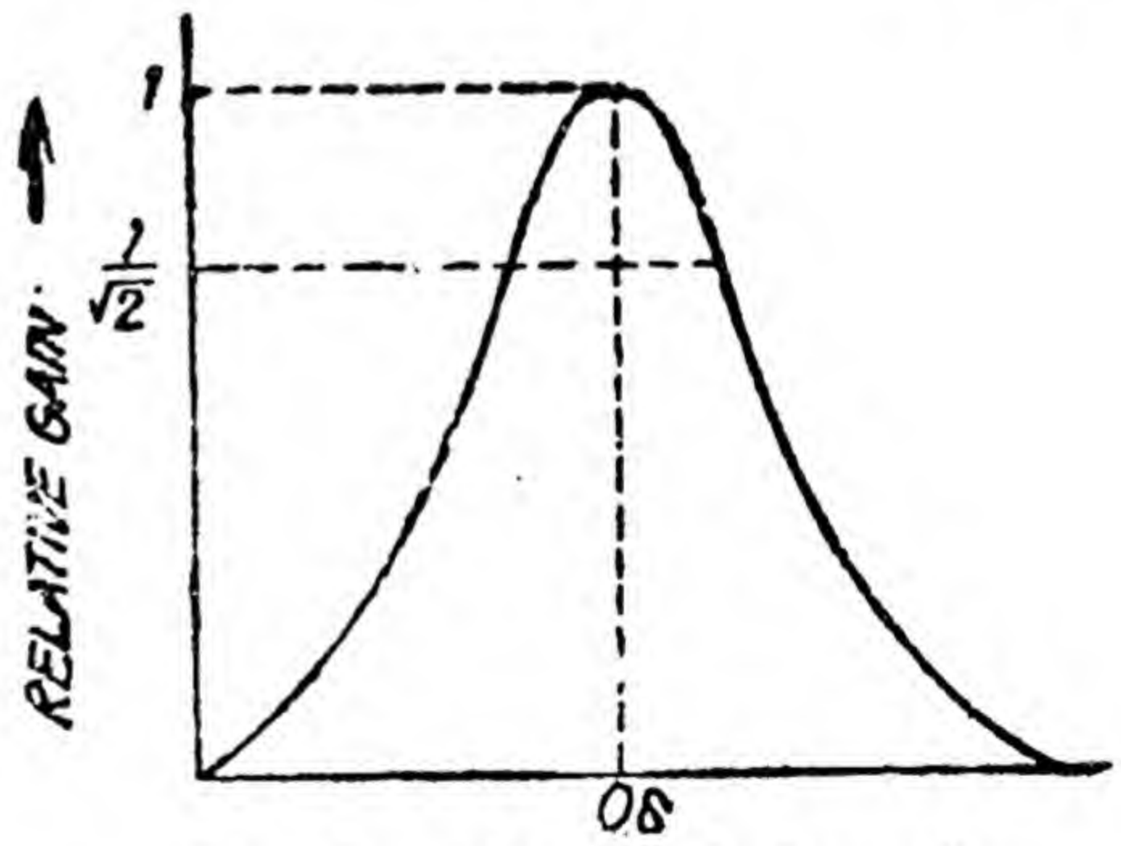


Fig. 5. Plot of relative gain.

(iii) Band width of the amplifier :

When $4\delta^2 Q_e^2 = 1$, $\delta Q = \pm \frac{1}{2}$. The positive value of δQ_e corresponds to the upper cut off frequency f_2 and the negative value of δQ_e corresponds to the lower cut off frequency f_1.

When $\delta Q_e = \frac{1}{2}$

or $$\frac{f_2 - f_0}{f_0} Q_e = \frac{1}{2}$$

$$\therefore \quad f_2 - f_0 = \frac{f_0}{2Q_e} \qquad \ldots(17)$$

where f_2 is the upper cut off frequency at which the gain is 70% of the gain at the resonant frequency f_0.

Similarly, when $\delta Q_e = -\frac{1}{2}$

$$\delta Q_e = \frac{f_0 - f_1}{f_0} Q_e = \frac{1}{2}$$

or $$f_0 - f_1 = \frac{f_0}{2Q_e} \qquad \ldots(18)$$

The bandwidth of the amplifier = B.W. $= f_2 - f_1$

$$\therefore \quad f_2 - f_1 = f_2 - f_0 + f_0 - f_1$$

$$= \frac{f_0}{2Q_e} + \frac{f_0}{2Q_e} = \frac{f_0}{Q_e}. \qquad \ldots(19)$$

Thus the bandwidth of a single tuned amplifier is inversely proportional to the Q_e of the circuit *i.e.*, as Q_e decreases bandwidth increases. For a narrow band amplifier, a large bandwidth is not desirable. A large bandwidth is an important consideration for wide band tuned amplifier such as used in television circuits. The Q_e in turn depends upon the ratio L/C and if ω_0 and L/C are fixed, then the increase in Q of the tuned circuit increases the effective Q_e and correspondingly decreases the bandwidth or increases the frequency selectivity of an amplifier.

13·5. TRANSFORMER COUPLING :

This type of coupling is commonly used in r f. amplifiers. The coils of transformer for the use in r.f. amplifiers are wound on a former of insulating material. Tight coupling between the primary and secondry cannot be obtained without the use of iron core. But the iron core cannot be used because of excessive eddy currents and hysteresis losses which would occur at high frequencies. We shall describe here two types of transformer coupled r.f. amplifiers.

(a) Secondary Tuned Transformer Coupled Amplifier : A

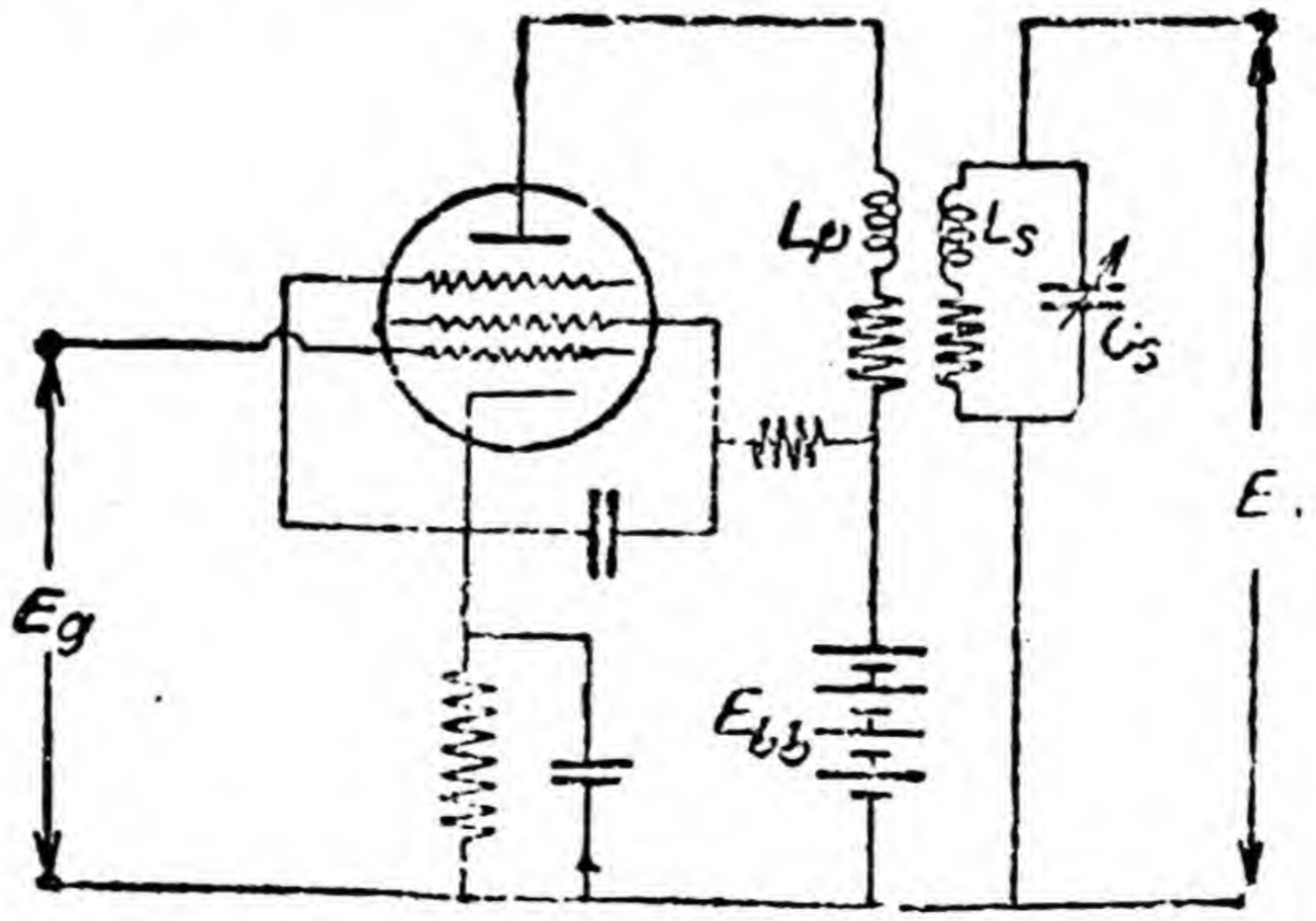


Fig. 4. Transformer coupled with tuned secondary r.f. amplifier.

simple circuit of a r.f amplifier employing transformer coupling with tuned secondary is shown in figure 5. The circuit is similar to fig 2 (*a*) except that a transformer is used to couple the output voltage to the next stage. Here no grid leak resistance R_g for second stage is necessary as the d.c. path between the grid and cathode is provided by the secondary winding itself. These are extensively used in radio receivers because they affect the sensitivity, selectivity, and fidelity to a considerable extent.

The equivalent circuit of single tuned transformer coupled amplifier is shown in figure 5. In this circuit R_1 and R_2 are the resistances of primary and secondary coils respectively, r_p is the plate resistance which is greater than R_1 and R_2. M is the mutual inductance and C_2, the total shunt capacitance on the secondary including the input capacitance of the following stage. Let $\mathbf{I}_1$ and $\mathbf{I}_2$ be the currents flowing in primary and secondary circuits respectively. The equations for the two loops are :

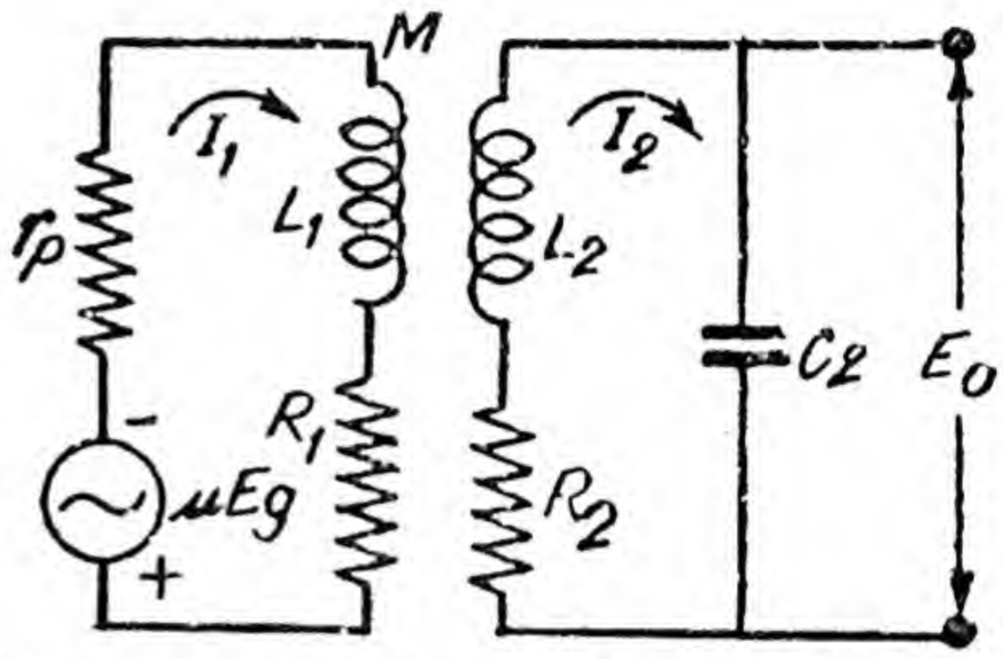


Fig. 5. Equivalent circuit of single tuned transformer coupled amplifier.

$$-\mu \mathbf{E}_g = \mathbf{I}_1 (r_p + R_1 + j\omega L_1) + j\omega M \mathbf{I}_2 \quad \ldots(1)$$

and
$$0 = \mathbf{I}_2 \left(R_1 + j\omega L_2 + \frac{1}{j\omega C_2} \right) + j\omega M \mathbf{I}_1. \quad \ldots(2)$$

From equation (2),

$$\mathbf{I}_1 = - \frac{\mathbf{I}_2 \left(R_2 + j\omega L_2 + \frac{1}{j\omega C_2} \right)}{j\omega M} \quad \ldots(3)$$

Substituting the value of $\mathbf{I}_1$ from equation (3) in equation (1), we have

$$-\mu \mathbf{E}_g = - \frac{I_2 \left(R_2 + j\omega L_2 + \frac{1}{j\omega C_2} \right) (r_p + R_1 + j\omega L_1)}{j\omega M} + j\omega M I_2$$

or
$$\mathbf{I}_2 = \frac{\mu \mathbf{E}_g \, j\omega M}{\left(R_2 + j\omega L_2 + \frac{1}{j\omega C_2} \right) (r_p + R_1 + j\omega L_1) + \omega^2 M^2}. \quad \ldots(4)$$

The output voltage is given by

$$\mathbf{E}_0 = \frac{I_2}{j\omega C_2} = \frac{\mu M \mathbf{E}_g / C_2}{\left(R_2 + j\omega L_2 + \frac{1}{j\omega C_2} \right) (r_p + R_1 + j\omega L_1) + \omega^2 M^2} \quad \ldots(5)$$

The gain of the amplifier $\mathbf{A} = \dfrac{\text{Output voltage}}{\text{Input voltage}} = \dfrac{\mathbf{E}_0}{\mathbf{E}_g}$

$$\therefore \quad A=\frac{\mu M/C_2}{\left(R_2+j\omega L_2+\frac{1}{j\omega C_2}\right)(r_p+R_1+j\omega L_1)+\omega^2M^2} \qquad ...(6)$$

Usually r_p is greater than R_1 and ωL_1, hence

$$A=\frac{\mu M/C_2}{r_p\left[R_2+j\left(\omega L_2-\frac{1}{\omega C_2}\right)\right]+\omega^2M^2} \qquad ..(7)$$

(i) Voltage gain at resonance :

When tne secondary circuit is tuned to resonance

$$\omega_0 L_2=\frac{1}{\omega_0 C_2}. \qquad ..(8)$$

Now considering equation (7) at resonance, we haae

$$(A)_{resonance}=\frac{\mu M/C_2}{r_pR_2+\omega_0^2M^2}. \qquad ...(9)$$

The value of Q at resonance is given by

$$Q_0=\frac{\omega_0L_2}{R_2}=\frac{1}{\omega_0C_2R_2}=\frac{1}{R_2}\sqrt{\left(\frac{L_2}{C_2}\right)} \qquad ...(10)$$

In terms of Q_e, the equation (9) can be written as

$$(A)_{resonance}=\frac{\mu M/r_pC_2R_2}{1+\omega_0^2M^2/r_pR_2}$$

$$=g_m.\frac{\omega_0MQ_0}{1+\omega_0^2M^2/r_pR_2} \qquad ...(11)$$

$$=g_m\,\omega_0\,MQ_e,$$

where $$Q_e=\frac{Q_0}{1+\frac{\omega_0^2M^2}{r_pR_2}}$$

and is the effective value of Q_0.

It appears from equation (11) that the voltage gain increases as M increases. But this is not so, because the term $\omega_0^2M^2$ becomes important with increasing M. There is a optimum value of M for the maximum gain which can be obtained from the condition

$$\frac{\partial(A)_{resonance}}{\partial M}=\frac{\mu/C_2}{r_pR_2+\omega^2M^2}-\frac{(\mu M/C_2)\,2\omega_0^2M}{(r_pR_2+\omega_0^2M^2)^2}=0.$$

Solving we have $(M)_{optimum}=\frac{\sqrt{(R_2\,r_p)}}{\omega_0}=\sqrt{(R_2.r_pL_2C_2)}$ ...(12)

(ii) Voltage gain at near about resonant frequency

For any frequency ω close to constant frequency ω_0, let us put $\delta=\frac{\omega}{\omega_0}-1$ and $Q_0=\frac{\omega_0L_2}{R_2}=\frac{1}{\omega_0C_2R_2}=\frac{1}{R_2}\sqrt{\left(\frac{L_2}{C_2}\right)}$.

$$\left(\omega L_2-\frac{1}{\omega C_2}\right)=\left(\frac{\omega\omega_0L_2}{\omega_0}-\frac{\omega_0}{\omega\omega_0C_2}\right)$$

$$= \omega_0 L_2 \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right) \text{ as } \omega_0 L_2 = \frac{1}{\omega_0 C_2}$$

$$= \sqrt{\left(\frac{L_2}{C_2}\right)}\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right)$$

$$= \sqrt{\left(\frac{L_2}{C_2}\right)}\left\{(1+\delta) - \frac{1}{(1+\delta)}\right\}.$$

Since ω is very nearly equal to ω_0, δ is small and $1/(1+\delta)$ can be expanded by Binomial theorem.

Thus

$$\frac{1}{(1+\delta)} = (1+\delta)^{-1} = 1 - \delta + \delta^2 - \delta^3 + \ldots \cong (1-\delta)$$

Therefore,

$$\left(\omega L_2 - \frac{1}{\omega C_2}\right) \approx 2\delta \sqrt{\left(\frac{L_2}{C_2}\right)} \approx 2\delta R_2 Q_0. \qquad \ldots(13)$$

Equation (7) now may be written as

$$\mathbf{A} = \frac{\mu M / C_2}{r_p R_2 \left[1 + j \frac{1}{R_2} 2\delta R_2 Q_0\right] + \omega^2 M^2}$$

$$= \frac{g_m \omega_0 M Q_0}{1 + \frac{\omega_0^2 M^2}{r_p R_2} + j\, 2\delta Q_0} = \frac{g_m \omega_0 M Q_0 / (1 + \omega_0^2 M^2 / r_p R_2)}{1 + j 2\delta Q_0 / (1 + \omega_0^2 M^2 / r_p R_2)}$$

$$= \frac{g_m \omega_0 M Q_0}{1 + j 2\delta Q_0}.$$

Hence $\dfrac{\mathbf{A}}{\mathbf{A}_{res.}} = \dfrac{1}{1 + j 2\delta Q_0}$. $\qquad \ldots(14)$

(iii) Bandwidth of the amplifier :

It can be seen that for transformer coupled amplifier equation (14) is the same as for single tuned direct coupled amplifier. Hence the bandwidth can be obtained in the similar manner.

Now we shall determine the upper and lower cut-off frequencies f_2 and f_1. At these frequencies $(\mathbf{A}/\mathbf{A}_{resonance}) = \dfrac{1}{\sqrt{2}}$.

$$\therefore \quad 4\delta^2 Q_0^2 = 1,$$

or $$\delta = \pm \tfrac{1}{2} \frac{1}{Q_0}.$$

The positive and negative values of δ correspond to the upper and the lower cut-off frequencies respectively so that

$$\delta = \frac{f_2 - f_0}{f_0} = \tfrac{1}{2} \frac{1}{Q_0}$$

or $$f_2 - f_0 = \frac{f_0}{2Q_0}$$

Similarly, considering the negative sign

$$f_0 - f_1 = \frac{f_0}{2Q_e}.$$

Thus the 3*db* bandwidth $= BW = f_2 - f_1 = f_2 - f_0 + f_0 - f_1$

$$= \frac{f_0}{2Q_3} + \frac{f_0}{2Q_e} = \frac{f_0}{Q_e}.$$

In case of pentode tubes used as tuned amplifier r_p is large and correspondingly $(M)_{optimum}$ should be large. To achieve large M, the distributed capacitance of the winding may become excessive, and the self resonant frequency may be so low that the coil may be of no use ; so the value of M is chosen below the optimum value.

(b) Tuned primary and tuned secondary radio frequency transformer-coupled amplifier (Band pass Amplifier) : In order to obtain high fidelity and high selectivity, the ideal response curve should have flat top and straight edges. The ideal response curve can be obtained by using two tank circuits tuned to the same frequency and coupled together. This circuit is called a band pass amplifier and is used as r.f. amplifier in superhet receiver. The circuit diagram is shown in fig. 6.

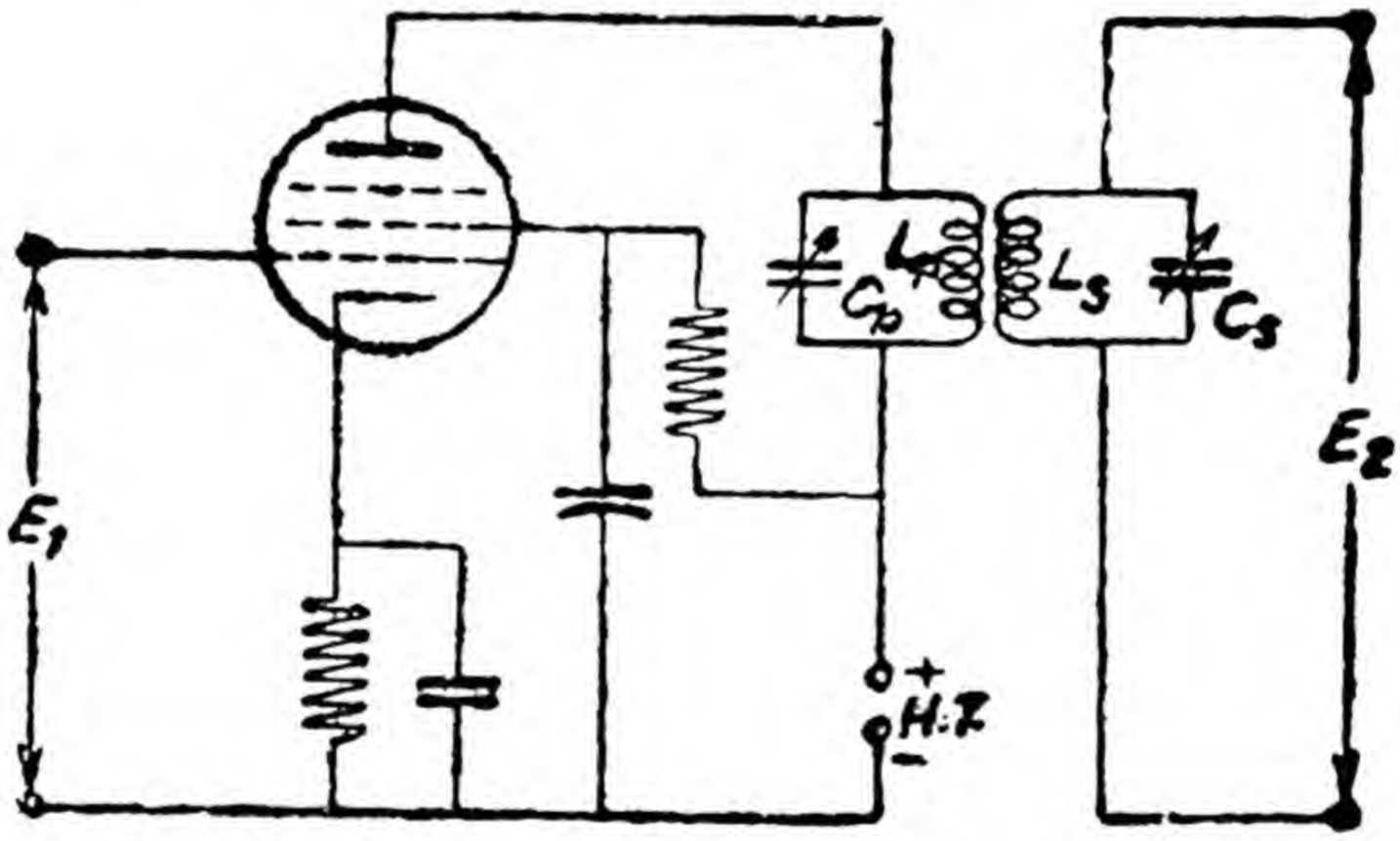


Fig. 6. Band Pass Amplifier.

(i) Voltage gain :

Voltage gain of this amplifier can be calculated as follows :

The equivalent circuit of the amplifier is given in figure 7 (*a*). C_P' is stray capacitance of the primary windings and C_P is the capacitance of the primary tuned circuit. The total shunt primary capacitance C_P'' will be

$$C_P'' = C_P' + C_P.$$

In calculations, we shall use only C_P and ignore C_P'.

The equivalent circuit of fig. 7 (*a*) can be simplified by applying Thevenin's theorem to the portion of the circuit lying left to the points *GF* of fig. 7 (a). The open circuit voltage of the equivalent generator is

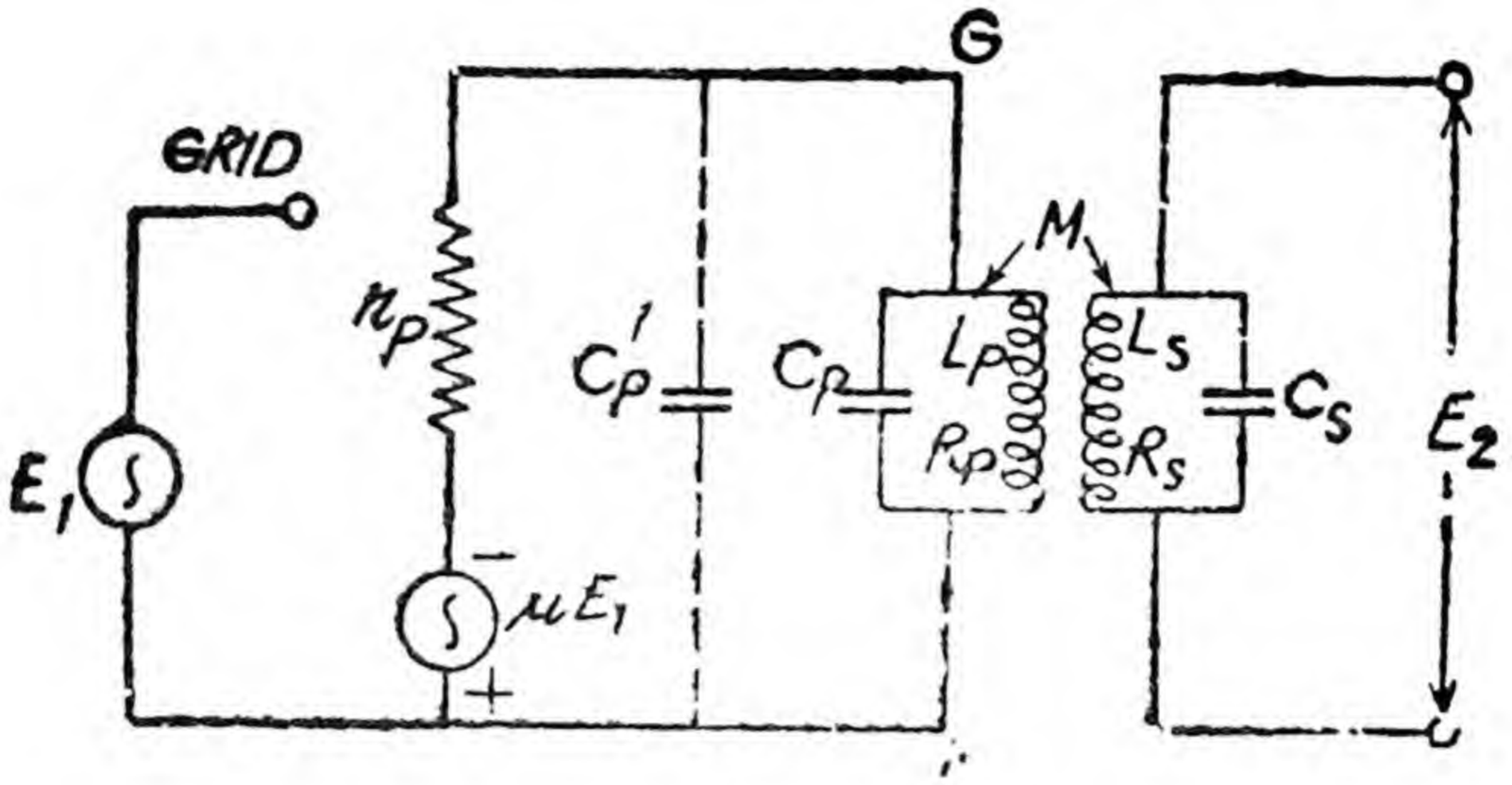


Fig. 7. (a) Equivalent circuit of double tuned amplifier.

$$\mathbf{E}_{oc}=\frac{\mu \mathbf{E}_1 \dfrac{1}{j\omega C_P}}{r_p+\dfrac{1}{j\omega C_P}}$$

But $r_p \gg \dfrac{1}{j\omega C_P}$ for pentode ; hence

$$\mathbf{E}_{oc}=\frac{\mu \mathbf{E}_1 \cdot \dfrac{1}{j\omega C_P}}{r_p}=\frac{g_m \mathbf{E}_1}{j\omega C_P}.$$

The internal impedance of the equivalent generator will be

$$\mathbf{Z}_i=\frac{r_p \cdot \dfrac{1}{j\omega C_P}}{r_p+\dfrac{1}{j\omega C_P}}=\frac{1}{j\omega C_P}$$

Thus with generator voltage E_{oc} and internal impedance $\mathbf{Z}_i$, the equivalent circuit becomes as shown in figure 7 (*b*).

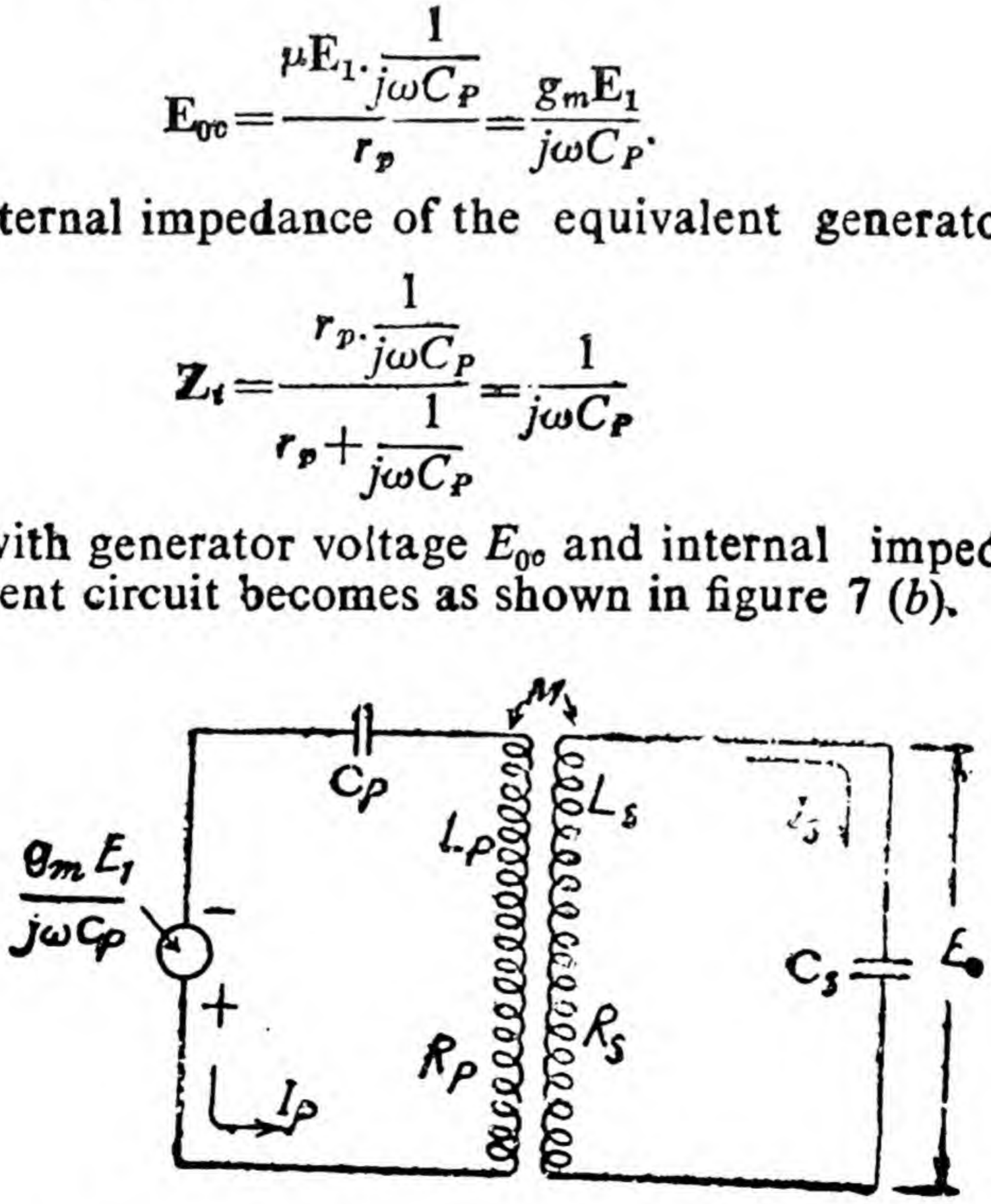


Fig. 7. (b) Simplified form of equivalent circuit of fig. 7 (a).

We know from network analysis (Chap. 3) that

$$\left.\begin{aligned} E &= Z_{11}I_1 + Z_{12}I_2, \\ 0 &= Z_{21}I_1 + Z_{22}I_2 \end{aligned}\right\}$$

In the present case ; this implies

$$\left.\begin{aligned} E_{0c} &= Z_{PP}I_P + Z_{PS}I_S, \\ 0 &= Z_{PS}I_P + Z_{SS}I_S. \end{aligned}\right\}$$

from which current in secondary will be

$$I_S = -\frac{E_{0c} \cdot Z_{PS}}{Z_{PP}Z_{SS} - Z^2_{PS}} \qquad \ldots(15)$$

Now,

$$\left.\begin{aligned} Z_{PP} &= R_P + j\left(\omega L_P - \frac{1}{\omega C_P}\right), \\ Z_{PS} &= j\omega M, \\ Z_{SS} &= R_S + j\left(\omega L_S - \frac{1}{\omega C_S}\right), \end{aligned}\right\}$$

Voltage gain of amplifier is given by

$$A = \frac{E_2}{E_1} = \frac{I_S \cdot \dfrac{1}{j\omega C_S}}{E_1}.$$

Putting the value of I_S from equation (15),

$$A = \frac{-E_{0c}\, Z_{PS}}{Z_{PP}Z_{SS} - Z^2_{PS}} \cdot \frac{\dfrac{1}{j\omega C_S}}{E_1}$$

Again substitute for E_{0c} and Z_{PS},

$$\therefore \quad A = \frac{-\dfrac{g_m E_1}{j\omega C_P}\, j\omega M \cdot \dfrac{1}{j\omega C_S}}{Z_{PP}\, Z_{SS} - Z^2_{PS}},$$

$$= \frac{\dfrac{g_m}{j\omega C_P}\, j\omega M \cdot \dfrac{1}{j\omega C_S}}{Z^2_{PS} - Z_{PP}\, Z_{SS}} \qquad \ldots(16)$$

Primary and secondary are tuned to the same frequency, *i.e.*

$$\omega_0 = \frac{1}{\sqrt{(L_P C_P)}} = \frac{1}{\sqrt{(L_S C_S)}}. \qquad \ldots(17)$$

Also

$$\left.\begin{aligned} Q_P &= \frac{\omega_0 L_P}{R_P} = \frac{1}{\omega_0 C_P R_P}, \\ Q_S &= \frac{\omega_0 L_S}{R_S} = \frac{1}{\omega_0 C_S R_S}. \end{aligned}\right\} \qquad \ldots(18)$$

And coupling dissipation factor a is given by

$$a^2 = K^2 Q_P Q_S = \frac{M^2}{L_P L_S} \cdot \frac{\omega_0 L_P}{R_P} \cdot \frac{\omega_0 L_S}{R_S} = \frac{M^2 \omega_0^2}{R_P R_S}$$

or $$a = K\sqrt{(Q_P\ Q_S)} = \frac{\omega_0 M}{\sqrt{(R_P R_S)}},$$

where K is the coefficient of coupling.

Further let δ be the drift in frequency expressed as a fraction of resonant frequency ω_0 *i.e.*

$$\delta = \frac{\omega - \omega_0}{\omega_0}$$

or $$\omega = \omega_0 (1+\delta).$$

Now the expression for Z_{PP} and Z_{SS} can be modified as

$$\left.\begin{aligned} Z_{PP} &= R_P + j\left(\omega L_P - \frac{1}{\omega C_P}\right) \\ &= R_P + j\left[\omega_0(1+\delta)\, L_P - \frac{1}{\omega_0 (1+\delta)\, C_P}\right] \\ &= R_P + j\left[Q_P R_P (1+\delta) - \frac{R_P Q_P}{(1+\delta)}\right] \\ &= R_P + jQ_P R_P [(1+\delta) - (1+\delta)^{-1}] \\ &= R_P (1 + j\, 2\delta Q_P). \\ \text{Similarly,}\quad Z_{SS} &= R_S (1 + j\, 2\delta Q_S). \\ \text{and}\quad Z_{PS} &= j\omega M. \end{aligned}\right\} \quad \text{...(19)}$$

Putting the values of Z_{PP}, Z_{PS} and Z_{SS} from equation (18) in equation (16), we get

$$A = \frac{\frac{g_m}{j\omega C_P} \cdot \frac{1}{j\omega C_S}\, j\omega M}{-\omega^2 M^2 - R_P R_S (1 + j.2\delta Q_P)(1 + j\, 2\delta Q_S)}$$

$$= \frac{j g_m \left(\frac{M}{\omega_0 C_P C_S}\right)}{\omega^2 M^2 + R_P R_S [1 + j.2\delta\, (Q_P + Q_S) - 4\delta^2 Q_P Q_S]}$$

For frequencies close to resonant frequency, we take

$$\omega \approx \omega_0.$$

$$A = \frac{j g_m \left(\frac{M}{\omega_0 C_P C_S}\right)}{\omega_0 M^2 + R_P R_S [1 + j.2\delta\, (Q_P + Q_S) - 4\delta^2 Q_P Q_S]}$$

$$= \frac{j.g_m \frac{M}{\omega_0 C_P C_S} \cdot \frac{1}{R_P R_S}}{\frac{\omega_0 M^2}{R_P + R_S} + [1 + j.2\delta (Q_P + Q_S) - 4\delta^2 Q_P Q_S]}$$

$$= \frac{j\, g_m \frac{\omega_0 M}{\sqrt{(R_P R_S)}} \cdot \frac{1}{\omega_0 C_P C_S} \cdot \frac{1}{\omega_0 R_P R_S} \cdot \sqrt{(R_P R_S)}}{1 + \frac{\omega_0^2 M^2}{R_P R_S} + j.2\delta\, (Q_P + Q_S) - 4\delta^2 Q_P Q_S}$$

$$= \frac{j.g_m.a.\dfrac{1}{\omega_0 C_P R_P}.\dfrac{1}{\omega_0 C_S R_S}\sqrt{(R_P R_S)}}{1+a^2+j.2\delta\,(Q_P+Q_S)-4\delta^2 Q_P Q_S}.$$

$$= \frac{j.g_m.a.Q_P Q_S\sqrt{(R_P R_S)}}{1+a^2+j.2\delta\,(Q_P+Q_S)-4\delta^2 Q_P Q_S}. \qquad ...(20)$$

The gain at resonance will be obtained by setting $\delta=0$.

$$\therefore \quad \mathbf{A}_{r\,eso.} = \frac{jg_m a\ Q_P Q_S\sqrt{(R_P R_S)}}{1+a^2}$$

The above expression is generally required in a very popular form obtained by putting

$a^2=K^2 Q_P Q_S$ in equation (20)

$$\mathbf{A}_{r\,eso.} = \frac{jg_m\ K\sqrt{(Q_P Q_S)}.Q_P Q_S.\sqrt{(R_P R_S)}}{1+K^2 Q_P Q_S}$$

$$= \frac{jg_m.K.\sqrt{\left(\dfrac{\omega_0 L_P}{R_P}\right)}\sqrt{\left(\dfrac{\omega_0 L_S}{R_S}\right)}.\sqrt{(R_P R_S)}}{K^2+\dfrac{1}{Q_P Q_S}}$$

$$= \frac{jg_m\ K\omega_0\sqrt{(L_P L_S)}}{K^2+\dfrac{1}{Q_P Q_S}}$$

$$[\mathbf{A}_{r\,eso.}] = \frac{g_m K\ 2\pi f_0.\sqrt{(L_P L_S)}}{K^2+\dfrac{1}{Q_P Q_S}} \qquad ...(21)$$

The expression shows that voltage amplification is dependent upon K, the coefficient of coupling. Therefore, when K is of such a value so as to transfer maximum energy, voltage amplification is maximum. This amount of coupling is called *critical coupling* and is expressed by relation

$$K_c = \frac{1}{\sqrt{(Q_P Q_S)}}.$$

Thus maximum amplification at resonant frequency will be

$$(\mathbf{A}_{r\,eso.})_{max.} = \frac{j.g_m.\dfrac{1}{(\sqrt{Q_P Q_S})}.2\pi f_0\sqrt{(L_P L_S)}}{\dfrac{1}{Q_P Q_S}+\dfrac{1}{Q_P Q_S}}$$

$$= j.g_m.\pi f_0.(L_P L_S Q_P Q_S)^{1/2}$$

$$= j\,[\tfrac{1}{2} g_m.\omega_0\sqrt{(L_P L_S Q_P Q_S)}]. \qquad ..(22)$$

It is shown that when primary and secondary circuit Q's were approximately the same, the width of the response band was

determined by the coefficient of coulping K according to the relation

$$\text{width of pass band} = \sqrt{2}Kf \text{ provided } Q_P = Q_S.$$

It is also found that when primary and secondary Q values are the same the response curve has the maximum possible single peak flatness in the vicinity of resonance. This is the position of critical coupling K_c. For $K > K_c$, response curve shows double humps. The phenomenon has been discussed in the chapter on resonant circuits.

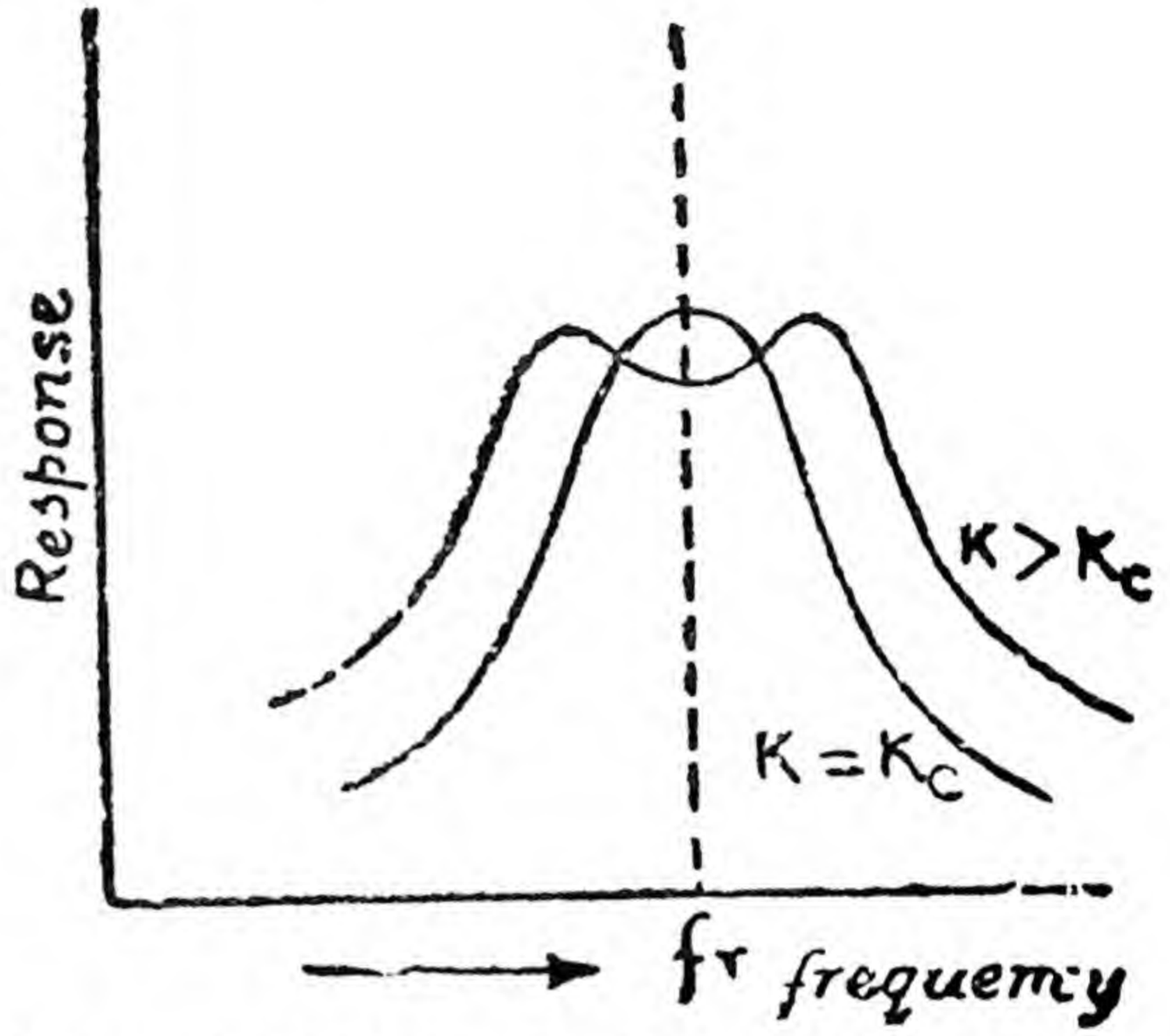


Fig. 8. Response curve of typical amplifier.

The important characteristics of tuned voltage amplifiers are :

(*i*) amplification at resonance.

(*ii*) variation of amplification with frequency in the immediate vicinity of resonance, and

(*iii*) discrimination against frequencies differing appreciably from resonance.

(ii) Maximum voltage gain of over coupled double tuned amplifier :

If $a > 1$, the circuit is said to be overcoupled. In this case the positions of the peaks of the response curve can be determined by differentiation of the square of the absolute value of A and equating it to zero. The squa reof the absolute value of A is taken since the gain is maximum without any consideration to the phase. Hence from equation (20)

$$|A|^2 = [g_m Q_P Q_S \sqrt{(R_P R_S)}]^2 \frac{a^2}{[2\delta\,(Q_P + Q_S)]^2 + [1 + a^2 - 4\delta^2 Q_P Q_S]^2}$$

Equating $\frac{d\,|A|^2}{d\delta}$ to zero, we get

$$\frac{d}{d\delta}\left[(1 + a^2 - 4\delta^2 Q_P Q_S)^2 + \{2\delta\,(Q_P + Q_S)\}^2\right] = 0$$

or $$2\{1 + a^2 - 4\delta^2 Q_P Q_S\}\,(-4Q_P Q_S \times 2\delta) + 2\,\{2\delta(Q_P + Q_S)\}\,\{2\,(Q_P + Q_S)\} = 0$$

$$-16 Q_P Q_S \delta\,\{1 + a^2 - 4\delta^2 R_P Q_S\} + 8\delta\,(Q_P + Q_S)^2 = 0$$

$$2Q_P Q_S\{1 + a^2 - 4\delta^2 Q_P Q_S\} = (Q_P + Q_S)^2$$

or $$\{1+a^2-4\delta^2 Q_P Q_S\}=\frac{\{Q_P+Q_S\}^2}{2Q_P Q_S} \qquad \ldots(23)$$

Rearranging equation (23), we have

$$\delta^2=\frac{1+a^2}{4Q_P Q_S}-\frac{(Q_P+Q_S)^2}{8Q_P^2 Q_S^2}$$

$$\therefore \quad \delta=\pm\tfrac{1}{2}\sqrt{\left\{\frac{1+a^2}{Q_P Q_S}-\frac{(Q_P+Q_S)^2}{2\,(Q_P Q_S)^2}\right\}}$$

$$=\pm\tfrac{1}{2}\sqrt{\left\{\frac{1}{Q_P Q_S}+K^2-\tfrac{1}{2}\left(\frac{Q_P+Q_S}{Q_P Q_S}\right)^2\right\}}$$

Usually Q_P is kept equal to Q_S or very close to Q_S so that $\sqrt{(Q_P Q_S)}=(Q_P+Q_S)/2$.

$$\therefore\ \delta=\pm\tfrac{1}{2}\sqrt{\left\{K^2-\frac{1}{Q_P Q_S}\right\}}$$

or $$\delta\sqrt{(Q_P Q_S)}=\pm\tfrac{1}{2}\sqrt{(K^2 Q_P Q_S-1)}=\pm\tfrac{1}{2}\sqrt{(a^2-1)}. \qquad \ldots(24)$$

Substituting the value of δ from equation (24), in equation (20), we get

$$\mathbf{A}_{max.}=j\,g_m Q_P Q_S\sqrt{(R_P R_S)}\,\frac{a}{1+a^2+2j\,\sqrt{(a^2-1)}-(a^2-1)}$$

$$=j\,g_m Q_P Q_S\sqrt{(R_P R_S)}\,\frac{a}{2\,[1+j\,\sqrt{(a^2-1)}]}$$

The numerical value of $\mathbf{A}_{max.}$ is given by

$$\mathbf{A}_{max.}=\tfrac{1}{2}g_m Q_P Q_S\sqrt{(R_P R_S)} \qquad \ldots(25)$$

(iii) Bandwidth of double tuned amplifier :

The bandwith of a double tuned amplifier when $a=1$ (optimum condition) and with $Q_P=Q_S=Q$ may be calcuiated from

$$\frac{\mathbf{A}}{A_{reso.}}=\frac{1}{1-2\delta^2 Q^2+2j\delta Q}. \qquad \ldots(26)$$

The magnitude of $(\mathbf{A}/\mathbf{A}_{reso.})$ is given by

$$\frac{A}{A_{reso.}}=\frac{1}{\sqrt{\{(1-2\delta^2 Q^2)^2+4\delta^2 Q^2\}}}$$

$$=\frac{1}{\sqrt{(1+4\delta^4 Q^4)}}. \qquad \ldots(27)$$

The bandwidth is defined as the width of a band of frequencies over which power amplification does not fall below $1/\sqrt{2}$ of the maximum amplification. The fraction of $1/\sqrt{2}$ corresponds to $3db$ loss in amplification, hence it is also called $3db$ power bandwidth. If f_1 and f_2 be the two frequencies for which the amplification is $1/\sqrt{2}$ of maximum value, then $3db$ bandwidth

$$B=f_2-f_1.$$

For frequency range over which the gain does not fall below $1/\sqrt{2}$ of its maximum value, we have

$$4\delta^4 Q^4 = 1 \quad \text{or} \quad \delta = \pm \frac{1}{\sqrt{2}} \cdot \frac{1}{Q}$$

Thus $$f_0 - f_1 = \frac{1}{\sqrt{2}} \frac{f_0}{Q}$$

and $$f_2 - f_0 = \frac{1}{\sqrt{2}} \frac{f_0}{Q}$$

Thus the bandwith $= f_2 - f_1 = \frac{2}{\sqrt{2}} \frac{f_0}{Q} = \sqrt{2} \frac{f_0}{\sqrt{(Q_P Q_S)}}$.

In this way by introducing two or more tuned circuits. an improvement in response curve is available as compared to signal tuned circuit.

13·6. GAIN BANDWIDTH PRODUCT OF TUNED AMPLIFIERS :

In case of tuned amplifier, voltage gain at resonance are given by the following expressions :

Single tuned capacitance coupled amplifier : $A_{res.} = g_m \omega_0 L Q_e$

Single tuned transformer coupled amplifier : $A_{res.} = g_m \omega_0 M Q_e$

Double tuned amplifier : $A_{reso.} = j \frac{1}{2} g_m Q_P Q_S \sqrt{(R_P R_S)}$

The bandwidth of these amplifiers are also given below :

Single tuned capacitance coupled amplifier : $B = \frac{f_0}{Q_e}$

Single tuned transformer coupled amplifier : $B = \frac{f_0}{Q_e}$

Double tuned amplifier $$B = \sqrt{2} \frac{f_0}{\sqrt{(Q_P Q_S)}}.$$

The bandwidth product of these amplifiers can be calculated by multiplying the voltage gain and bandwidth. Now we shall calculate the bandwidth product in different cases :

(i) Single tuned capacitance coupled amplifier :

$$A_{reso.}\ B = g_m \omega_0 L Q_e \times (f_0/Q_e) = g_m \omega_0 L f_0$$

$$= \frac{g_m \omega_0^2 L}{2\pi}$$

We know that $\omega_0^2 = (1/LC)$ or $\omega_0^2 L = (1/C)$

$$\therefore \quad A_{res.}\ B = \frac{g_m}{2\pi C}.$$

In order to increase the bandwidth product, the value of C should be reduced. In the lowest limit, the capacity C is reduced of the interelectrode capacity, $(C_{in} + C_{out})$ of the tube and the limiting gain bandwidth product is

$$A_{res.}\ B = \frac{g_m}{2\pi\ (C_{in} + C_{out})},$$

where C_{out} is the output capacitance of the amplifier and C_{in} is input capacitance of the next stage.

(ii) Single tuned transformer coupled amplifier :

The gain bandwidth product in this case is given by

$$A_{res.}\ B = g_m \omega_0 M Q_e \times (f_0/Q_e) = g_m \omega_0 M f_0$$

$$= \frac{g_m \omega_0^2 M}{2\pi}$$

Again $$a = \frac{M}{\sqrt{(L_P L_S)}} = \frac{M}{L} = 1 \quad \text{or} \quad M = L$$

Thus $A_{res.}\ B = \frac{\omega_0^2 L g_m}{2\pi} = \frac{g_m}{2\pi C}$.

In the limiting case $C = C_{in} + C_{out}$.

$$\therefore \quad A_{res.}\ B = \frac{g_m}{2\pi\ (C_{in} + C_{out})}$$

(iii) Double tuned amplifier :

The gain band width product of double tuned amplifier with parameter $a = 1$ and $Q_P = Q_S = Q$ is given by

$$A_{res.}\ B = \tfrac{1}{2} g_m\ Q_P Q_S \sqrt{(R_S R_P)} \times \sqrt{2} f_0 / \sqrt{(Q_P Q_S)}$$

$$= \frac{1}{\sqrt{2}}\ g_m \sqrt{(Q_P Q_S R_p R_S)} \cdot (\omega_0/2\pi)$$

$$= \frac{1}{\sqrt{2}} \frac{g_m Q R \omega_0}{2\pi}$$

$$= \frac{1}{2\sqrt{2}} \frac{g_m \omega_0^2 L}{\pi} \text{ as } QR = \omega_0 L$$

$$= \frac{g_m}{2\sqrt{2\pi C}} \qquad \omega_0^2 L = 1/C$$

$$= \frac{g_m}{2\sqrt{2\pi\ (C_{out} + C_{in})}}$$

This expression shows that the gain bandwidth product of a double tuned amplifier is $\sqrt{2}$ times that for a singie tuned circuit.

13·7. CASCADED TUNED AMPLIFIER :

Sometimes we are interested in high gain of the amplifier then several identical stages of the tuned amplifiers are used. The over-all voltage gain is the product of the voltage gains of the individual stages. At the same time the high voltage gain is accompanied by a narrower bandwidth than for a single stage. The decrease in bandwidth can be determined as follows :

Case (i) n stage of a single tuned amplifier : Let us consider that n identical stages of a single tuned direct coupled amplifier are connected in cascade. Our problem is to determine the overall gain and bandwidth of such an amplifier.

We know that the relative gain of a single tuned amplifier with respect to the gain at resonant frequency f_0 is given by the expression

$$\left(\frac{A}{A_{res}}\right)=\frac{1}{\sqrt{[1+(2\delta Q_e)^2]}} \qquad \ldots(1)$$

Now the gain of n stage cascade amplifier becomes

$$\left(\frac{A}{A_{res}}\right)^n=\left[\frac{1}{\sqrt{[1+(2\delta Q_e)^2}}\right]^n=\frac{1}{[1+(2\delta Q_e)^2]^{n/2}} \qquad \ldots(2)$$

When the relative gain ($A/A_{res.}$) is put to $1/\sqrt{2}$, the factor δ in equation (2) corresponds to two extreme frequencies. Hence in order to determine the bandwidth, the denominator should be equal to $\sqrt{2}$.

$$\therefore \quad [1+(2\delta Q_e)^2]^{n/2}=\sqrt{2}$$

or
$$1+(2\delta Q_e)^2=2^{1/n} \text{ or } 2\delta Q_e=\pm(2^{1/n}-1)^{1/2} \qquad \ldots(3)$$

Substituting the value of δ in equation (3), we have

$$\frac{2(f-f_0)}{f_0}\,Q_e=\pm(2^{1/n}-1)^{1/2}.$$

Now $f_0-f_1=\frac{f_0}{2Q_e}\sqrt{(2^{1/n}-1)}$

and $f_2-f_0=\frac{f_0}{2Q_e}\sqrt{(2^{1/n}-1)}.$

Thus the bandwidth of an n-stage identical amplifier is

$$B_{1n}=f_2-f_1=f_2-f_0+f_0-f_1$$
$$=\frac{f_0}{2Q_e}\sqrt{(2^{1/n}-1)}+\frac{f_0}{2Q_e}\sqrt{(2^{1/n}-1)}$$
$$=\frac{f_0}{Q_e}\sqrt{(2^{1/n}-1)}=B_1\sqrt{(2^{1/n}-1)},$$

where B_{1n} is the bandwidth of n stages cascade amplifier and B_1 is the bandwidth for single stage. Equation (4) shows that the bandwidth of n stages cascade amplifier is multiplied by a factor $\sqrt{(2^{1/n}-1)}$. When $n=2$, the factor $\sqrt{(2^{1/n}-1)}=0{\cdot}643$ and for $n=3$, its values is $0{\cdot}510$. Thus the bandwidth is reduced to $64{\cdot}3\%$ for 2 stages of cascade amplifier and 51% for stages amplifier. In order to maintain a prescribed 3 *db* bandwidth, it is necessary to reduce the Q of the tuned circuits continuously as the stages are increased. ...(4)

Case (ii) n-stages of double tuned amplifier : Let there be n stages in a double tuned amplifier with critical coupling $a=1$ and Q_e is effective Q of the tuned circuit in each stage. The relative voltage gain of n stages of double tuned amplifier will be

$$\left(\frac{A}{A_{res.}}\right)^n=\frac{1}{[1+4\delta^4Q^4]^{n/2}}. \qquad \ldots(5)$$

To find the corresponding bandwidth, it is noted that

$$(1+4\delta^4Q^4)^{n/2}=\sqrt{2}$$

or
$$\delta Q=\pm\sqrt[4]{\left(\frac{2^{1/n}-1}{4}\right)}$$

or $$\frac{f-f_0}{f_0}=\pm\frac{1}{2Q}\sqrt[4]{(2^{1/n}-1)}.$$

The negative sign corresponds to the lower cut off frequency f_1, and the positive sign corresponds to the upper cut off frequency f_2.

$$\therefore \quad f_0-f_1=\frac{f_0}{2Q}\sqrt[4]{(2^{1/n}-1)}$$

and $$f_2-f_0=\frac{f_0}{2Q}\sqrt[4]{(2^{1/n}-1)}.$$

Thus the bandwidth of n identical stage double tuned amplifier is

$$B_{2n}=f_2-f_1=\frac{f_0}{Q}\sqrt[4]{(2^{1/n}-1)}=B_2\sqrt[4]{(2^{1/n}-1)}, \qquad ...(7)$$

where B_{2n} is the bandwidth of n stage cascade double tuned amplifier and B_2 is the bandwidth for each single stage. Equation (7) shows that the bandwidth is reduced by the factor $\sqrt[4]{(2^{1/n}-1)}$. The bandwidth reduction factor is tabulated in the following table :

TABLE

Number of stag n	Bandwidth reduction factor $\sqrt[4]{(2^{1/n}-1)}$
1	1
2	0·802
3	0·713
4	0·659
5	0·622

Thus for a two stage double tuned amplifier, 3*db* bandwidth reduces to 80·2% of the value of one stage. This reduction is smalier than the corresponding reduction in single tuned amplifier.

13·8. STAGGERED TUNED AMPLIFIERS :

In order to have high gain wideband amplification one may use a single tuned or a double tuned amplifier with a coil having a small value of Q. The desired gain can be obtained by the use of several cascaded stages. For a specified gain bandwidth product, double tuned cascade stages are preferred in comparison to the single tuned one. But it is very difficult to make alignment of the double tuned circuits. The alternate method to obtain a large bandwidth and other desirable characteristics of the double tuned circuits by a single tuned circuit is the *staggered tuned circuits*. The method is as follows :

In staggered tuned circuits, two single tuned cascade amplifiers having a certain bandwidth are taken. The resonant frequencies of the each tuned circuit of each stage are so adjusted that they are separated by an amount equal to the bandwidth of each

stage. Since the resonant frequencies are displaced or staggered, they are known as staggered tuned circuits. Figure 9 shows the combined gain characteristics of the two stages.

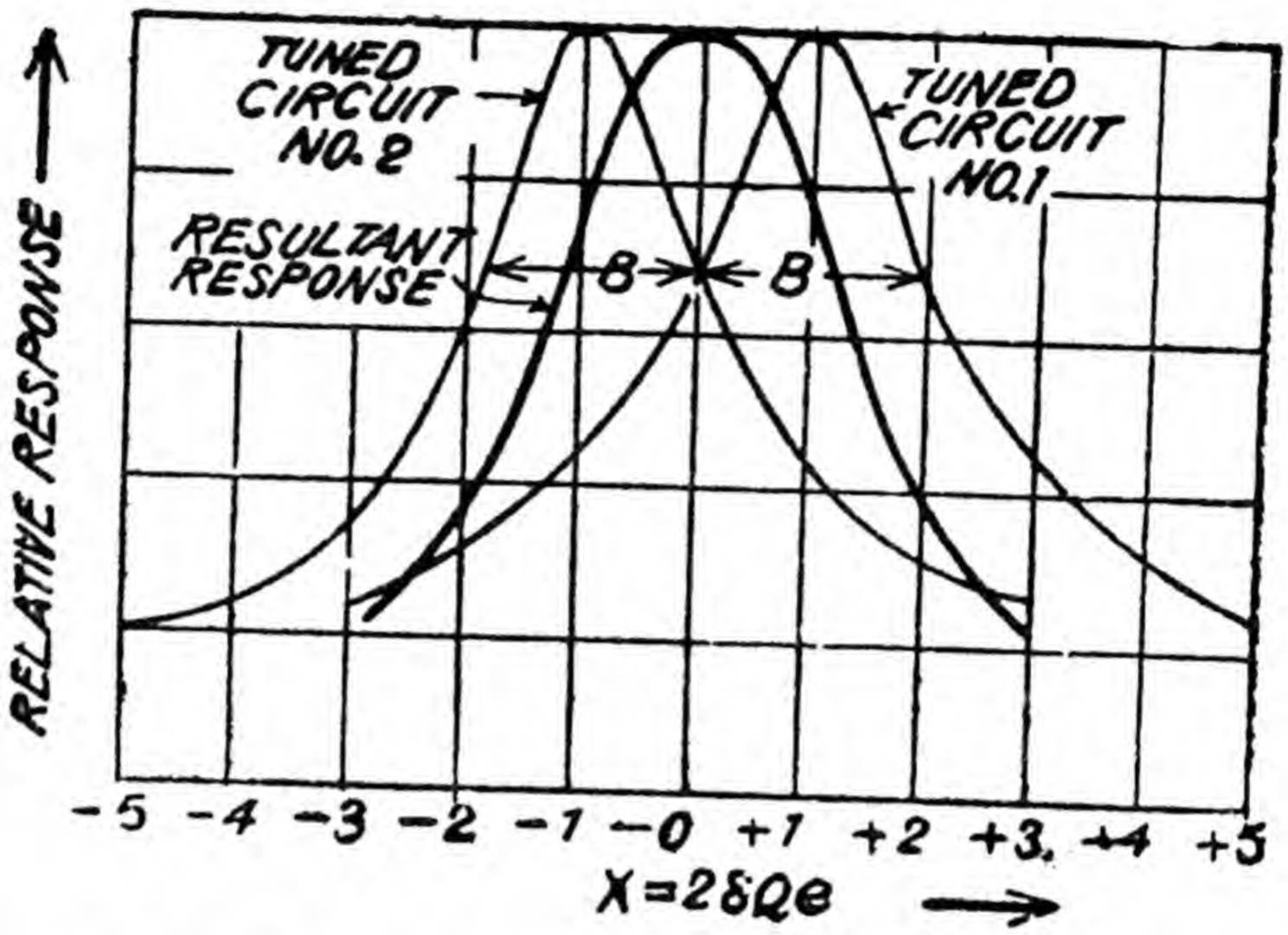


Fig. 9. The adjustment of frequency of a staggered tuned pair.

The resultant staggered pair will have a bandwidth that is $\sqrt{2}$ times as great as that of each of the individual single tuned circuits making up the pair. The overall selectivity function will be identical in form with that of a single stage double tuned system. Now we shall obtain the expression for the overall response characteristic of the staggered tuned pair. The relative gain of a single tuned direct coupled amplifier is given by the expression

$$\frac{A}{A_{res.}} = \frac{1}{1+j\,2\delta Q_e} = \frac{1}{1+jx}, \qquad \ldots(1)$$

where $2\delta Q_e = x.$

The bandwidth between the 70% relative gain points is

$$B = 2\delta f_0 = \frac{f_0}{Q_e}. \qquad \ldots(2)$$

Since one stage is tuned to a frequency $\delta_0 f_0$ below f_0 and the other stage is tuned to a frequency $\delta_0 f_0$ above f_0, the corresponding selectivity functions of the circuits are

$$\left(\frac{A}{A_{res.}}\right)_1 = \frac{1}{1+j\,(x-1)}, \qquad \ldots(3)$$

and

$$\left(\frac{A}{A_{res.}}\right)_2 = \frac{1}{1+j\,(x+1)}. \qquad \ldots(4)$$

By multiplying the relative gains of the two amplifiers, the overall gain function becomes

$$\left(\frac{A}{A_{res}}\right)_{pair} = \left(\frac{A}{A_{res}}\right)_1 \left(\frac{A}{A_{res}}\right)_2$$

$$= \frac{1}{2 - x^2 + 2jx} \qquad ..(5)$$

The magnitude of the resulting function is

$$\left(\frac{A}{A_{res}}\right)_1 \left(\frac{A}{A_{res}}\right)_2 = \frac{1}{\sqrt{(4+x^4)}} = \frac{1}{2} \frac{1}{\sqrt{(1+4\delta_0^4 Q^4)}}, \qquad ...(6)$$

where δ_0 is the value of δ referred to the new frequency ω_0 and Q is the value of Q_c for each circuit referred to ω_0. The comparison of equation (6) with double tuned circuit shows that the forms of variation are identical.

13·9. TUNED POWER AMPLIFIERS :

We have seen that low frequency or high frequency voltage amplification is always done under class *A* because the amplified output is free from distortion. When power output is the main consideration, a tuned power amplifier is used. The tuned power amplifier converts the d.c. power from the plate supply source into r.f. power. Thus the amplifier circuit should be designed in such a manner that the conversion of d.c. power into r.f power may take place with maximum efficiency. The most efficient condition is the class *C* operation in which the grid is biased beyond cut-off. When a sinusoidal signal is applied between grid and cathode, the plate current flows only for that part of the signal voltage cycle which makes the grid voltage less than the cut-off value. Now the plate current consists of a series of pulses The output voltage has no resemblance with input signal voltage and if the anode load is a resistor, the output wave is gross distorted. However, if the anode load is a parallel tuned circuit, tuned to the signal frequency, the output voltage is still sinusoidal even though the plate current is nonsinusoidal. This can be explained as follows :

The current pulse from supply voltage source E_{bb} flowing into the tuned circuit are periodic and non sinusoidal. It can be shown by Fourier series that they consists of fundamental and a large number of harmonics. The tuned circuit offers large impedance to the fundamental component of the current and hence a large output voltage is developed across its terminals. At the same time, the tuned circuit offers a very low impedance at harmonic frequencies and thus there will be negligible output at these frequencies. Thus a signal of single frequency or a narrow band of frequencies can be amplified from a lower level to a high power level by a tuned class *C* amplifier with small amount of distortion.

The tuned power amplifier can also be operated under class *B* which has some advantages over class *C*. Firstly. there are smaller harmonic contents because the plate current pulses are exactly half-sinusoids for a sinusoidal input voltage. Secondly, the output voltage is proportional to the magnitude of the input signal because class *B* is a linear operation. Due to this fact a class *B* tuned

amplifier is used to amplify signals of varying amplitude such as modulated signals. Under class *B* operation the envelope wave shapes of plate current and input voltage are identical while under class *C* operation, the envelope waveform of the resulting plate current does not correspond to that of input signal and hence very much distorted.

13·10. BASIC CIRCUIT OF TUNED POWER AMPLIFIER :

The circuit arrangement of tuned power amplifier is shown in fig. (10). The circuit arrangement is the same as that o a single tuned direct coupled voltage amplifier. The essential differences are the magnitude of grid supply voltage E_{cc}, signal voltage e_g, and the amount of power involved.

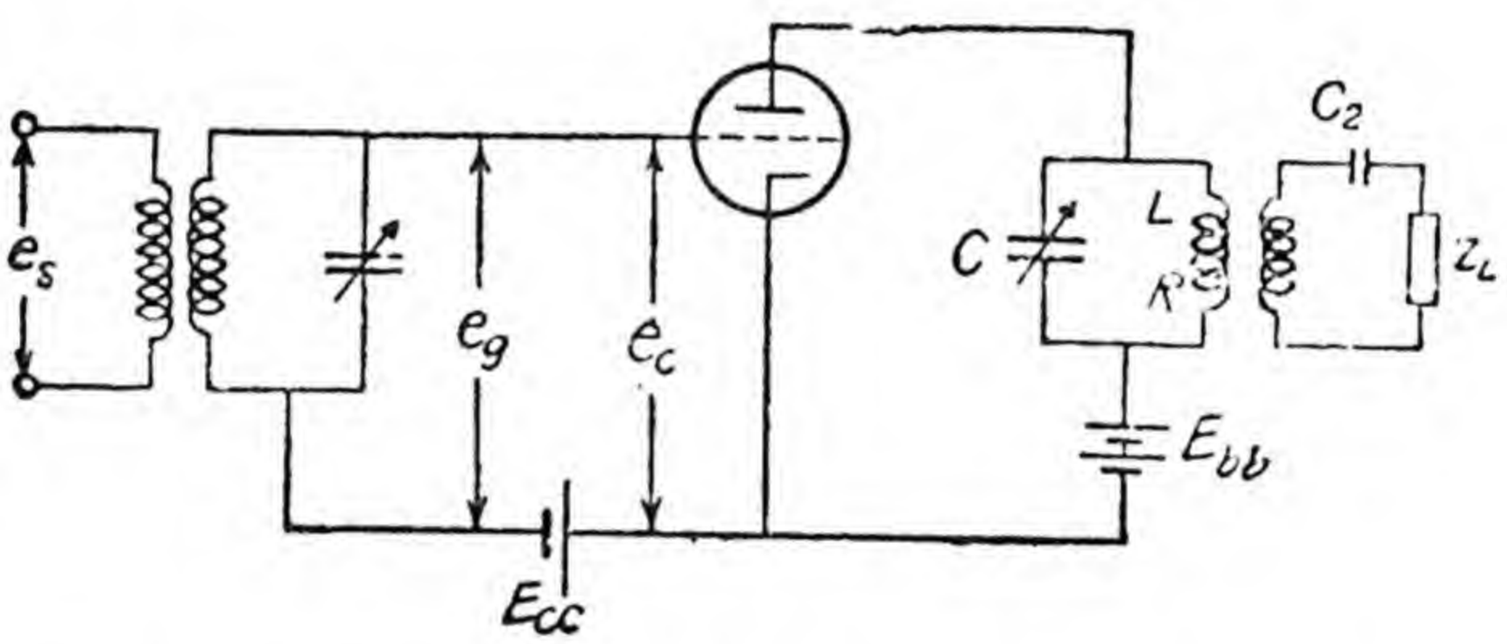


Fig 10. Basic circuit o a tuned power amplifier.

Under class *C* operation, the valve has to handle a large power hence triodes are preferred than pentodes. The plate can be kept below safe temperature by a water cooling system (in high power transmitters) whereas the temperature of screen grid of a pentode can not be so limited. Therefore, the size and rating of a pentode is limited by heating of the screen. The triodes are free from this limitation. (For low power amplifiers tetrodes and pentodes may be used). Triodes have a drawback of a large grid to plate capacitance (C_{gp}) through which unwanted feed back occurs but this can be eliminated by neutralising circuits.

In the cathode circuit a fixed bias battery E_{cc} is used instead of R_kC_k combination for self biasing. This is due to the fact that no d c. plate current flows and so no bias voltage can be developed if R_kC_k combination is used. The grid leak bias which develops only when the input signal is present can be used but it is also unsuitable because under quiescent condition the grid bias would be large enough to damage the tube. In actual practice a combination of fixed bias and grid leak bias is used for satisfactory operation. A combination of fixed bias and cathode bias is also not used because of losses that may occur in biasing resistor R_k due to heavy plate current.

The input signal e_s is applied to the grid through a tuned

circuit, tuned to the signal frequency. The load Z_L is inductively coupled to the plate tuned circuit. It Z_L is purely resistive the effect of coupling would be to increase R and reduce the Q factor of the tuned circuit. If the Q of the tuned plate circuit is 10 or more, the impedance of the tank circuit to the second or higher harmonics will be very low. Thus the effect of harmonic generation in the tube plate current is suppressed by the tuned plate load.

The amplifier can be operated under class B or class C by adjusting the bias E_{cc} to plate current cut-off or beyond plate current cut off respectively. Class B amplifier is used to amplify a narrow band of frequencies of differing amplitudes while class C amplifier is confined to narrow band of frequencies of constant amplitudes.

13·11. ANALYTICAL ANALYSIS OF CLASS B TUNED AMPLIFIER :

Under class B operation, the grid bias voltage is adjusted to cut off value so that no plate current flows with zero signal voltage while plate current flows for positive half of the a.c. signal applied. These amplifiers are used in radio transmitter for power amplification of modulated carrier. The analysis of the amplifier involves the determination of factors like d.c. power fed from the supply source, the a.c. power output, efficiency, grid driving potential, the shape of the plate current wave and the corresponding harmonic components of the plate current etc. with permissible limits of plate dissipation, grid dissipation and other limiting factors. The analysis is based on the following simplifying assumptions :

(*i*) triode characteristics are linear (this is most important, without this circuit linear method of analysis is invalid).

(*ii*) tuned plate circuit is tuned to the input signal frequency.

(*iii*) grid input signal is sinusoidal.

(*iv*) the bias is such that the plate current is zero for zero signal voltage.

(*v*) Q of the tuned circuit is high enough so that the impedance of the tank circuit is negligible at harmonic frequencies.

For the analysis, following specific procedure is followed :

(i) Representation of tube characteristic in analytical form :

We know that plate current is a function of plate potential and grid potential *i.e.*,

$$i_b = f(e_c\ e_b).$$

The above expression does not give a direct relationship as function f is unknown. The actual rela ionship can be derived by

Taylor's series expansion and is given by

$$i_b=k\left(e_c+\frac{e_b}{\mu}\right)^n, \text{ where } k \text{ and } \mu \text{ are constants.}$$

As the plate characteristics are straight line and equally spaced, $n=1$ and hence

$$i_b=k\left(e_c+\frac{e_b}{\mu}\right).$$

i_b=instantaneous total plate current,
e_c=instantaneous total grid voltage,
e_b=instantaneous total plate voltage.

The partial differential of i_b with respect to e_c is

$$\frac{\partial i_b}{\partial e_c}=k$$

However, $\frac{\partial i_b}{\partial e_c}=\left(\frac{\Delta i_b}{\Delta e_c}\right)_{e_b=\text{const.}}=g_m$

$$\therefore \quad i_b=g_m\left(e_c+\frac{e_b}{\mu}\right) \text{ provided } i_b \geqslant 0 \qquad \ldots(1)$$

This is the first term of Taylor expansion for current. Equation (1) represents the triode charcteristics in terms of the equivalent diode voltage $\left(e_c+\frac{e_b}{\mu}\right)$. If i_b is plotted against $\left(e_c+\frac{e_b}{\mu}\right)$ a straight line with slope g_m is obtained.

***(ii)* Cut-off bias :** Under the cut-off condition ($i_b=0$), if E_{bb} and E_{cc} are plate supply voltage and grid bias voltage respectively then $e_b=E_{bb}$ and $e_c=E_{cc}$.. From equation (1), we have

$$0=g_m\left(E_{co}+\frac{E_{bb}}{\mu}\right)$$

i.e., the cut-off bias $E_{co}=-\frac{E_{bb}}{\mu}$. ...(2)

***(iii)* Equivalent circuit :** Since Q of the tuned circuit is large, its impedance to harmonic frequencies of the grid signal is assumed to be zero. Thus for a sinusoidal grid signal the plate voltage will also be sinusoidal and 180° out of phase with grid voltage. Now the instantaneous potentials can be written as

$$\begin{aligned} e_g &= E_{gm}\cos\omega t \\ e_c &= E_{co}+E_{gm}\cos\omega t \\ e_b &= E_{bb}-E_{pm}\cos\omega t \end{aligned} \qquad \ldots(3)$$

Substituting these values in equation (1), we get

$$i_b=g_m\left(E_{co}+E_{gm}\cos\omega t+\frac{E_{bb}}{\mu}-\frac{E_{pm}}{\mu}\cos\omega t\right)$$

$$=g_m\left(E_{gm}\cos\omega t-\frac{E_{pm}}{\mu}\cos\omega t\right)\text{ as }\left(E_{co}+\frac{E_{bb}}{\mu}\right)=0$$

$$=g_m\left(E_{gm}-\frac{E_{pm}}{\mu}\right)\cos\omega t$$

$$=I_{bm}\cos\omega t \qquad \ldots(4)$$

The equation (4) for the current is valid when ωt is lying in the first and fourth quadrant. For other values $i_b=0$. Hence,

$$i_b=I_{bm}\cos\omega t \qquad -\pi/2<\omega t<\pi/2$$

$$i_b=0 \qquad \pi/2<\omega t<3\pi/2 \qquad \ldots(5)$$

where $$I_{bm}=g_m\left(E_{gm}-\frac{E_{pm}}{\mu}\right)$$

The average value of plate current is

$$(I_b)_{av.}=\frac{1}{2\pi}\int_0^{2\pi} i_b\,d(\omega t)=\int_{-\pi/2}^{+\pi/2} I_{bm}\cos\omega t\,d(\omega t)$$

$$=\frac{I_{bm}}{\pi} \qquad \ldots(6)$$

The fundamental component of plate current is

$$I_{p1m}=\frac{1}{\pi}\int_0^{2\pi} i_b\cos\omega t\,d(\omega t)$$

$$=\frac{2}{\pi}\int_0^{\pi/2} I_{bm}\cos^2\omega t\,d(\omega t)=\frac{I_{bm}}{2} \qquad \ldots(7)$$

At resonance $Z(\omega_0)=R_0$ is resistive and the output voltage at fundamental frequency is

$$E_{p1m}=I_{p1m}.R_0=\frac{I_{bm}}{2}R_0$$

$$=\frac{R_0}{2}g_m\left(E_{gm}-\frac{E_{p1m}}{\mu}\right)$$

or $$E_{p1m}+\frac{R_0}{2}g_m\frac{E_{p1m}}{\mu}=\frac{R_0}{2}g_m E_{gm}$$

$$E_{p1m}\left(1+\frac{R_0}{2r_p}\right)=\frac{R_0}{2}g_m E_{gm}$$

$$\therefore\ E_{p1m}=R_0\frac{\mu E_{gm}}{R_0+2r_p}$$

or $$I_{p1m}=\frac{\mu E_{gm}}{R_0+2r_p} \qquad \ldots(8)$$

The r.m.s. value of the fundamental component current is

$$I_{p1}=\frac{\mu E_g}{R_0+2r_p} \qquad \ldots(9)$$

The equivalent a.c. circuit for the fundamental component of the linear class *B* tuned amplifier as suggested by equation (9) is shown in fig. (11). This circuit is only applicable for the calculation of fundamental component of the plate current of angular frequency ω_0. The voltage gain of the amplifier is

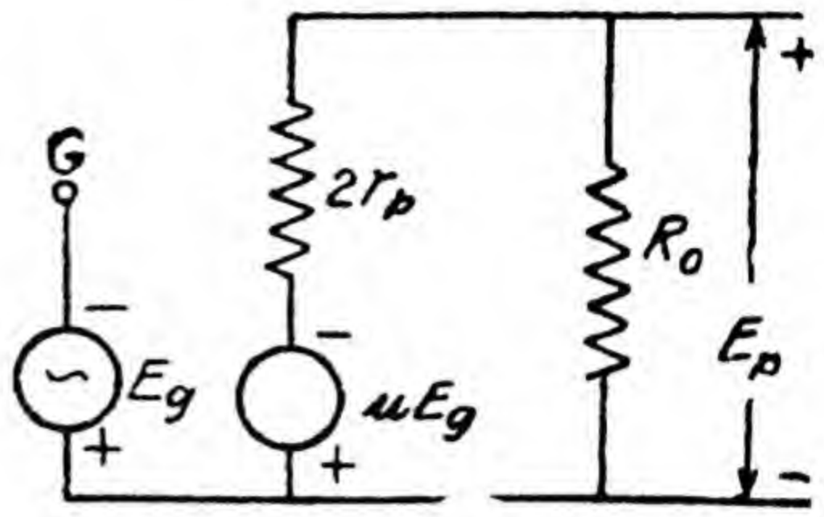


Fig. 11. Equivalent circuit of Class B tuned amplifier for fundamental angular freq. ω_0.

$$A=\frac{E_p}{E_g}=-\frac{\mu R_0}{2r_p+R_0} \qquad \ldots(10)$$

(*iv*) **Power output and efficiency :** The power delivered to the tuned circuit is

$$P_{ac}=I_{p1}^2 R_0=\left(\frac{\mu E_g}{R_0+2r_p}\right)^2 R_0 \qquad \ldots(11)$$

and power supplied by the plate voltage supply source E_{bb}, is,

$$P_{bb}=E_{bb}.I_b=E_{bb}\,\frac{I_{pm}}{\pi}=E_{bb}.\,\frac{2I_{p1m}}{\pi} \qquad \because \quad I_{bm}=2\,I_{p1m}$$

$$=E_{bb}.\,\frac{2}{\pi}\left(\frac{\mu E_{gm}}{R_0+2r_p}\right) \qquad \ldots(12)$$

Plate circuit efficiency $\eta_p=\dfrac{P_{ac}}{P_{bb}}$

$$=\frac{E_p.I_{p1}}{E_{bb}.I_b}=\frac{E_p.I_{p1}}{E_{bb}.\left(\dfrac{I_{bm}}{\pi}\right)}$$

$$=\frac{E_p.I_{p1}.\pi}{E_{bb}.2I_{p1m}}=\frac{E_p.I_{p1}.\pi}{E_{bb}(2\sqrt{2}\,I_{p1})}$$

$$=\frac{E_p}{E_{bb}}.\,\frac{\pi}{2\sqrt{2}}=\frac{E_{pm}}{E_{bb}}.\,\frac{\pi}{4}$$

$$=78{\cdot}5\,\frac{E_{pm}}{E_{bb}}\ \% \qquad \ldots(13)$$

Where E_{pm} is the maximum amplitude of fundamental component of the plate voltage.

The plate dissipation (assuming no a.c. losses in tuned circuit) is

$$P_d=P_{bb}-P_{ac}=E_{bb}.\,\frac{2}{\pi}\left(\frac{\mu E_{gm}}{R_0+2r_p}\right)-\frac{\mu^2 E_{gm}^2 R_0}{2(R_0+2r_p)^2} \qquad \ldots(14)$$

Alternately, it can be expressed as

$$P_d=P_{bb}\left(1-\frac{P_{ac}}{P_{bb}}\right)=P_{bb}\,(1-\eta_p) \qquad \ldots(15)$$

From equation (11) it is clear that the a.c. power output is proportional to the square of input voltage (E_g). Hence grid drive should be as large as possible. When $(e_c)_{max} \triangleright (e_b)_{min}$, the relationship derived above for P_{ac} no longer holds good because the grid is positive and draws a larger current and consequently the plate current decreases. Now the power output P_{ac} no longer increases as the square of E_g. Due to this fact the power amplification decreases and harmonic in plate current increases in magnitude. Thus $(e_c)_{max}$ is never made to exceed $(e_b)_{min}$.

The output power, efficiency and plate dissipation can be calculated from the optimum condition $(e_c)_{max.} = (e_b)_{min.}$

We have,

$$(e_c)_{max.} = E_{cc} + E_{gm}$$
$$(e_b)_{min.} = E_{bb} - E_{pm}$$

$\therefore$ $$E_{cc} + E_{gm} = E_{bb} - E_{pm}$$

or $$E_{gm} + E_{pm} = E_{bb} - E_{cc}$$

$$= E_{bb} + \frac{E_{bb}}{\mu} \qquad \because \quad E_{cc} = -E_{bb}/\mu$$

$$= E_{bb}\left(1 + \frac{1}{\mu}\right)$$

$\therefore$ $$E_{gm} = E_{bb}\left(1 + \frac{1}{\mu}\right) - E_{pm}$$

$$= E_{bb}\left(1 + \frac{1}{\mu}\right) - I_{p1m}\, R_0$$

$$= E_{bb}\left(1 + \frac{1}{\mu}\right) - \frac{\mu E_{gm}}{2r_p + R_0} \cdot R_0$$

or $$E_{gm}\left(1 + \frac{\mu R_0}{2r_p + R_0}\right) = E_{bb}\left(1 + \frac{1}{\mu}\right)$$

$$E_{gm} = E_{bb}\left(1 + \frac{1}{\mu}\right)\left[\frac{2r_p + R_0}{2r_p + R_0\,(1+\mu)}\right] \quad ...(16)$$

The equation (16) shows that a large value of grid drive (required for high power output) can be used if E_{bb} is made large. As the maximum allowable d.c. plate voltage for specified tube is fixed by the manufacturer, hence E_{bb} can not be changed at will.

Now from equations (9) and (16), we have

$$I_{p1} = \frac{F_{bb}\,(\mu+1)}{\sqrt{2}} \cdot \frac{1}{2r_p + (\mu+1)\;R_0} \quad ...(17)$$

and from equations (5) and (6)

$$(I_b)_{av} = \frac{I_{bm}}{\pi} = \frac{2I_{p1m}}{\mu} = \frac{2\sqrt{2}\;I_{p1}}{\mu}$$

$$= \frac{2\sqrt{2}}{\pi}\;\frac{\mu E_g}{2r_p + R_0}$$

or $$(I_b)_{av.}=\frac{2}{\pi}\frac{\mu E_{gm}}{2r_p+R_0} \quad \ldots(18)$$

Substituting the value of E_{gm} from eq. (16) in equation (18), we have

$$(I_b)_{av.}=\frac{2}{\pi}E_{bb}(\mu+1).\frac{1}{2r_p+(\mu+1)R_0} \quad \ldots(19)$$

Now efficiency corresponding to optimum condition *i.e.* $(e_b)_{max.}=(e_b)_{min.}$ is

$$\eta_p=\frac{\pi}{4}\frac{E_{pm}}{E_{bb}}=\frac{\pi}{4}\frac{I_{p1m}R_0}{E_{bb}}$$

$$=\frac{\pi}{4}\frac{(\mu+1)R_0}{2r_p+(\mu+1)R} \quad \ldots(20)$$

The plate dissipation $P_d=P_{bb}-P_{ac}$

$$=E_{bb}(I_b)_{av}-I_{p1}^2R_0$$

$$=\frac{2}{\pi}\frac{(\mu+1)E_{bb}^2}{2r_p+(\mu+1)R_0}-\frac{(\mu+1)^2E_{bb}^2R_0}{2[2r_p+(\mu+1)R_0]^2} \quad \ldots(21)$$

The equation gives allowable plate dissipation in terms of tube parameters and impedance of the load (R_0) at resonant frequency.

The value of R_0 for optimum condition can be obtained from the following expression

$$R_0^2+\left[\frac{4r_p}{(\mu+1)}+\frac{E_{bb}^2}{P_d}\left(\frac{1}{2}-\frac{1}{\pi}\right)\right]R_0+\left[\frac{4r_p}{(\mu+1)^2}+\frac{4r_pE_{bb}^2}{(\mu+1)P_d}\right)=0 \quad \ldots(22)$$

23·12. TUNED CLASS C AMPLIFIER

A class *C* amplifier is one in which high output and plate efficiency are the primary considerations The grid is negatively biased to a point considerably beyond cut-off with the result that plate current flows in pulses and lasts for less than half a cycle. There is practically no output for small grid signal voltage because the plate curent is zero ; thus the amplitude of the output voltage is not proportional to the input voltage. However, the output is sinusoidal for the sinusoidal input, because of the tuned load. Class *C* amplifiers are useful only to amplify a signal of constant amplitude. Plate efficiency is about 85%.

The basic circuits for class *C* amplifier are shown in figure 12. There are two types : (*i*) series fed, and (*ii*) shunt fed.

In series fed types the d.c. component of plate current flows through d.c supply and circuit in series. In shunt fed the d.c. component of plate circuit flows through radio frequency choke (RFC) while a.c. component through tank circuit and blocking capacitor C_c. There are two advantages of shunt fed. Firstly in this case less insulation is required because one side of tank

circuit is earthed instead of the whole coil being above ground potential. Secondary, the d.c. supply is not short-circuited even if there is an arc between the coil and earth.

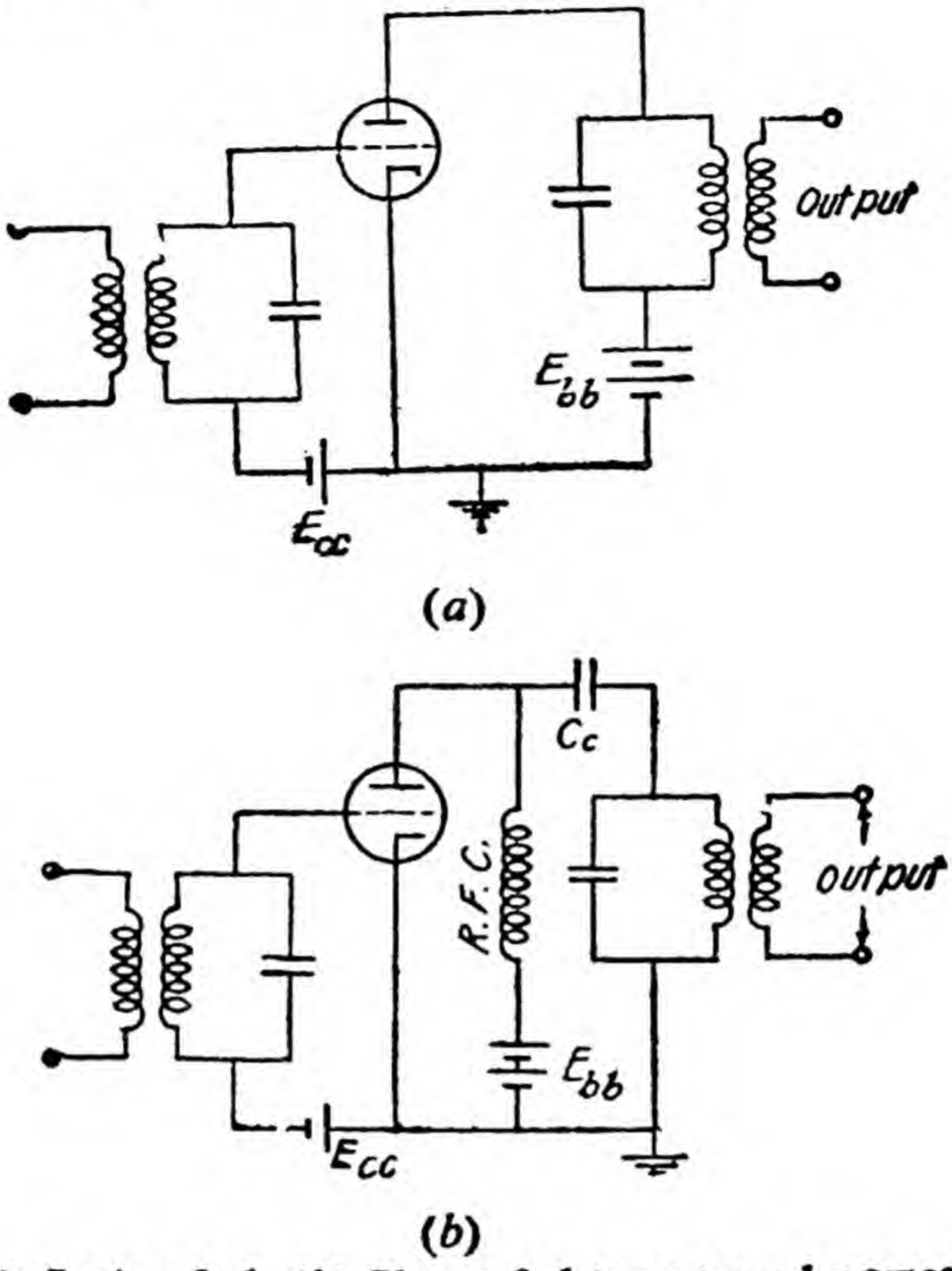


Fig. 12. (*a*) Series fed, (*b*) Shunt fed type tuned power amplifier.

Waveforms of grid voltage, plate voltage and plate current of both types are similar and are shown in figure (13). As the tuned circuit offers a pure resistance at resonant frequency, the a.c. plate voltage is 180° out of phase with grid voltage.

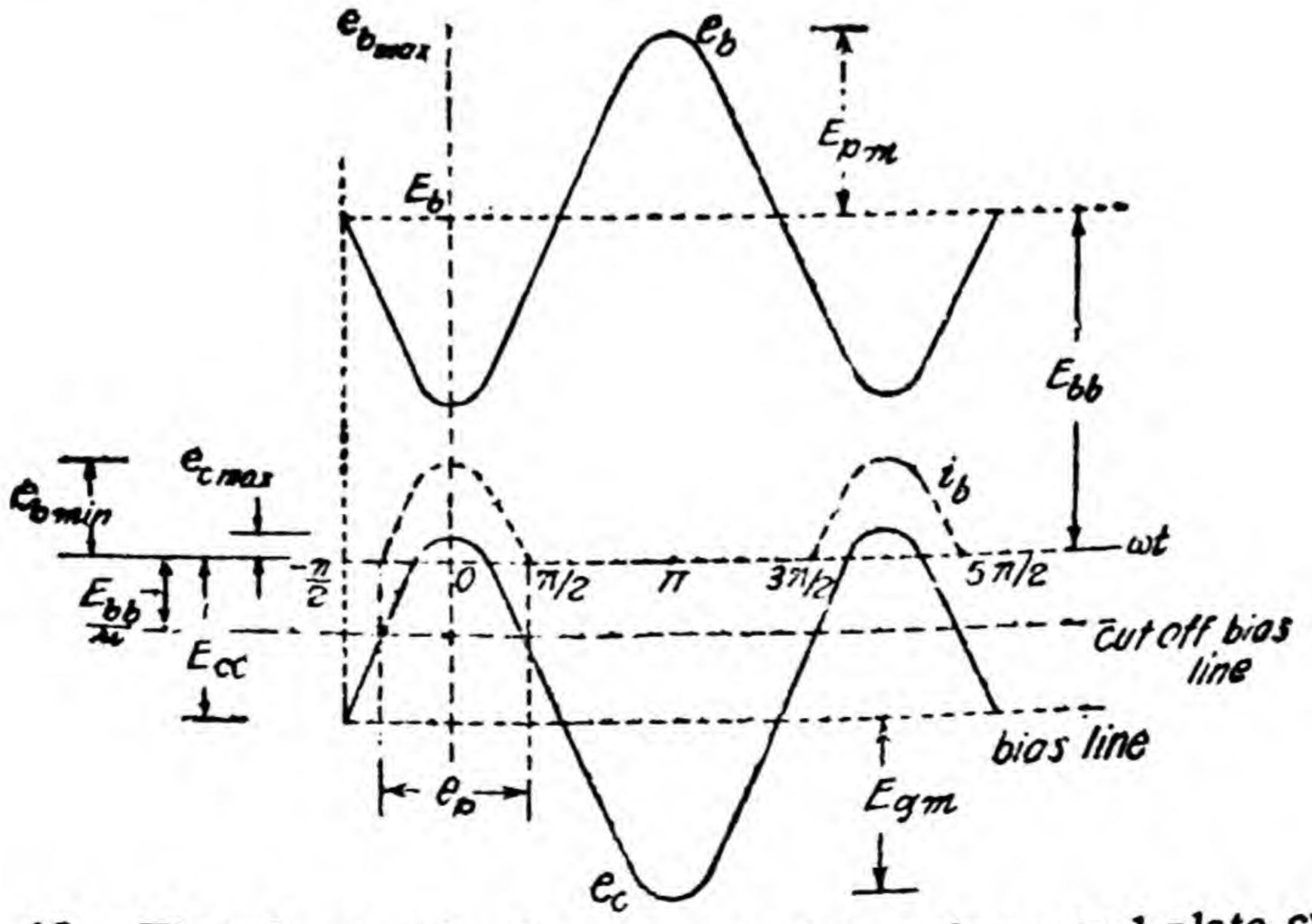


Fig. 13. Waveforms of grid voltage, plate voltage and plate current in class C tuned amplifier.

The following important points may be noted for the operation of class *C* tuned power amplifier :

1. The plate current flows for a period less than half the a.c. cycle when plate voltage is near about its minimum.
2. As the plate current flows for less than 180°, it is more distorted than in class *B* tuned amplifier and consequently it is more rich in harmonics
3. As the plate current flows in the pulses, the power flow to the tuned circuit of the plate is not continuous
4. The grid and plate voltages are sinusoidal as they appear across the tuned circuit. They are 180° out of phase. The load has no reactive component when tuned to resonance.
5. The maximum positive value of grid voltage is approximately equal to the minimum value of plate voltage (*i.e.*, $(e_c)_{max.}=(e_b)_{min.}$ as in class *B* tuned amplifier. If this is not so the plate dissipation and grid losses would increase which result in a decrease of plate efficiency.
6. When the grid voltage is positive with respect to cathode, the grid current flows. This amount to large power flow in grid-cathode circuit. The power is derived from the input (or driver) signal source. Therefore, driver circuit should be designed to furnish the power required to drive class *C* amplifier.
7. Class *C* tuned power amplifier can handle only a signal of constant amplitude. It is not suitable to amplify signals of varying amplitude such as amplitude modulated signals.

13·13. ANALYSIS OF CLASS C POWER AMPLIFIER :

The aim of this section is to consider the different steps which lead to the approximate design of class *C* power amplifier. *The first step is to compute the ratio of peak current to average current and the ratio of fundamental to the average current for various angular widths of the plate current.* For this purpose it is assumed that the plate pulse is a portion of cosine wave as shown in figure (14) having the same frequency as that of the input signal. The continuous curve indicates the plate current pulse and dotted part shows the extension of cosine curve. Here θ_p is the conduction angle. With *O* as the origin, the equation of the curve can be represented as $i=I\cos\omega t$ Now the instantaneous value of the plate current is

$$i_b=i(\omega t)-I\cos\theta_p/2$$

$$=I\cos\omega t-I\cos\theta_p/2$$

or $$i_b=I(\cos\omega t-\cos\theta_p/2). \quad ...(1)$$

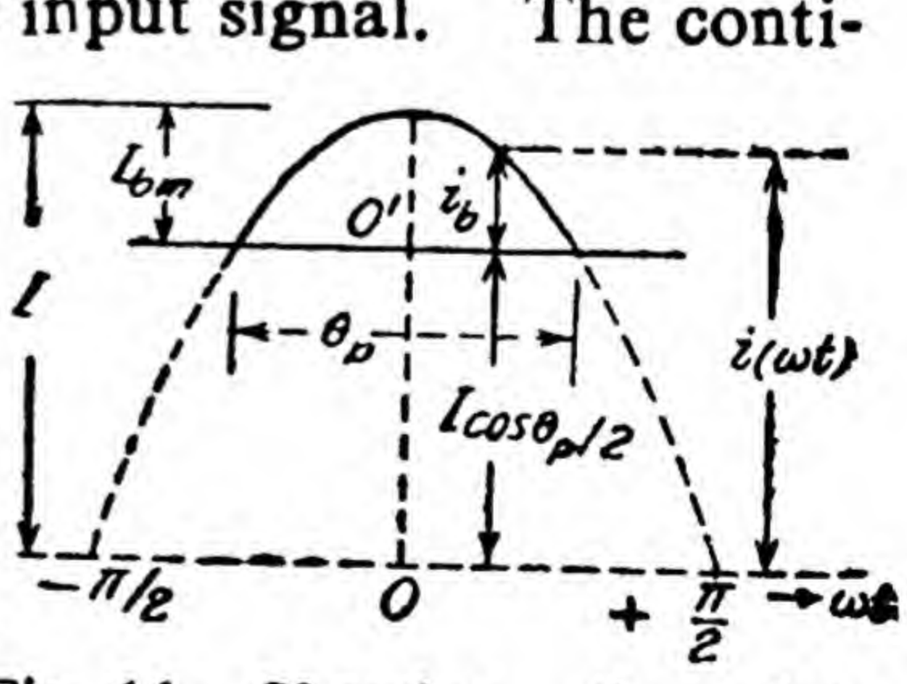


Fig. 14. Showing the plate current which is a part of cosine curve.

Using Fourier Series the plate current can be expressed as

$$i_b=(I_b)_{av.}+I_{p1m}\cos\omega t+I_{p2m}\cos 2\omega t+...$$

$$\therefore \quad (I_b)_{av.} = \frac{1}{2\pi}\int_0^{2\pi} i_b \, d(\omega t)$$

$$= \frac{1}{2\pi}\int_{-\theta_p/2}^{+\theta_p/2} I(\cos \omega t - \cos \theta_p/2)\, d(\omega t)$$

$$= \frac{I}{2\pi}\left(\sin \omega t - \cos \theta_p/2 . \omega t\right)_{-\theta_p/2}^{p/2}$$

$$= \frac{I}{\pi}(\sin \theta_p/2 - 2.\cos \theta_p/2). \qquad \ldots(2)$$

From figure (14),

$$I_{bm} = I - I \cos \theta_p/2. \qquad \ldots(3)$$

From equations (2) and (3),

$$\frac{I_{bm}}{(I_b)_{av,}} = \frac{\pi(1-\cos \theta_p/2)}{(\sin \theta_p/2 - \theta_p/2.\cos \theta_p/2)} \qquad \ldots(4)$$

The fundamental component of the plate current is given by

$$I_{p1m} = \frac{1}{\pi}\int_0^{2\pi} i_b \cos \omega t.\, d(\omega t)$$

$$= \frac{1}{\pi}\int_{-\theta_p/2}^{\theta_p/2} I(\cos \omega t - \cos \theta_p/2) \cos \omega t . d(\omega t)$$

$$= \frac{1}{\pi}\int_{-\theta_p/2}^{\theta_p/2} (\cos^2 \omega t - \cos \theta_p/2.\cos \omega t)\, d(\omega t)$$

$$= \frac{1}{\pi}\int_{-\theta_p/2}^{\theta_p/2} [\tfrac{1}{2}(1+\cos 2\omega t) - \tfrac{1}{2}\{\cos(\omega t + \theta_p/2) + \cos(\omega t - \theta_p/2)\}]\, d(\omega t)$$

$$= \frac{I}{\pi}\left[\tfrac{1}{2}\left(\omega t + \frac{\sin 2\omega t}{2}\right) - \tfrac{1}{2}\left\{\sin(\omega t + \theta_p/2) + \sin(\omega t - \theta_p/2)\right\}_{-\theta_p/2}^{\theta_p/2}\right]$$

$$= \frac{I}{\pi}\left[\frac{\theta_p}{2} - \tfrac{1}{2}\sin \theta_p\right]$$

$$\text{or} \quad I_{p1m} = \frac{I}{\pi}\left[\frac{\theta_p}{2} - \sin \theta_p/2.\cos \theta_p/2\right]. \qquad \ldots(5)$$

$$\text{Now } I_{p1} = \frac{I}{\sqrt{2}\pi}\left[\frac{\theta_p}{2} - \sin \theta_p/2 \cos \theta_p/2\right] \quad \because \quad I_{p1m} = \sqrt{2}.I_{p1}.$$

$$\therefore \quad \frac{I_{p1}}{(I_b)_{av.}} = \frac{1}{\sqrt{2}} \cdot \frac{(\theta_p/2 - \sin \theta_p/2.\cos \theta_p/2)}{(\sin \theta_p/2 - \theta_p/2.\cos \theta_p/2)}. \qquad \ldots(6)$$

The second step is to compute the average and fundamental components of grid current : As grid current—grid voltage characteristic is not linear, the grid current pulse is somewhat different from that of plate current pulse. To a first approximation the grid current pulse may be assumed to be a sine squared function as shown in figure (5) by a dotted rectangular pulse. It can be seen from the figure that conduction angle of grid current pulse θ_g is slightly less than the conduction angle θ_p of the plate current pulse.

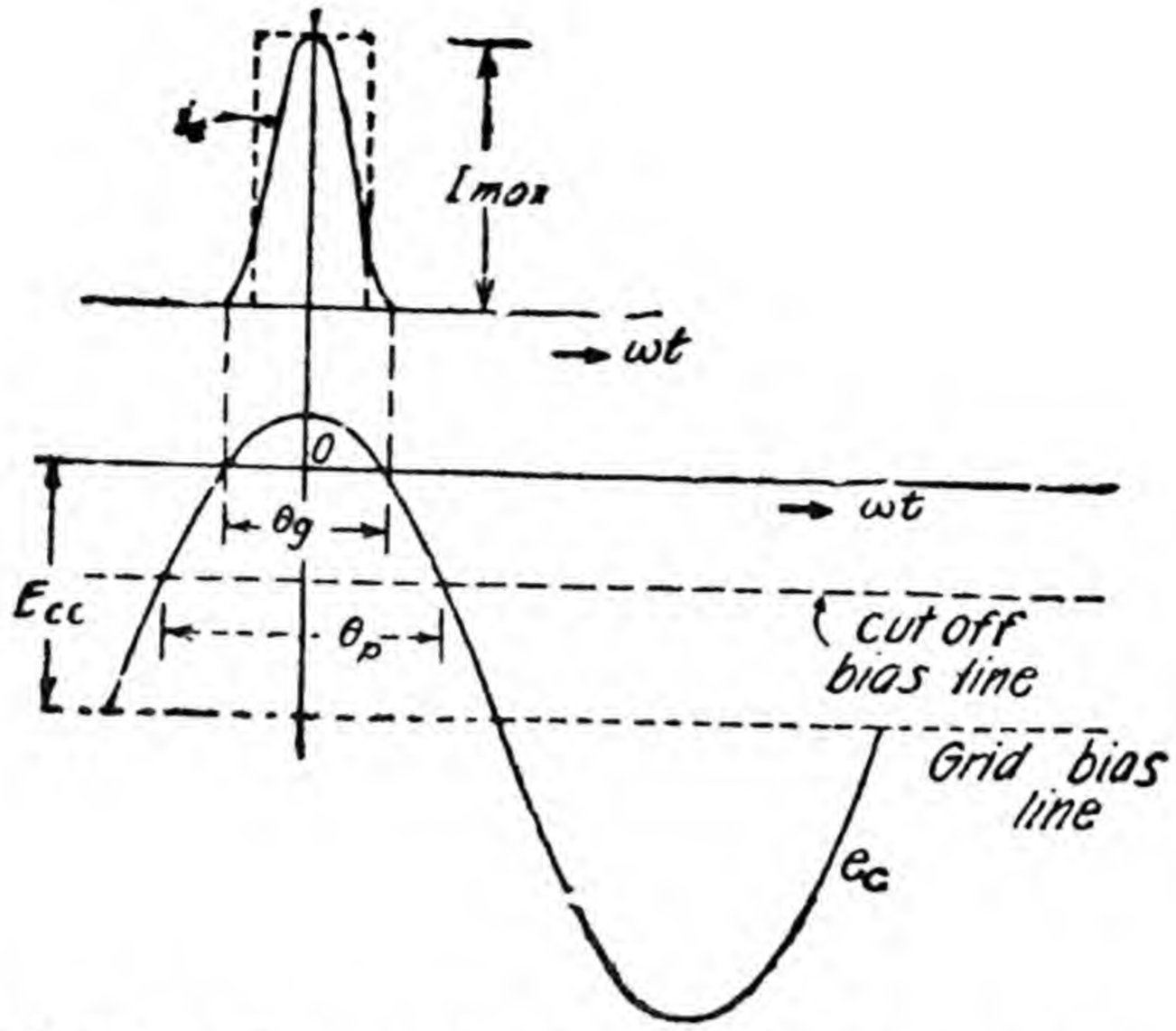


Fig. 15. Grid input voltage and grid current pulses of class C amplifier.

The grid current pulse can be represented from Fourier series analysis as

$$i_o=(I_o)_{av.}+I_{g1m}\cos\omega t+I_{g2m}\cos 2\omega t+ \ldots$$

The average and fundamental components can be calculated as follows :

$$(I_o)_{av.}=\frac{1}{\pi}\int_0^{2\pi} i_o\, d(\omega t)$$

$$=\frac{1}{\pi}\int_0^{\theta_g/2} i_c\, d(\omega t) \qquad \ldots(7)$$

and

$$I_{g1m}=\frac{1}{\pi}\int_0^{2\pi} i_c\cos\omega t.d(\omega t)$$

$$=\frac{2}{\pi}\int_0^{\theta_g/2} i_c\cos\omega t.d(\omega t)$$

$$=\frac{2}{\pi}\int_0^{\theta_g/2} i_c.d(\omega t), \qquad \ldots(8)$$

because the grid current flows for a relatively small portion of the cycle in the neighbourhood of $\theta_g=0$ and hence the value of cos ωt can be taken as unity.

Therefore $$I_{g1m}=2\,(I_o)_{av.}$$

or $$I_{g1}=\frac{I_{g1m}}{\sqrt{2}}=\sqrt{2}\,(I_o)_{av.}$$

Now $$\frac{I_{g1}}{(I_o)_{av.}}=1{\cdot}4 \text{ (when } \theta_g=0). \qquad \ldots(9)$$

Similarly, the value of $(I_g)_{max.}/(I_o)_{av}$ can be calculated but this is somewhat complicated.

The third step is to compute the grid bias : In order to show the different factors which determine the grid bias, the graphs of figure (13) are reproduced in figure (16). It is observed from this

figure that $\theta_g/2$ is that value of ωt at which the instantaneous grid voltage is zero and the grid current drop to partically zero. Similarly, $\theta_p/2$ is that value of ωt at which the plate current drops to zero.

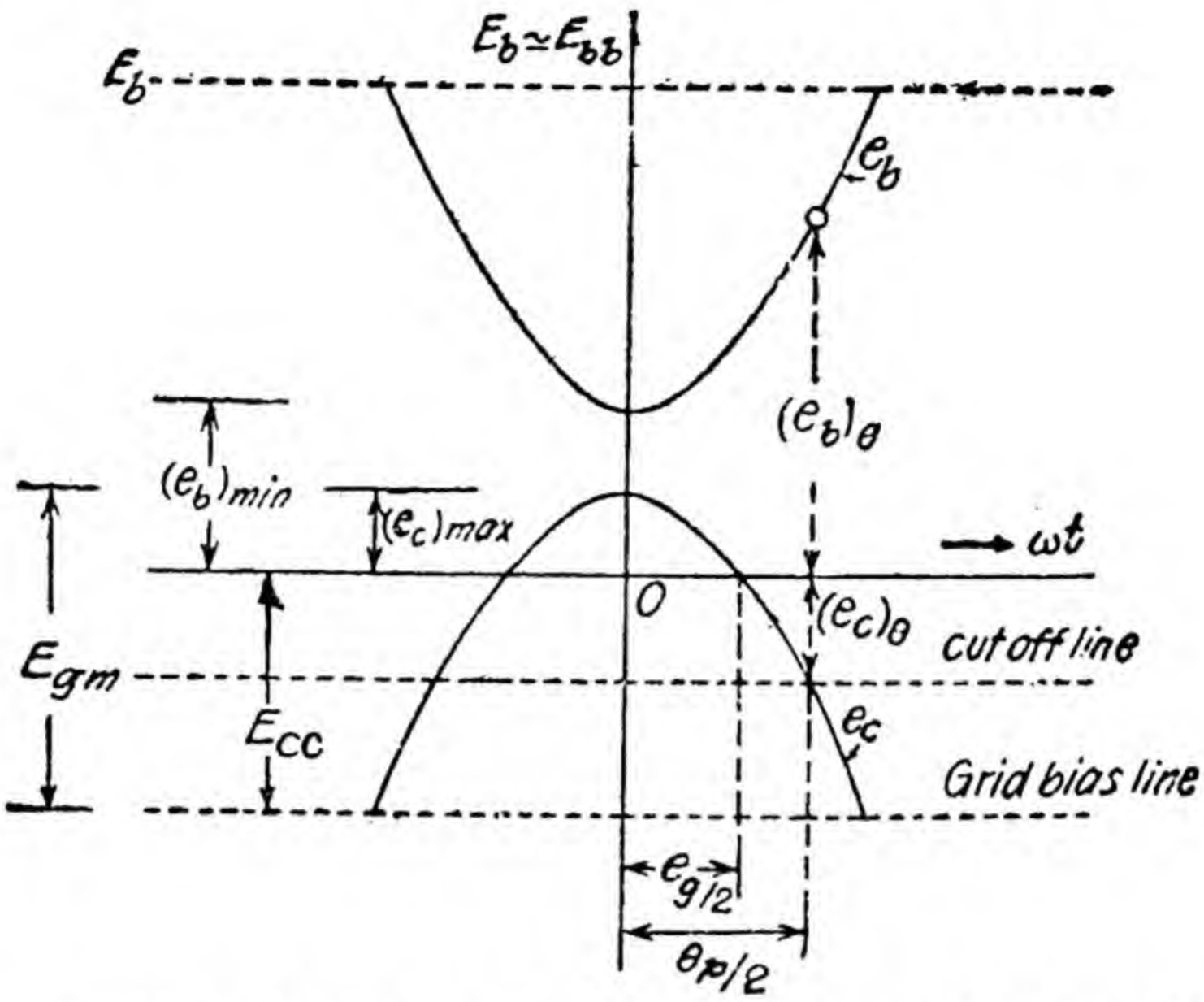


Fig. 16. Grid and plate voltage waveforms of a class C amplifier.

The plate current is zero at $\omega t = \theta_p/2$, hence the corresponding grid voltage, which is, of course, the cut off grid bias is

$$(e_c)_{\theta=\theta_p/2} = -(e_b)_{\theta=\theta_p/2}\,\mu$$

or simply
$$(e_c)_\theta = -\frac{(e_b)_\theta}{\mu}. \qquad \text{..(10)}$$

The instantaneous values of grid and plate voltages are expressed as

$$e_c = E_{cc} + E_{gm}\cos\omega t$$

or
$$(e_c)_\theta = E_{cc} + E_{gm}\cos\theta_p/2 \quad (\because\ \omega t = \theta = \theta_p/2) \qquad \text{...(11)}$$

and
$$e_b = E_{bb} - E_{pm}\cos\omega t$$

or
$$(e_b)_\theta = E_{bb} - E_{pm}\cos\theta_p/2. \qquad \text{...(12)}$$

Also
$$E_{pm} = E_{bb} - (e_b)_{min.} \quad \text{and} \quad E_{gm} = (e_c)_{max.} - E_{cc}. \qquad \text{...(13)}$$

Substituting equation (13) in equations (11) and (12), we get

$$(e_c)_\theta = E_{cc} + [(e_c)_{max.} - E_{cc}]\cos\theta_p/2 \qquad \text{...(14)}$$

and
$$(e_b)_\theta = E_{bb} - [E_{bb} - (e_b)_{min}]\cos\theta_p/2. \qquad \text{...(15)}$$

Using equation (10), we have

$$E_{cc} + [(e_c)_{max.} - E_{cc}]\cos\theta_p/2.$$

$$= -\frac{1}{\mu}\,[E_{bb} - \{E_{bb} - (e_b)_{min}\}\cos\theta_p/2]$$

or
$$E_{cc} = -\frac{1}{1-\cos\theta_p/2}\left[\frac{E_{bb}}{\mu}(1-\cos\theta_p/2) + \left\{\frac{(e_b)_{min.}}{\mu} + (e_c)_{max.}\right\}\cos\theta_p/2\right] \qquad \text{...(16)}$$

Thus the grid bias can be calculated from equation (16).

The last step is to compute the grid driving power and grid dissipation. Instantaneous grid power supplied by the driving source (or signal) is $e_g.i_c$. Hence the average grid power supplied by the driving source is

$$P_g=\frac{1}{2\pi}\int_0^{2\pi} e_g.i_c\ d\,(\omega t).$$

Now we assume that the grid potential is at its maximum value when the grid current flows and does not vary appreciably during this interval.

Hence $$P_g=\frac{E_{gm}}{2\pi}\int_0^{2\pi} i_c\ d\,(\omega t)\simeq E_{gm}.\,(I_c)_{av}. \qquad ...(17)$$

The grid dissipation *i.e.*, the loss in the cathode grid circuit is given by

$$P_c=\frac{1}{2\pi}\int_0^{2\pi} e_c\ i_c\ d\,(\omega t)=\frac{1}{2\pi}\int_0^{2\pi}(E_{cc}+e_g)\ i_c\ d\,(\omega t)$$
$$=E_{cc}\,(I_c)_{av}.+E_{gm}\,(I_c)_{av}.$$
$$=P_{cc}+P_g \qquad ...(18)$$

The first term is the amount of power absorbed by the d.c. grid battery source from the signal because

$$P_{cc}=\frac{1}{2\pi}\int_0^{2\pi} E_{cc}\ i_c\ d\,(\omega t)=E_{cc}.(I_c)_{av}.$$

Actually E_{cc} is negative, hence the grid dissipation is

$$P_c=P_g-|\,P_{cc}\,|. \qquad ...(19)$$

Example. *Design a class C radio frequency power amplifier for continuous working using a high power triode tube with the following specifications :*

$\mu=25$, $E_{bb}=2{,}000\ V$, $E_{cc}=250\ V$, $(I_b)_{av}=500mA$, $(I_c)_{av}=75mA$.

$I_c/(I_c)_{av}=1{\cdot}25$, *Plate dissipation* $=250$ *watts.*

Plate efficiency $=7{\cdot}5\%$

$(e_b)_{min}=(e_c)_{max}=200V$ *for a current* $1{\cdot}9$ *amp.*

We know that $$\eta_p=\frac{P_{ac}}{P_{bb}}=\frac{P_{ac}}{P_d+P_{ac}}$$

or $$P_{ac}\,(1-\eta_p)=P_d\ \eta_p$$

$$\therefore \quad P_{ac}=\frac{\eta_p P_d}{1-\eta_p}=\frac{0{\cdot}75\times 250}{1-0{\cdot}75}=750 \text{ watt.}$$

Total power input to the circuit $P_{bb}=P_{ac}+P_d$

$$=750+250=1{,}000 \text{ watt.}$$

Now $$E_{bb}\ (I_b)_{av}.=1000$$

or $$(I_b)_{av}.=\frac{1000}{2000}=0{\cdot}50 \text{ amp.}$$

Hence the specifications for the design of amplifier are

$E_{bb}=2000V$, $(I_b)_{av.}=0{\cdot}5A$ and $\theta_p/2=75°$ (say)

(i) from equation (4)

$$\frac{I_{bm}}{(I_b)_{av.}}=\frac{\pi\,(1-\cos\theta_p/2)}{(\sin\theta_p/2-\theta_p/2.\cos\theta_p/2)}=3{\cdot}8$$

$\therefore$ $(I_b)_{max.}=3{\cdot}8\times0{\cdot}5=1{\cdot}9$ amp.

(ii) $(e_b)_{min}=(e_c)_{max.}=200$ volt.

(iii) $E_{pm}=E_{bb}-(e_b)_{min}=2000-200=1800$

$\therefore$ $E_p=\dfrac{E_{pm}}{\sqrt{2}}=0{\cdot}707$ $E_{pm}=0{\cdot}707\times1800$ volt. $=1272{\cdot}6$ volt.

(iv) From equation (6) $I_p/(I_b)_{av.}=1{\cdot}20$

$\therefore$ R.M.S. plate current $=I_p=1{\cdot}20\times(I_b)_{av.}$

$=1{\cdot}20\times0{\cdot}5=0{\cdot}60$ amp.

(v) A.C. power output $=E_p\times I_p=1272{\cdot}6\times0{\cdot}60=763{\cdot}56$ watt.

(vi) Total power input of the plate circuit $=E_{bb}\times(I_b)_{av.}$

$=2000\times0{\cdot}5$

$=1000$ watt.

(vii) Plate circuit efficiency $=\eta_p=\dfrac{763{\cdot}56}{1000}=76{\cdot}4\%$

(viii) The grid bias E_{cc} is

$$E_{cc}=-\frac{1}{1-\cos\theta_p/2}\left[\frac{E_{bb}}{\mu}\left(1-\cos\theta_p/2\right)+\left\{\frac{(e_b)_{min.}}{\mu}+(e_c)_{max.}\right\}\cos\theta_p/2\right]$$

$$=-\frac{1}{1-0{\cdot}26}\left[\frac{2000}{25}\,(1-0{\cdot}26)+\left(\frac{200}{25}+200\right)0{\cdot}26\right]$$

$$=-\frac{1}{0{\cdot}74}\left[80\times0{\cdot}74+208\times0{\cdot}26\right]$$

$$=\frac{1}{0{\cdot}74}\times113{\cdot}28=-153\text{ volt.}$$

(ix) R.M.S. value of the grid voltage

$E_g=0{\cdot}707$ $E_{pm}=0{\cdot}707$ $[(e_c)_{max}-E_{cc}]$

$=0{\cdot}767\,[200+153]$

$=249{\cdot}571$ volt.

(x) $I_c/(I_c)_{av.}=1{\cdot}25$ $\therefore$ $I_c=1{\cdot}25\times(I_c)_{av.}$

$=1{\cdot}25\times0{\cdot}075$

$=0{\cdot}09375$

(xi) Grid driving power $=E_g.I_c=249{\cdot}571\times0{\cdot}02375$

$=$ **23·4 watts.**

13·14. SELF OSCILLATIONS IN AMPLIFIERS :

Due to the inter electrode capacities between different electrodes the feed back of energy takes place which produces self-oscillation

in amplifiers, especially at high frequencies. The most important inter-electrode capacity from this point of view is that existing between grid and plate denoted by C_{gp}. Consider the simple case of triode having tuned circuits, one in between grid and filament and other in plate circuit as shown in fig. 17. The two tuned circuits are connected together by virtue of C_{gp}. The transference of energy takes place from the tuned plate circuit back to the input circuit. Greater the value of C_{gp} the higher is the coefficient of coupling between the circuits and greater is the amount of energy, fed back. Due to this energy self-oscillations take place. Various methods are used for the prevention of self-oscillations in radio frequency amplifiers. The most important method uses the neutralising circuits.

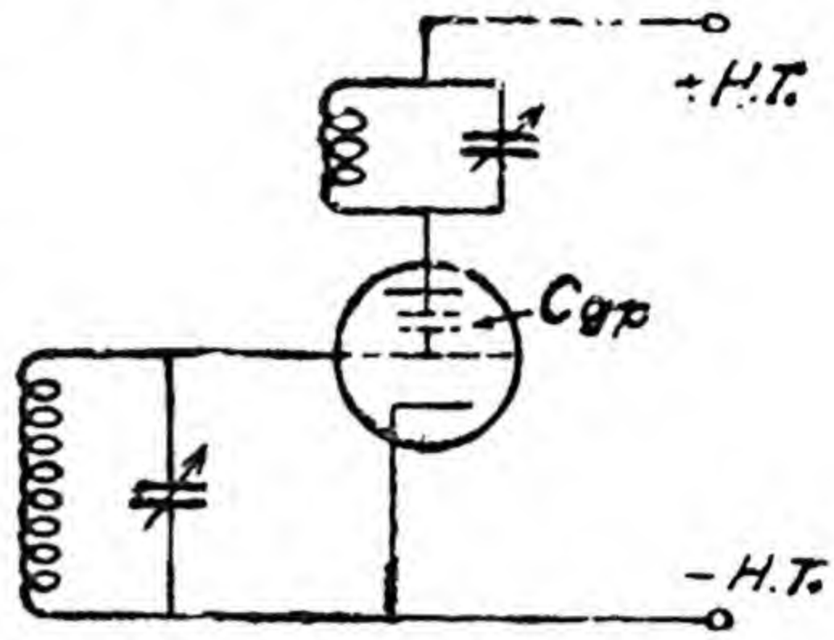


Fig. 17. Triode with tuned circuits.

Neutralising circuits :

Figure 18 (*a*) and 18 (*b*) show the form of the bridge circuit.

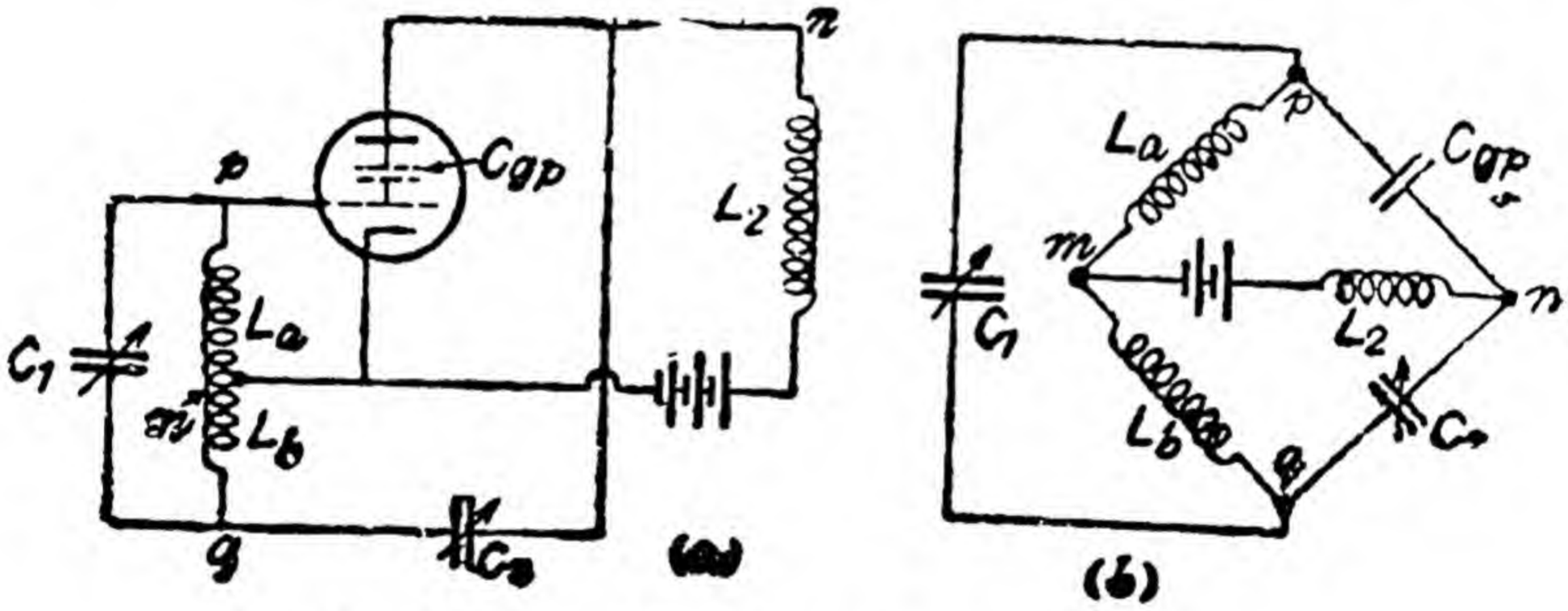


Fig. 18. (*a*) Neutralising circuit.
(*b*) Its equivatent circuit.

The cathode terminal instead of connecting at the lower end of the input coil, is connected to the intermediate point which divides the inductance into two parts L_a and L_b. The lower end (q) of the input coil is connected to the plate through balancing condenser C_n. The terminals p and q of the input circuit, and m and p of the output circuit constitute two pairs of diagonally opposite points of a bridge as shown in fig. 13 (*b*). When the bridge is balanced,

$$\frac{L_a}{L_b} = \frac{C_n}{C_{gp}}.$$

This balance is independent of frequency.

This neutralising circuit is commonly used in radio frequency power amplifiers in transmitting sets.

13·15. NOISE IN AMPLIFIERS :

In the absence of any input signal, all amplifiers give some output and thus unwanted output is referred to as noise. The noise is introduced by a number of sources. The important sources are (*i*) inherent particle nature of electron flow in tubes and (*ii*) due to circuit sources external to the tubes. Here we shall discuss the noise produced due to different sources in brief.

(i) **Thermal noise** : Every electric conductor produces irregularly varying voltage across its terminals due to the random motion of the free electrons present in conductor caused by thermal action. This effect is known as circuit noise, resistance noise or Johnson noise. The noise energy is distributed over all frequencies and is believed due to impact of electrons with molecules, each impact producing a very short pulse. The thermal noise voltage developed in a resistor R is found to be

$$E_{noise}=\sqrt{(4kRT\Delta f)},$$

where k=Boltzmann's constant,

T=temperature, K,

R=resistance, ohm, and

Δf=bandwidth accepted by the circuit. Because all the circuit components have resistance, Johnson noise is present in all circuit components.

(ii) **Shot noise** : This is due to the random emission of electrons from a cathode. The noise is reduced when the tube operates under space charge limited conditions because the space charge reservoir of electrons tends to smooth out the individual emission pulses. This noise is uniformly distributed over the useful frequency range. For triodes working under space charge limited conditions, the equivalent noise resistance due to the shot noise is given by

$$R_g=\frac{2\cdot5}{g_m},$$

where g_m is the transconductance of the tube.

(iii) **Gas noise** : The positive ions produced due to the ionization of the remaining gas in tube, causes the random current disturbances as they penetrate the space charge. This noise is not appreciable above 1 megacycles.

(iv) **Partition noise** : The noise is produced due to the irregular division of current between the two or more electrodes. It is arbitrary whether the electron goes to screen or to plate. Pentodes are more noisy in comparison to triodes because in a pentode one has partition noise in addition to shot effect noise. The equivalent grid resistance R_{eq} representing the noise of a negative grid pentode amplifier is given approximately by

$$R_{eq}=\frac{I_b}{I_b+I_{c_2}}\left(\frac{2\cdot5}{g_m}+\frac{20I_{c_2}}{g_m}\right)$$

where I_b = d.c. plate current,

I_{o2} = d.c. screen current,

g_m = transconductance.

(v) Microphone noise : The noise is produced due to the mechanical vibration of parts of circuit. Considering the case of a tube, when it is struck, the electrodes vibrate and produce current variation or noise.

(vi) Electric and Magnetic pickup : Stray electrostatic and magnetic fields are produced, externally due to adjacent power leads which induce voltage in grid circuit. This induced voltage is amplified and produces noise.

(vii) Filament hum : When directly heated filament operates with a.c then through the interelectrode capacitance between cathode and grid hum voltage is introduced which produces noise. Using indirectly heated filament this noise can be minimised.

(viii) Plate supply hum : This hum is produced due to a.c. present in the d.c. plate supply.

NUMERICAL EXAMPLES

1. *A tuned voltage amplifier has a triode having* $\mu = 25$, $r_p = 20.000$ *ohms. The plate tuned circuit consists of a coil of inductance* $L = 200\mu H$, $R = 10$ *ohms, tuning condenser* $C = 300\mu\mu F$., $E_{g1} = 0{\cdot}1$ *volts r.m.s. at the resonant frequency. Calculate :*

(a) resonant frequency,
(b) impedance of the tuned circuit,
(c) output voltage at resonance,
(d) Q of the tuned circuit at resonance.

(a) Resonant frequency

$$f_0 = \frac{1}{2\pi\sqrt{(LC)}} = \frac{1}{2\times 3{\cdot}14\times\sqrt{(200\times 10^{-6}\times 300\times 10^{-12})}}$$

$$= \mathbf{649{\cdot}7\ kc/s.}$$

(b) Impedance of the tuned circuit at resonance

$$Z = \frac{L}{RC} = \frac{200\times 10^{-6}}{300\times 10^{-12}\times 10}$$

$$= \mathbf{6{\cdot}67\times 10^4\ ohms.}$$

(c) Output voltage at resonance

$$E_{g2} = \frac{\mu E_{g1} Z}{r_p + Z} = \frac{25\times 0{\cdot}1\times 6{\cdot}67\times 10^4}{2\times 10^4 + 6{\cdot}67\times 10^4}$$

$$= \mathbf{1{\cdot}92\ volts.}$$

(d) $$Q_0 = \frac{\omega_0 L}{R} = \frac{2\pi\times 649{\cdot}7\times 10^3\times 200\times 10^{-6}}{10}$$

$$= \mathbf{81{\cdot}6.}$$

2. *The tuned voltage amplifier is a pentode with* $g_m = 1{\cdot}5$ *mA/volt and* $r_p = 10^6$ *ohms. The tuned circuit consists of a coil of inductance* $L = 200\mu H$ *and resistance* $R = 12$ *ohms in shunt with a tuning condenser adjusted to 500 pf. Calculate* (i) *resonant frequency,* (ii) *Q of the tuned circuit at resonance,* (iii) *impedance of the tuned circuit at resonance,* (iv) *effective Q of the resonant circuit including* r_p, (v) *voltage gain at resonance. If the frequency of the applied signal is increased to 10 KHz above the resonance, calcluate the magnitude and phase angle of voltage gain.*

(*i*) Resonant frequency $f_0 = \dfrac{1}{2\pi\sqrt{(LC)}}$

$$= \frac{1}{2\pi\,[200 \times 10^{-6} \times 500 \times 10^{-12}]^{1/2}}$$

$$= 503\ KHz.$$

(*ii*) Q of the tuned circuit at resonance Q_0 is given by

$$Q_0 = \frac{1}{R}\sqrt{\left(\frac{L}{C}\right)} = \frac{1}{12}\sqrt{\left(\frac{300 \times 10^{-6}}{500 \times 10^{-12}}\right)} = 52{\cdot}7$$

(*iii*) Impedance of tuned circuit at resonance is given by

$$Z_0 = \omega_0 L Q_0 = \frac{1}{CR} = Q_0^2 R$$

$$= (52{\cdot}5)^2 \times 12$$

$$= 34 \times 10^3 \text{ ohms.}$$

(*iv*) Z_0 and r_p are in parallel

$$\therefore \quad \frac{1}{Z_t} = \frac{1}{Z_0} + \frac{1}{r_p} = \frac{r_p + Z_0}{Z_0 r_p}$$

or
$$Z_t = \frac{Z_0 r_p}{r_p + Z_0} = \frac{34 \times 10^3 \times 10^6}{10^6 + 34 \times 10^3.}$$

Effective Q of the plate circuit Q_e is given by

$$Q_e = \frac{Z_t}{X_e} = Z_t\sqrt{\left(\frac{C}{L}\right)} = 32{\cdot}8 \times 10^3\sqrt{\left(\frac{500 \times 10^{-12}}{200 \times 10^{-6}}\right)}$$

$$= 52{\cdot}48.$$

(*v*) Voltage gain at resonance $A_{r\,cs.} = -g_m Z_t$

$$= -\frac{1{\cdot}5}{1000} \times 32{\cdot}8 \times 10^3$$

$$= -49{\cdot}20.$$

At a frequency 10 *KHz* above resonance

$$\delta = \frac{f - f_0}{f_0} = \frac{10 \times 10^3}{503 \times 10^3} = {\cdot}019.$$

Relative voltage gain is given by

$$\frac{\mathbf{A}}{\mathbf{A}_{r\,os}}=\frac{1}{1+j2\delta Q_e}=\frac{1}{1+j2\times\cdot019\times52\cdot48}$$

$$=\frac{1}{1+j\,(1\cdot9\gamma)}$$

or $$\mathbf{A}=\frac{-49\cdot20}{1+j\,(1\cdot99)}\cdot$$

Magnitude $A=\frac{49\ 20}{\sqrt{[1+(1\cdot9\gamma)^2]}}=22\cdot4$.

Phase angle of $\mathbf{A}=\phi=180°-\tan^{-1}(1\cdot99)=180°-62°=118°$.

3. *A tuned amplifier uses a pentode valve whose parameters under the operating conditions are mutual conductance* $g_m=5mA/V$, *plate resistance* $r_p=0\cdot5\ M\Omega$. *The anode load consists of a tuned circuit having an inductance coil of* $25\mu H$ *with* Q *value of 50 in parallel with a capacitance and tuned to 4MHz. The output of the amplifier is coupled to the next stage having an input resistance of* $0\cdot5\ M\Omega$ *through a coupling capacitor of negligible reactance. Calculate the gain of the stage at resonance.*

According to the problem $r_p=0\cdot5\Omega$; $R_g=0\cdot5\ M\Omega$.

Impedance of tuned circuit at resonance is given by

$$Z_0=\omega_0 LQ_0=2\pi\times4\times10^6\times250\times10^{-6}\times50$$

$$=314{,}000\Omega=0\ 31\ M\Omega.$$

$$\text{Now}\ \frac{1}{Z_t}=\frac{1}{Z_0}+\frac{1}{r_p}+\frac{1}{R_g}$$

$$=\frac{1}{0\cdot314\times10^6}+\frac{1}{0\cdot5\times10^6}+\frac{1}{0\cdot5\times10^6}$$

$$=[3\cdot183+2+2]\times10^{-6}=7\cdot183.10^{-6}\ \text{mhos}$$

$$\text{Voltage gain}=\frac{E_0}{E_g}=-g_m\ Z_t=-\frac{5\times10^3}{7\cdot183}\times10^6$$

$$=-695$$

4. *In the capacitor coupled tuned voltage amplifier, the value of circuit elements are as below ;*

$\mu=60$, $r_p=30K\Omega$, $L=400\mu H$, $C'=380\mu\mu F$, $R=10\Omega$, $R_g=500K\Omega$. *Additional shunt capacitance* $C_s=20\mu\mu f$. *Reactance of* C_c *is negligibly small. Calculate (i) frequency of resonance of plate circuit, (ii)* Q_0 *of the LCR circuit at resonance* $(C=C'+C_s)$. *(iii) effective value of* Q_0 *of plate circuit including* r_p *and* R_g, *(iv) voltage gain at resonance, (v) half power bandwidth, (vi) gain bandwidth product, (vii) voltage gain at frequency of 20 KHz below resonance.*

(*i*) $C=C'+C_s=380+20=400\mu\mu f$.

$$\text{Frequency of resonance}\ f_0=\frac{1}{2\pi\sqrt{(LC)}}$$

$$=\frac{1}{2\pi\sqrt{(400\times10^{-6}\times400\times10^{-12})}}$$

$$=389\times10^3 Hz.$$

(*ii*) $$Q_0=\frac{1}{R}\sqrt{\left(\frac{L}{C}\right)}=\frac{1}{10}\sqrt{\left(\frac{400\times10^{-6}}{400\times10^{-12}}\right)}=100.$$

(*iii*) $\omega_0=2\pi f_0=2\pi\times389\times10^3=2{\cdot}5\times10^6$ radian per sec.

$$Q_e=\frac{Q_0}{1+\omega_0 L Q_0\left(\frac{1}{r_p}+\frac{1}{R_g}\right)}$$

$$=\frac{100}{1+2{\cdot}5\times10^6\times400\times10^6\times100\left(\frac{1}{30\times10^3}+\frac{1}{500\times10^3}\right)}$$

$$=22{\cdot}05.$$

(*iv*) $$\mathbf{A}_{res}=-g_m\omega_0 L Q_e=-\frac{60}{30\times10^3}\times2{\cdot}5\times10^6\times400\times10^{-6}\times22{\cdot}05$$

$$=-44{\cdot}10.$$

(*v*) Half power bandwidth $$B=\frac{f_0}{Q_e}=\frac{389\times10^3}{22{\cdot}05}$$

$$=17{\cdot}6\ KHz.$$

(*vi*) Gain bandwidth product $=B\times A_{res}$.

$$=(17{\cdot}6\times10^3)\times44{\cdot}10$$

$$=775{\cdot}8\times10^3$$

(*vii*) At 20 *KHz.* below resonance

$$\delta=\frac{20\times10^3}{389\times10^3}=0{\cdot}0502.$$

Voltage gain $$\mathbf{A}=\frac{\mathbf{A}_{res}}{1+j.2\delta Q_e}=\frac{-44{\cdot}10}{1+j.2\times0{\cdot}0502\times22{\cdot}05}$$

$$=\frac{-44{\cdot}10}{1+j\times2{\cdot}205}$$

Magnitude $$\mathbf{A}=\frac{44{\cdot}10}{\sqrt{[1+(2205,^2]}}=17{\cdot}1.$$

5. *In a transformer coupled single tuned voltage amplifer,* $\mu=15$, $r_p=10{,}000$ *ohms,* $L_s=200\mu H$, $R_s=8\,\Omega$, $C_s=25-400$ *pf.* $M=20\mu H$, *and* C_s *is adjusted to* 200 *pf. Calculate* (*i*) *resonant freqnency* f_0, (*ii*) Q_0 *of the tuned circuit at resonance,* (*iii*) *voltoge gain at resonance,* (*iv*) Q_e *of the tuned circuit including* r_p *reflected into the secondary circuit,* (*v*) *optimum value of mutual inductance M,* (*vi*) *half power bandwidths.*

(*i*) $$f_0=\frac{1}{2\pi\sqrt{(L_sC_s)}}=\frac{1}{2\pi\sqrt{(2{\cdot}0\times10^{-6}\times200\times10^{-12})}}$$

$$=79{\cdot}57\times10^4\ \text{c/s.}$$

(*ii*) $$Q_0=\frac{1}{R_s}\sqrt{\left(\frac{L_s}{C_s}\right)}=\frac{1}{8}\sqrt{\left(\frac{200\times10^{-6}}{200\times10^{-12}}\right)}$$

$$=\frac{1}{8}\times 10^3=125.$$

(*iii*) $$A_0=\frac{\mu M\omega_0^2 L_2}{R_s r_p+\omega_0^2 M^2}=\frac{\mu M/C_s}{R_s r_p+\omega_0^2 M^2}$$

$$=\frac{15\times 20\times 10^{-6}/200\times 10^{-12}}{8\times 10^4+10^4}$$

$$=16{\cdot}67.$$

(*iv*) $$Q_0=\frac{Q_0}{1+\omega^2 M^2/R_s r_p}=\frac{125}{1+10^4/8\times 10^4}=111{\cdot}1$$

(*v*) To determine maximum value of M, differentiate A_0 with respect to M and equate it to zero.

We know that

$$A_0=\frac{\mu M/C_s}{R_s r_p+\omega_0^2 M^2}.$$

Applying the above mentioned condition, we have

$$M=\frac{\sqrt{(r_p R_s)}}{\omega_0}$$

$$=\frac{\sqrt{(10{,}000\times 8)}}{2\pi\times 79{\cdot}57\times 10^4}$$

$$=55{\cdot}56\mu H.$$

(*vi*) Band power bandwidth

$$B=f_0/Q_0$$

$$=\frac{79{\cdot}75\times 10^4}{111{\cdot}1}$$

$$\doteq 7{\cdot}16\times 10^3\ c/s.$$

6. *A single tuned transformer coupled amplifier uses a valve with* $\mu=40$ *and* $r_p=10\ K\Omega$. *The secondary of the transformer has a coil of inductance* $250\mu H$ *and resistance* 10Ω *and is tuned with a total capacitance (including wiring and input capacitance of the subsequent stage) of* $300\ \mu\mu F$. *If the mutual inductance between the primary and the secondary is* $25\mu H$, *calculate (i) the resonant frequency, (ii) the Q of the tuned circuit, (iii) voltage gain at resonance, (iv) voltage gain at* $10KHz$ *off-resonance above the resonant frequency.*

(*i*) The resonant frequency f_c

$$=\frac{\omega_0}{2\pi}=\frac{1}{2\pi}\frac{1}{\sqrt{(L_s C_s)}}$$

$$=\frac{1}{2\times 3{\cdot}14\sqrt{(250\times 10^{-6}\times 300\times 10^{-12})}}$$

$$=580\ KHz.$$

(*ii*) $$Q=\frac{\omega_0 L_s}{R_s}=\frac{3{\cdot}65\times 10^6\times 250\times 10^{-6}}{10}$$ (here $\omega_0=3{\cdot}65\times 10^6$)

$$=91{\cdot}25$$

$$(iii)\quad Q_e=\frac{Q}{1+\dfrac{\omega_0^2M^2}{r_pR_s}}=\frac{91\cdot25}{1+\dfrac{(3\cdot65\times10^6)^2\,(25)^2\times10^{-12}}{10\times10^3\times10}}$$

$$=\frac{91\cdot25}{1+0\cdot085}=84\cdot5$$

$$\therefore\quad A_{res.}=g_m\omega_0MQ_e$$

$$=\frac{40}{10\times10^3}\times3\cdot65\times10^6\times25\times10^{-6}\times84\,5$$

$$=30\cdot85$$

$$(iv)\quad \delta=\frac{\omega-\omega_0}{\omega_0}=\frac{f-f_0}{f_0}=\frac{10\times10^3}{580\times10^3}=0\cdot0172$$

$$\therefore\quad A=\frac{A_{res.}}{1+j.2\delta Q_e}=\frac{30\cdot85}{1+j2\times0\cdot0172\times84\cdot5}=\frac{30\cdot85}{1+j2\cdot91}$$

$$|A|=\frac{30\cdot85}{3\cdot06}=10$$

angle $\phi=\tan^{-1}(-2\delta Q_e)=\tan^{-1}(-2\cdot91)=-71°$

$$\therefore\quad A=10\angle-71°$$

7. *In a double tuned voltage amplifier.* $L_p=L_s=200\ \mu H$, $R_p=R_s=10\,\Omega$, $C_s=30-500\ \mu\mu F$ *variable. Coupling dissipation factor* $a^2=K^2Q_pQ_s=0\cdot81$. C_s *is adjusted to* 450 $\mu\mu F$. *Calculate (i) frequency of resonance, (ii)* Q *of primary and secondary tuned circuits, (iii) voltage gain at resonance, (iv) tuned circuit voltage gain at a frequency* 2*KHz, above resonant frequency.*

(*i*) Frequency of resonance

$$f_0=\frac{1}{\pi\sqrt{(L_2C_2)}}$$

$$=\frac{1}{2\pi\sqrt{(200\times10^6\times450\times10^{-12})}}$$

$$=531\ KHz.$$

$$(ii)\quad Q_s=\frac{\omega_0L_s}{R_s}=\frac{1}{R_s}\sqrt{\left(\frac{L_s}{C_s}\right)}=\frac{1}{10}\sqrt{\left(\frac{200\times10^{-6}}{450\times10^{-12}}\right)}$$

$$=66\cdot7$$

$$Q_p=Q_s=66\cdot7.$$

$$(iii)\qquad A_{res.}=\frac{-ja\,g_mQ_pQ_s\sqrt{(R_pR_s)}}{1+a^2}$$

$$=\frac{-j\,3\times10^{-3}\sqrt{(0\cdot81)}\times66\cdot7\times66\cdot7\times10}{1+0\cdot81}$$

$$=-j.67\cdot4$$

Magnitude $A_{res}=67\cdot4.$

(*iv*) At frequency 2*KHz* above resonant frequency

$$\delta=\frac{2}{531}=3\cdot78\times10^{-3}$$

Hence the voltage gain is given by

$$A=\frac{-j\, g_m\, a\, Q_v Q_s \sqrt{(R_p R_s)}}{1+a^2-4\delta^2\, Q_p Q_s+j2\delta\,(Q_p+Q_s)}$$

$$=\frac{-j\ 3\times 10^{-3}\sqrt{(0{\cdot}81)}\times 66{\cdot}7\times 66{\cdot}7\times 10}{1+{\cdot}81-4\,(3{\cdot}78\times 10^{-3}\times 67{\cdot}7)^2+2j\times 3{\cdot}78\times 10^{-3}\times 133{\cdot}4}$$

$$=\frac{-j\ 120}{1{\cdot}55+j\ 1{\cdot}01}$$

$$\text{Magnitude } A=\frac{120}{\sqrt{[(1{\cdot}55)^2+(1{\cdot}01)^2]}}=65.$$

EXERCISES AND PROBLEMS

1. What do you mean by radio frequency voltage amplifier ? Why are pentodes preferred for R. F. amplification ?
2. Derive an expression for the amplification produced by a tuned R. F. amplifier.
3. Explain the terms 'anode conversion efficiency' or plate efficiency used in connection with a radio frequency power amplifier. On what factors does this efficiency depend ?
4. Prove that in case of a secondary tuned transformer coup'ed amplifier, the voltage gain is given by

$$\mathbf{A}=\frac{g_m\,\omega_0\,M\,Q_e}{1+j.2\delta.Q_e}$$

where the symbols have their usual meanings.

5. Describe a double tuned transformer coupled amplifier. Obtain an expression for the voltage gain for this amplifier.
6. What do you mean by bandwidth of an amplifier ? Obtain expressions for the bandwidth in the following cases :
   (i) Single tuned capacitance coupled amplifier.
   (ii) Single tuned transformer coupled amplifier.
   (iii) Double tuned amplifier.
   Also obtain their bandwidth products.
7. What are cascaded tuned amplifiers ? Explain the decrease in bandwidth of cascaded tuned amplifiers.
8. How the bandwidth is increased in staggered tuned circuits ? Obtain expression for the overall response characteristic of the staggered tuned pair.
9. What is a class *C* operation ? Give a simple analysis of class *C* amplifier performance.
10. Write short notes on the following :
    (i) Neutralisation,
    (ii) Class *C* amplifiers as harmonic generator.
    (iii) Tank circuit efficiency.
    (iv) Valve performance at high frequencies.
11. In a tuned voltage amplifier, $L=500\mu H$, $R=10$ ohms tuning condenser $C=500\mu\mu F$, $E_i=0{\cdot}2$ volt r.m.s. at resonant frequency. Calculate (i) resonant frequency, (ii) impedance of tuned circuit at resonance, (iii) voltage

gain at resonance, (iv) output voltage at resonance, (v) Q of tuned circuit at resonance, (vi) effective Q of tuned circuit including r_p.

12. The primary and secondary windings of an *R.F.* transformer have inductances of $80\mu H$ and $20\mu H$ respectively. The mutual inductance between windings is 10 μH and their resistance is negligible. Across its secondary is connected a 40 ohm resistor.
What impedance would be measured between the primary terminals at an angular frequeney of 2×10^5 radains per second ?
Find also the value of capacitance which should be connected in series with the primary winding to make the input circuit resonant at this frequency.

13. A transformer coupled single tuned voltage amplifier uses a triode having $\mu=40$ and $r_p=15$ K ohms. The secondary winding has inductance $L_s=400\ \mu H$ and $R_s=12$ ohms. Mutual incuctance $M=40\ \mu H$ The variable tuning condenser across the secondary is adjusted to 200 $\mu\mu F$. Calculate (i) frequency of resonance f_0, (ii) Q_0 of the circuit at resonance (iii) effect.ve Q of the tuned circuit including r_p reflected to the secondary side (iv) voltage gain at resonance, (v) optimum value of mutual inductance to get the maximum voltage gain at resonance, (vi) voltage gain at resonance with optimum mutual inductance.

14. A coil of induction $400\mu H$ and resistance 5 ohms is tuned to resonance at angular frequency of $\omega=10^6$ radians per second by a capacitor in parallel with it. The resonant circuit is used as the anode load of pentode of $g_m=6{\cdot}0$ mA/V in a tuned R. F. amplifier. Draw the circuit giving an indication of all component values and calculate the voltage gain (i) at resonance, (ii) at a frequency of 10% above the resonant frequency ? Find also the bandwidth between the frequencies at which the voltage gain is $1/\sqrt{2}$ of that at resonance.
The effect of the anode slope resistance may be neglected.

14

SPECIAL AMPLIFIERS

(Wideband and D. C. Amplifers)

The wideband amplifier is the amplifier which has almost constant gain over a wide frequency range (from few cycles per sec. to few mega cycles per sec). We have studied RC coupled amplifier and its frequency response curve. The frequency response curve of a RC coupled amplifier may be improved by reducing the shunt capacity (major portion is due to interelectrode capacites) and shunt resistance (a parallel combination of r_p, R_L and R_g). We know that interelectrode capacities of a pentode are far smaller than a triode, hence to increase the frequency response a pentode is preferred. Secondly by decreasing R_L, shunt resistance may be decreased but at the same time the over all gain will be decreased. The R C coupled amplifier having high value of the ratio $g_m/(C_{pk}+C_{gk})$ may be used as a wideband amplifier. In most of the cases, to increase the upper frequency range or lower frequency range or both, compensating circuits are used. First of all we shall describe uncompensated wideband amplifier and then wideband amplifiers with high frequency and low frequency compensations. These amplifiers are used in television, radar, pulse modulation systems etc.

14·1. UNCOMPENSATED WIDEBAND AMPLIFIER :

Figure 1 shows the circuit diagram of an uncompensated wideband amplifier using a R C coupling between the stages. In the circuit C_0 is the output capacity of the amplifier tube and C_i, the input capacity of the next tube. Figure 2 shows the a c. equivalent

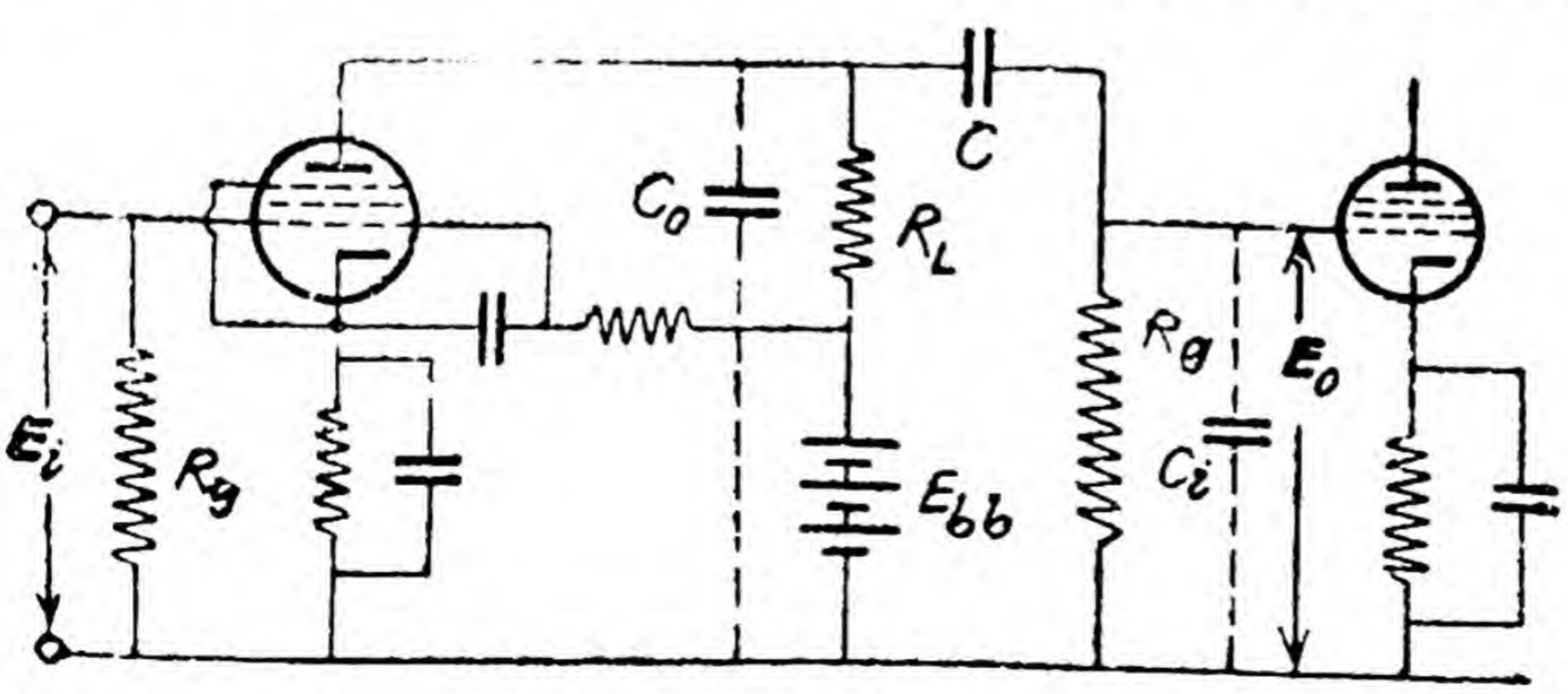


Fig .1. Uncompensated wideband amplifier.

circuit of figure 1. In the equivalent circuit C_t is the total shunt capacitance, *i.e.*, a combination of C_0 and C_i. The value of r_p is very high in case of a pentode and hence this is in shunt with R_L.

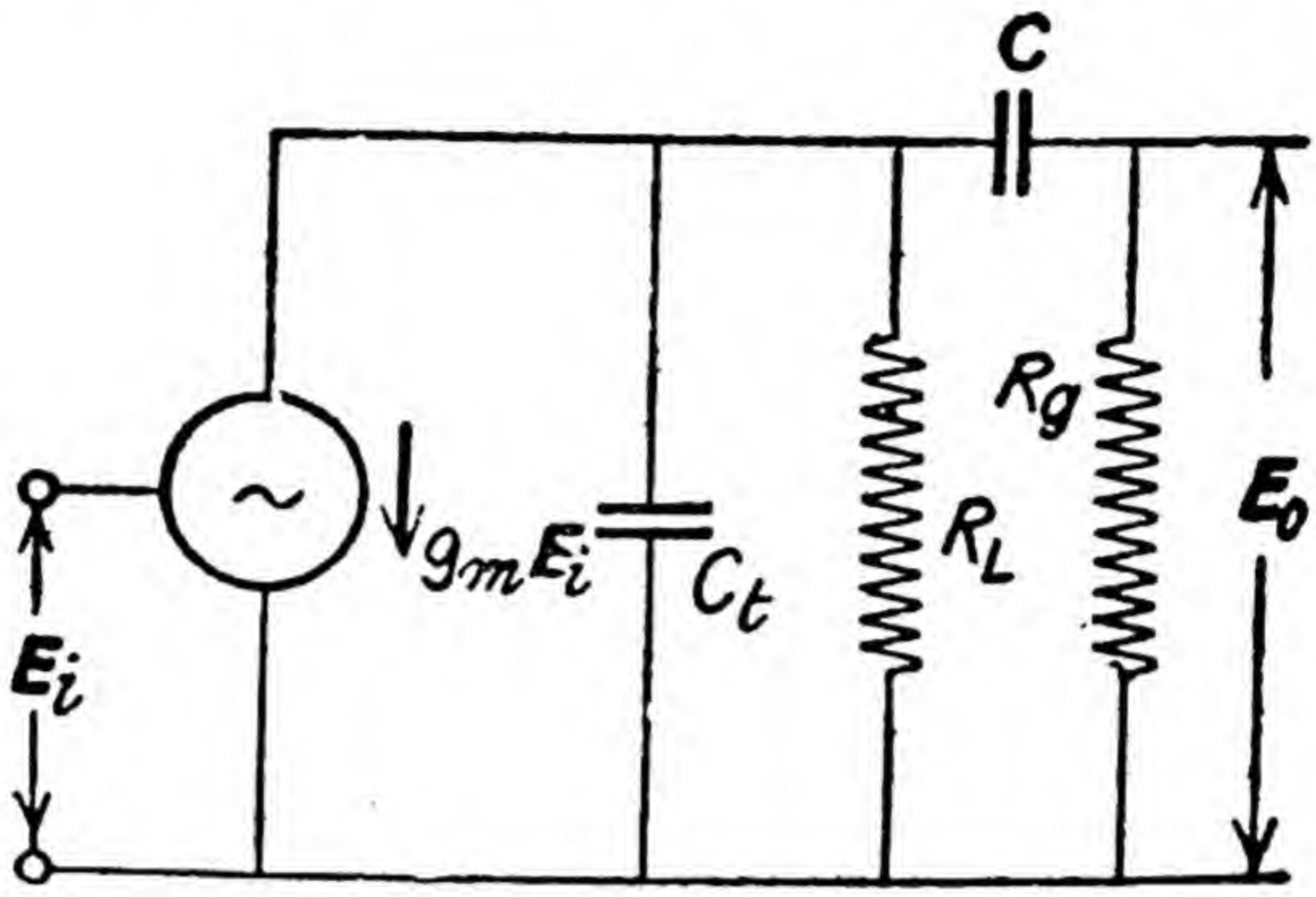


Fig. 2. A. c. equivalent circuit.

We shall study the performance of the amplifier at three frequency ranges :

(i) Middle frequency range :

In the middle frequency range the impedance of high value coupling capacitor C is very low and may be neglected. Further, the reactance of total shunt capacitance C_t is so large that it may also be omitted. The equivalent circuit of wideband amplifier in middle frequency range is given in fig. 3.

The total resistance R_t is a parallel combination of R_g and R_L and is given by

$$R_t = \frac{R_L R_g}{R_L + R_g} \quad \ldots(1)$$

The gain of the amplifier is given by

$$A_m = -g_m R_t.$$

$$|A_m| = g_m R_t. \quad \ldots(2)$$

Equation (2) shows that the gain is constant.

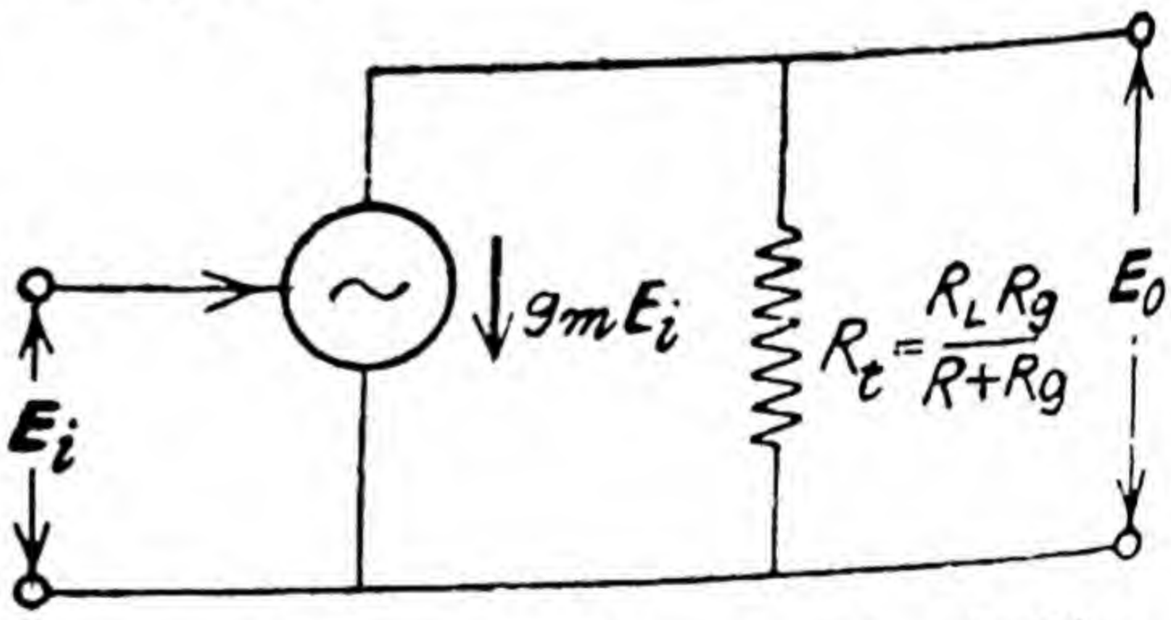


Fig. 3. A.c. equivalent circuit of wideband amplifier in middle frequency range.

(ii) High frequency range :

In the high frequency range, the reactance of coupling capacitor C is again negligibly small and hence it can be omitted. Fig. 4 gives the a.c. equivalent circuit.

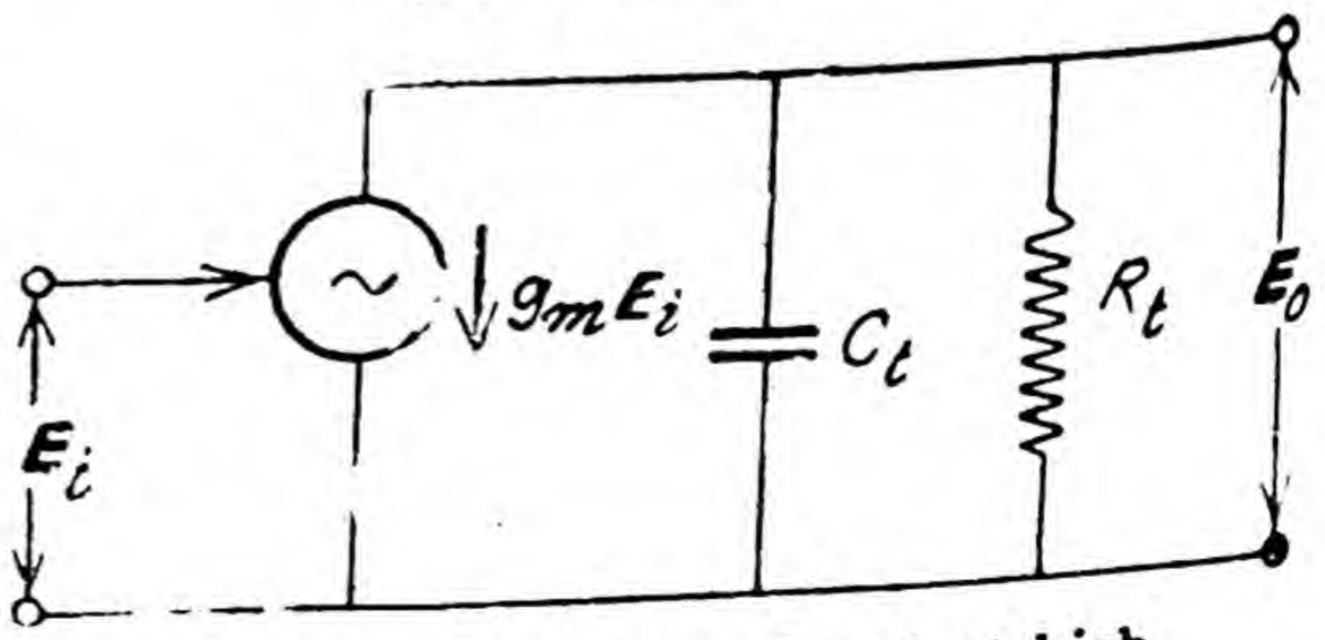


Fig. 4. A.c. equivalent circuit at high frequency range.

Tho total load impe-

dance Z_t in this case is a parallel combination of C_t and R_t.

$$\therefore \quad \frac{1}{Z_t}=\frac{1}{R_t}+\frac{1}{1/j\omega C_t}=\frac{1}{R_t}+j\omega C_t.$$

The output voltage $\mathbf{E}_0$, across R_t is given by

$$\mathbf{E}_0=g_m\mathbf{E}_tZ_t$$

$$=g_m\mathbf{E}_t\left(\frac{R_t}{1+j\omega C_tR_t}\right), \qquad \ldots(3)$$

The voltage gain at high frequency range is given by

$$\mathbf{A}_h=\frac{\mathbf{E}_0}{\mathbf{E}_t}=-g_m\mathbf{Z}_0$$

$$=\frac{-g_mR_t}{1+j\omega C_tR_t}$$

$$=\frac{-g_mR_t}{1+j2\pi f.C_tR_t}.$$

where f is the frequency of applied signal.

$$|A_h|=\frac{g_mR_t}{\sqrt{\{1+(2\pi fC_tR_t)^2\}}}. \qquad \ldots(4)$$

The equation (4) shows that for given values of R_t and C_t the response of the amplifier is a function of frequency.

The upper cross over frequency in this range is defined as

$$f_2=\frac{1}{2\pi C_iR_t}$$

$$\therefore \quad \mathbf{A}_h=\frac{-g_mR_t}{1+jf/f_2}$$

or

$$|A_h|=\frac{g_mR_t}{\sqrt{\{1+(f/f_2)^2\}}}. \qquad \ldots(5)$$

The relative gain is given by

$$\frac{A_h}{A_m}=\frac{1}{\sqrt{\{1+(f/f_2)^2\}}}. \qquad \ldots(6)$$

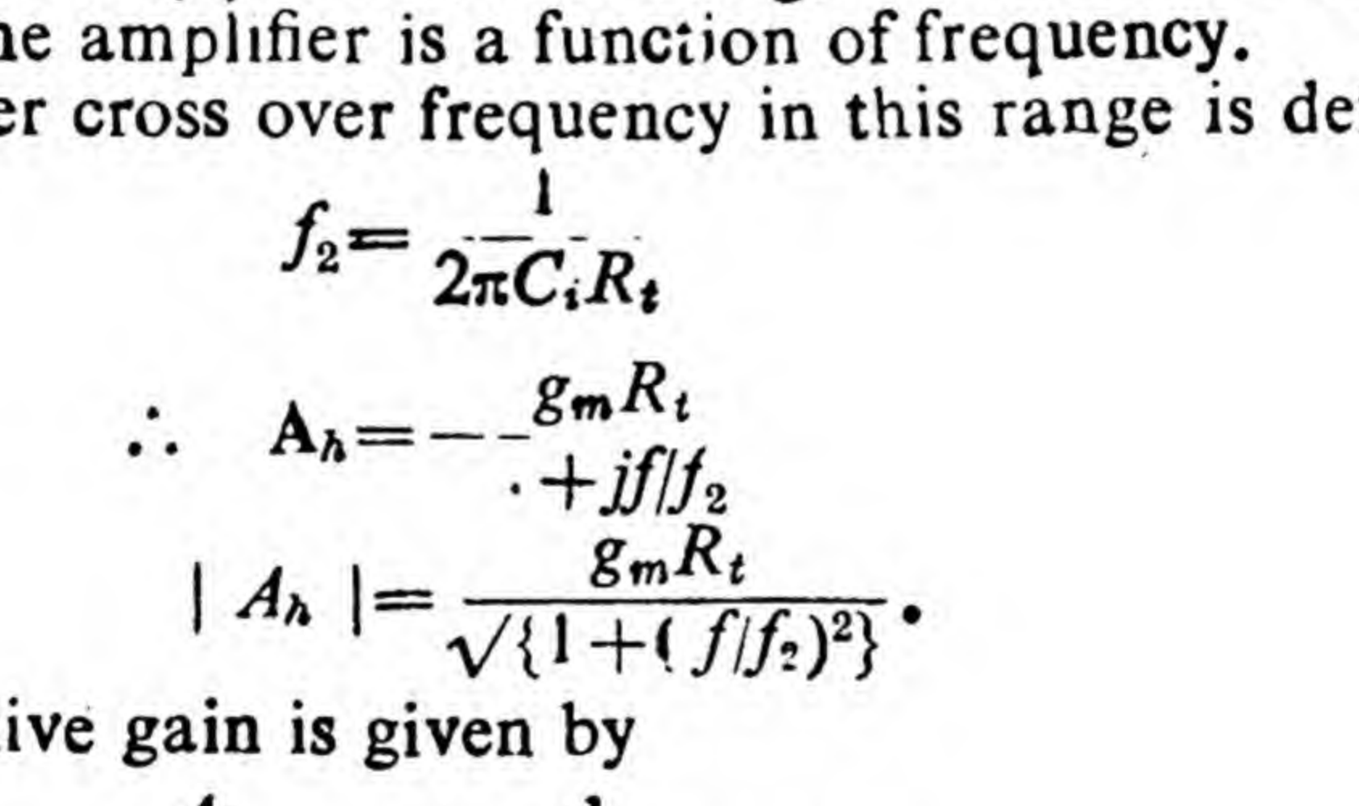


Fig .5 (a). Gain versus frequency characteristic of uncompensated and ideal wideband amplifier.

This equation shows that as the frequency f increases relative to the crossover frequency f_2, the gain A_h decreases. In figure 5 (*a*), a plot of relative gain versus relative frequency f/f_2 in case of wide band amplifier for high frequency range is shown. It is clear from the figure that the gain remains constant upto a relative frequency $f/f_2=0{\cdot}1$ and becomes ·707 *i.e.* $(1/\sqrt{2})$ at $f/f_2=1$.

For an uncompensated wideband amplifier, the components should be so chosen that the reference frequency f_2 is as large as possible. In order to maximise f_2, C_t and R_t should be made as small as possible because $f_2=1/(2\pi\ C_t R_t)$.

There are limits to which the value of R_t or C_t or both may be reduced. If R_t is made equal or less than $1/g_m$, then the gain of the amplifier is less than unity and it fails to function as a voltage amplifier. The total shunt capacitance C_t is a combination of several capacitances as given below

$$C_t=C_s+C_{pk}+C_{gk}+C_{gp}\ (1+A),$$

where C_s is the total stray capacitance to ground of wiring, tube sockets etc. and A is the voltage gain of the following stage. In pentodes C_{gp} is extremely small so that the term $C_{gp}\ (1+A)$ may be neglected C_{pk} and C_{gk} have their usual meanings

The phase shift in wideband amplifier is given by

$$\phi=180°-\tan^{-1}\ (2\pi fC_tR_t)$$
$$=180°-\tan^{-1}\ (f/f_2). \qquad \ldots(7)$$

In equation (7), 180° is a constant phase shift and $\tan^{-1}\ (f/f_2)$ is the additional phase shift. Fig. 5 (*b*) shows the variation of additional phase shift as a function of relative frequency f/f_2.

For values f/f_2 upto about unity, the additional phase shift is almost proportional to frequency, f itself because for small values of angle, tangent of an angle equals the angle itself.

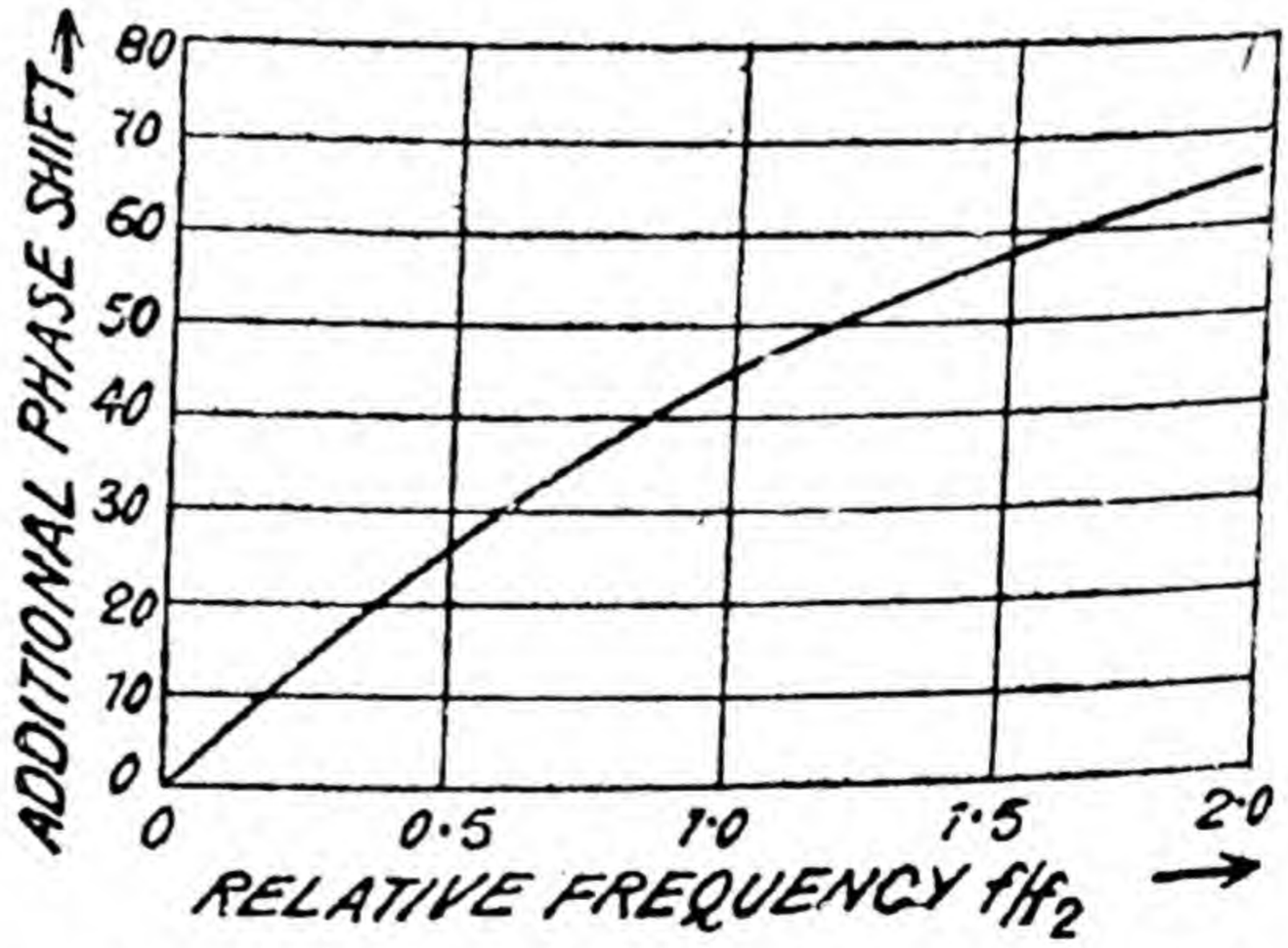


Fig. 5 (b). Additional phase shift versus relative frequency.

(iii) Low frequency range :

In low frequency range, the reactance of the shunt capacitance C_t being very low and may be neglected. In this case coupling capacity C cannot be neglected. In low frequency range, the equivalent circuit is shown in figure 6.

The input voltage E_t results in the output voltage E'_0 across

the load resistance R_L. A fraction K of E_0 appears across the grid leak resistance R_g to constitute the final voltage E_0.

$$K=\frac{R_g}{R_g+\frac{1}{j\omega C}}$$

$$=\frac{1}{1+\frac{1}{j\omega CR_g}} \quad ...(8)$$

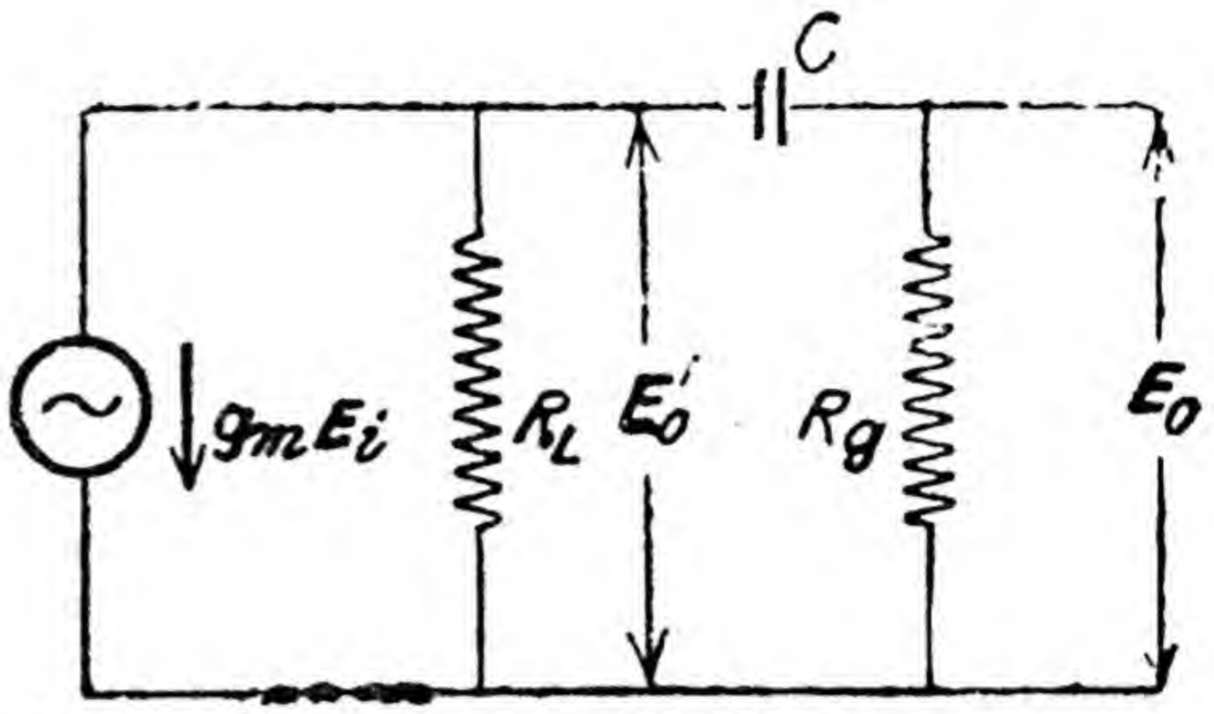


Fig 6. A.c. equivalent circuit of wide-band amplifier in low frequeny range.

The magnitude of K is given by

$$K=\frac{1}{\sqrt{\left\{1+\left(\frac{1}{\omega CR_g}\right)^2\right\}}} \quad ...(9)$$

Let f_1 denote the reference frequency such that

$$f_1=\frac{1}{2\pi CR_g} \quad ...(10)$$

Now $\mathbf{K}=\frac{1}{1-jf_1/f}$ and $K=\frac{1}{\sqrt{\{1+(f_1/f)^2\}}}$ ...(11)

The voltage $\mathbf{E_0}'$ developed across R_L is given by

$$\mathbf{E_0}'=-g_m\mathbf{E}_i R_L,$$

$$\mathbf{E}_0=\mathbf{E_0}'\mathbf{K}$$

or $$\mathbf{E}_0=-g_m\mathbf{E}_i R_L\mathbf{K}. \quad .. (12)$$

The voltage gain in this frequency range is given by

$$\mathbf{A}_i=\frac{\mathbf{E}_0}{\mathbf{E}_i}=-g_m R_L\mathbf{K}$$

$$=-g_m R_L\ \frac{1}{1-\frac{j}{2\pi fCR_g}}=\frac{-g_m R_L}{1-j(f_1/f)}$$

$$|A_i|=g_m R_L\ \frac{1}{\sqrt{\{1+(f_1/f)^2\}}}=\frac{g_m R_L}{\sqrt{\left\{1+\left(\frac{1}{2\pi CR_g}\right)^2\right\}}}. \quad ...(13)$$

The phase angle of voltage gain is given by

$$\phi=180°+\tan^{-1}(f_1/f)=180°+\tan^{-1}\left(\frac{1}{2\pi fCR_g}\right)$$

$$=180°+\cot^{-1}(f/f_1). \quad .. (14)$$

A plot of relative voltage gain $A_i/g_m R_L$ of wideband amplifier against relative frequency f/f_1 in low frequency range is shown in figure 7. In order to have a uniform voltage gain upto very low frequencies, the frequency f_1 should be made as small as possible.

We know that the value of the reference frequency is given by the following expression :

$$f_1 = \frac{1}{2\pi R_g C}.$$

Hence by employing the large value of R_g or C or both f_1 may be decreased. In practice the maximum values of R_g or C are limited. In most of the cases, the values of R_g and C are $1 M\Omega$ and $\cdot 05 \mu F$ respectively. Substituting these values, the reference frequency f_1 comes out to be 3·2 c/s.

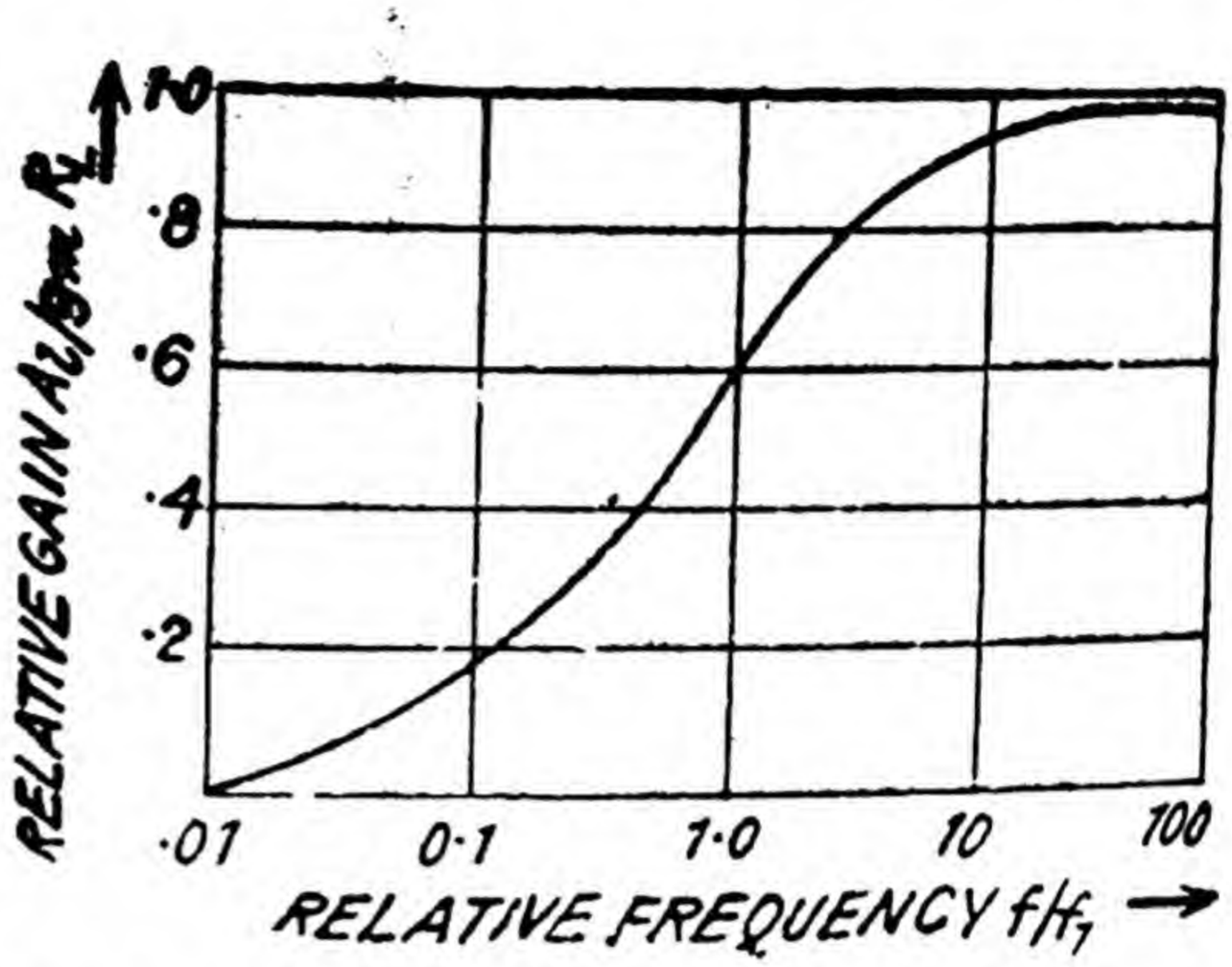


Fig 7. Relative gain versus relative frequency in low frequency range.

The additional phase shift in this frequency range is given by

$$\phi_a = \tan^{-1}(f_1/f) = \tan^{-1}\left(\frac{1}{2\pi f C R_g}\right)$$
$$= \cot^{-1}(f/f_1).$$

In fig. 8, additional phase angle ϕ_a is plotted against the relative frequency (f/f_1).

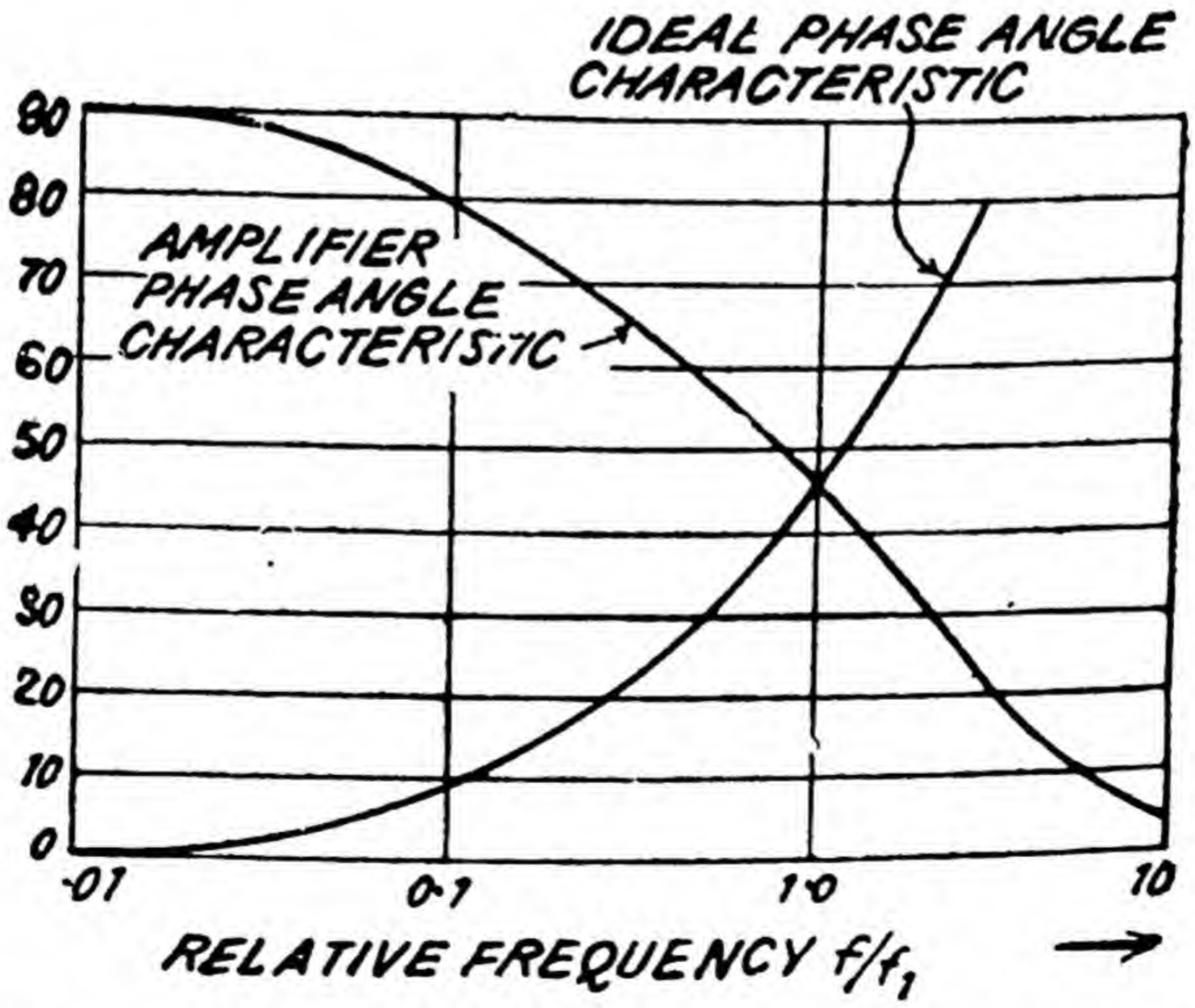


Fig. 8. The additional phase angle versus relative frequency f/f_1.

Figure 8 shows that the phase angle characteristic of the amplifier is opposite to the ideal curve. Thus the phase response of uncompensated wideband amplifier is poor and needs low frequency compesation.

14·2. HIGH FREQUENCY COMPENSATION IN WIDEBAND AMPLIFIERS :

In the above article we have considered the gain equations in an uncompensated wideband amplifier. These equations are like those used in RC coupled amplifier. Now our object is to consider the high frequency compensation. The following three methods are used : (*i*) Shunt peak high frequency compensation, (*ii*) Series peaked high frequency compensation and (*iii*) Shunt series high frequency compensation. The high frequency compensation consists in introducing an additional inductance of 100 to 500μH either in series with the load resistor or in series with the coupling capacitor or both the positions. The inductor resonates with total shunt capacitance C_t at a frequency somewhat above the upper limit of the video range resulting in a rising gain characteristic. Here we shall discuss the series peaked high frequency compensation.

Series peaked high frequency compensation :

At high frequencies, the impedance offered by the shunt capacitance is compar able to the load and hence it is of great importance. The shunt capacitance consists of two parts : (*i*) input capacitance of the following stage and (*ii*) output capacitance C_0 of the given stage. Figure 9 shows the circuit of wide band amplifier

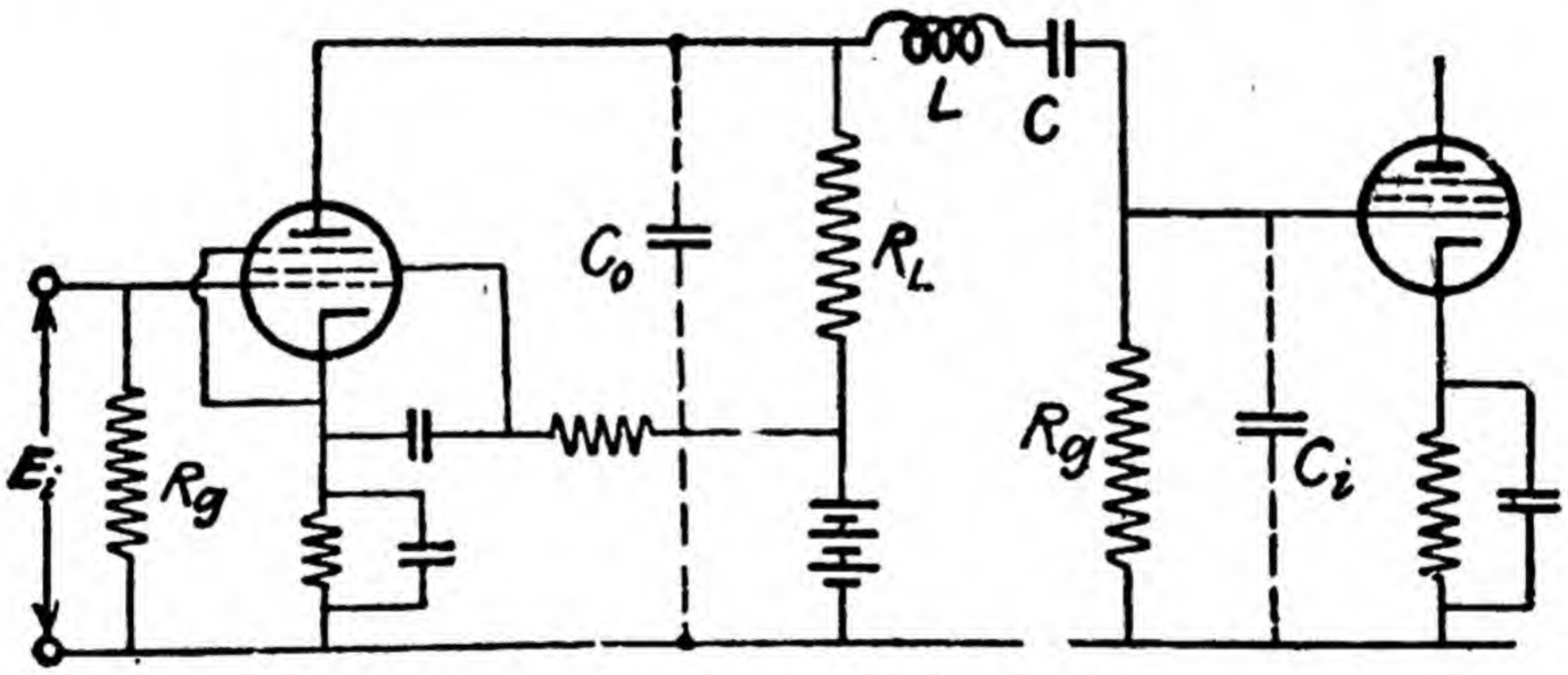


Fig. 9. Wideband amplifier with series peaked high frequency compensation.

using series compensation inductor L. The inductor isolated the two shunting capacitances and hence reduces the shunting capacitances across the amplifier tube. Thus for the given upper frequency, the gain is increased. Fig. 10 shows the equivalent circuit. The coupling capacitor C has been omitted from the a.c. equivalent circuit since its reactance in high frequency range is extremely small.

The impedance Z_1 of R_L and C_0 in parallel is given by

$$\frac{1}{Z_1}=\frac{1}{R_L}+\frac{1}{1/j2\pi fC_0}=\frac{1}{R_L}+j2\pi fC_0$$

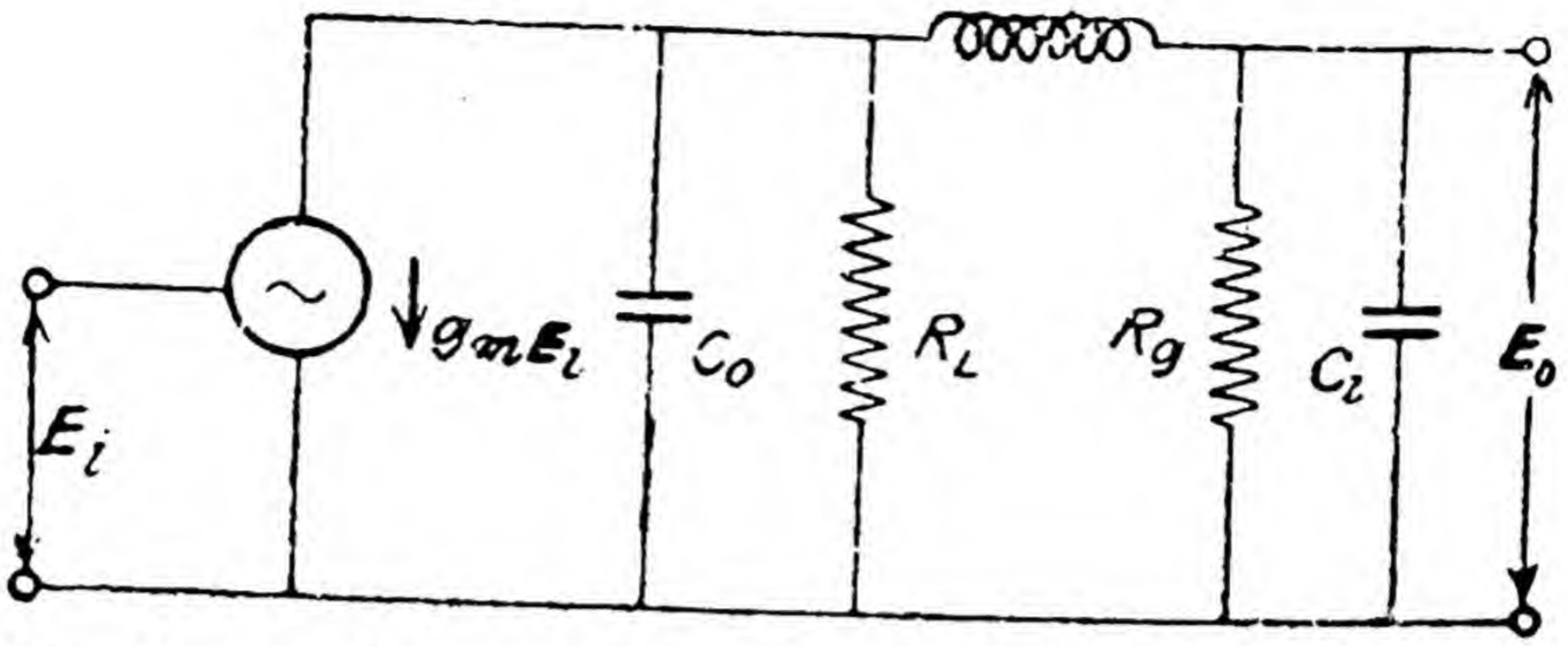


Fig. 10. A.c. equivalent circuit of series peaked wideband amplifier.

$$\therefore \quad Z_1 = \frac{R_L}{1 + j2\pi f C_0 R_L}. \qquad \ldots(1)$$

The voltage E_1 across Z_1 will be

$$E_1 = g_m E_i Z_1 = \frac{g_m E_i R_L}{1 + j2\pi f C_0 R_L}. \qquad \ldots(2)$$

This expression is the same as in uncompensated amplifier except the effective shunt capacitance is only C_0 instead of $C_t = (C_0 + C_i)$.

Considering the circuit only upto R_L, the upper crossover frequency is given by

$$f_0 = \frac{1}{2\pi R_L C_0}. \qquad \ldots(3)$$

Equation (3) shows that f_0 is higher than in case of uncompensated amplifier.

This voltage is now passed through the reactive voltage divider consisting of L and C_i. Here we have neglected R_g because it is larger in comparison with the reactance of C_i in shunt. The value of the large inductance is so selected that it resonates with C_i at frequency f_0. Due to the series resonance, the voltage across C_i tends to rise with frequency. This rising tendency compensates for the drop in voltage with frequency across R_L. Thus the frequency to constant gain increases beyond f_c. In this way the isolation of shunt capacitances C_0 and C_i results in the upper frequency limits.

Ex. *Design a video amplifier using a 6AK5 pentode such that both the uniform gain and the phase shift proportional to frequency are important upto 4 megacycles. The tube constants are :* $g_m = 5100$ *micro mhos.* $C_{out} = 2{\cdot}8$ *pf. and the wiring capacity and the input capacity of the next stage are* $9{\cdot}7$ *pf.,* $r_p = 0{\cdot}8$ *mega ohms.*

(a) Draw the circuit diagram for the above amplifier in which both high and low frequency compensation are made.

(b) Determine the values of the inductance L, the load resistance R_L *and the mid frequency gain.*

(*c*) *The lower cut off frequencies before and after compensation are* 20 *cycles and* 1·5 *cycles respectively. Choose the grid leak resistance* R_g *equal to* 0·4 *meg-ohm and find out the values of* C_c, C_b *and* R_b.

(*a*) The circuit diagram of a compensated amplifier at both frequencies is shown in figure 11.

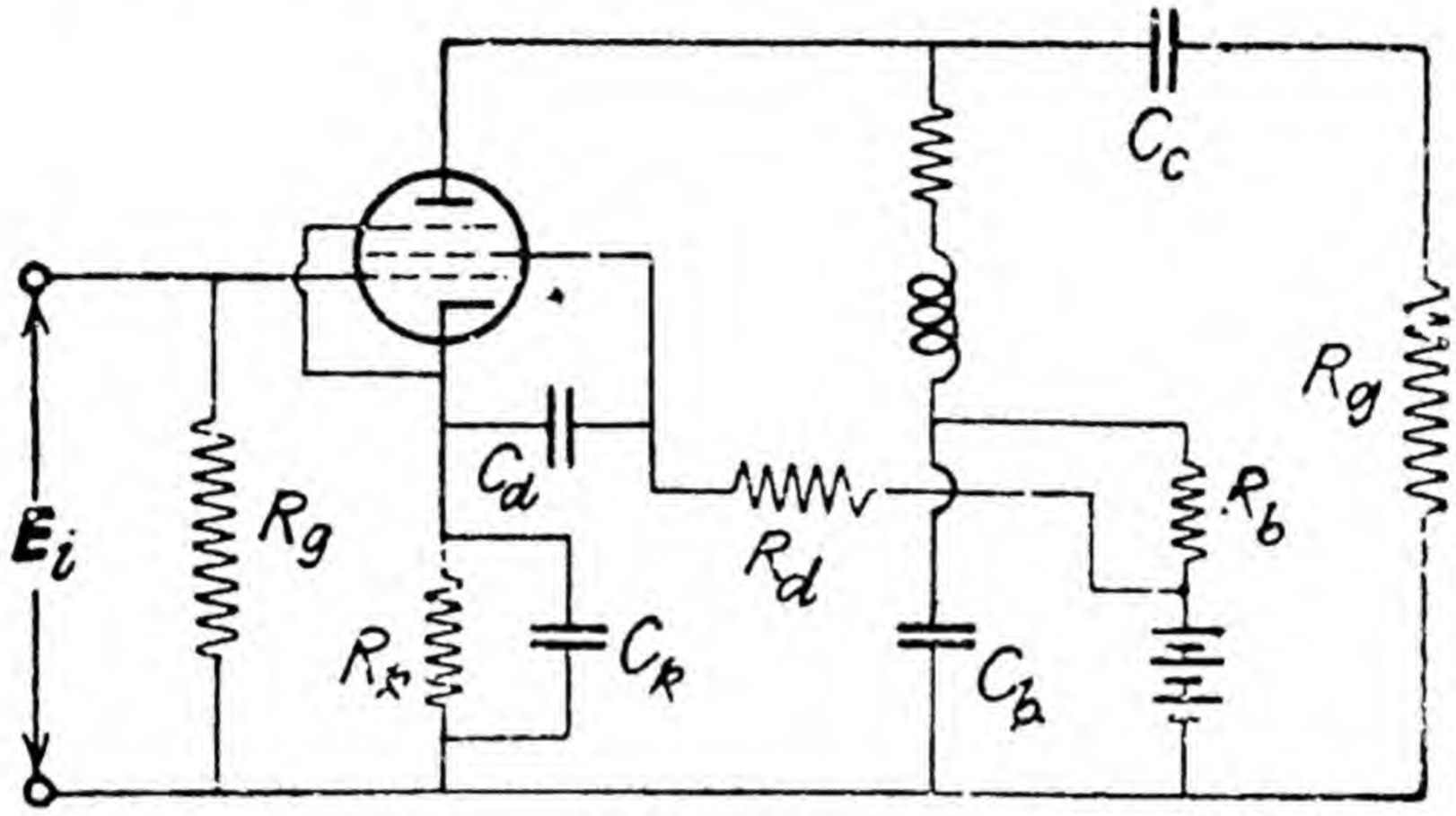


Fig. 11. Broad band amplifier with both l. f. and h. f. compensation.

(*b*) The uniform gain and phase shift proportional to frequency are important factors for determining the value of Q, a compromise value of $Q=0{\cdot}44$ is taken. At this value the gain remains practically equal to the mid frequency gain when f/f_2 is equal to 6·66 and $f=4$ megacycles.

$$\therefore \quad f/f_2=\frac{4\times 10^6}{f_2}=0{\cdot}66.$$

The upper cut off frequency f_2 of the uncompensated amplifier is

$$f_2=\frac{4\times 10^6}{0{\cdot}66}=6\times 10^6 \text{ cycles.}$$

We know that $f_2=\dfrac{1}{2\pi C_t R_L}$, where $C_t=C_{out}+C_{wiring}+C_{in}$ for the next stage $=2{\cdot}8+9{\cdot}7=12{\cdot}5$ pf.

$$\therefore \quad 6\times 10^6=\frac{1}{2\pi\times 12{\cdot}5\times 10^{-12}\times R_L}$$

or $$R_L=2100 \text{ ohms.}$$

Again $$Q=0{\cdot}44=\frac{2\pi f_2 L}{R_L}=\frac{2\times 3{\cdot}14\times 6\times 10^6 L}{2100}$$

or $$L=\frac{0{\cdot}44\times 2100}{2\times 3{\cdot}14\times 6\times 10^6}=21{\cdot}3\times 10^{-6}$$

$$=21{\cdot}3 \text{ micro-henry.}$$

The mid frequency gain is given by

$$A_{mid} = -g_m R_L = -5100 \times 10^{-6} \times 2100 = -10{\cdot}5.$$

(*c*) The lower cut off frequency of the uncompensated amplifier f_1 is given by

$$f_1 = \frac{1}{2\pi C_c R_g}$$

$$20 = \frac{1}{2 \times 3{\cdot}14 \times C_c \times 0{\cdot}4 \times 10^6}$$

$$C_c = \frac{1}{2 \times 3{\cdot}14 \times 20 \times 0{\cdot}4 \times 10^6}$$

$$= 0{\cdot}02 \text{ micro farad}$$

We know that

$$C_c R_g = C_b \frac{R_L R_b}{R_L + R_b}$$

$$0{\cdot}02 \times 10^{-6} \times 0{\cdot}4 \times 10^6 = C_b \frac{2100 R_b}{2100 + R_b} \quad \ldots(1)$$

and

$$f_3 = \frac{1}{2\pi C_b R_b}$$

i.e.

$$C_b R_b = \frac{1}{2\pi f_3} = \frac{1}{2\pi \times 3{\cdot}14 \times 1{\cdot}5} \cdot \quad \ldots(2)$$

Solving (1) and (2), we get

$R_b = 23$ kilo-ohms and $C_b = 14$ micro-farad.

14·3. DIRECT COUPLED AMPLIFIERS : LOFTIN-WHITE FORM :

Amplifiers in which no coupling reactive element, (*e.g.* if coupling capacitor in *RC* coupling be removed), is employed and output of first stage, is connected to the input of the next stage as shown in figure 12, are called direct coupl d amplifiers. Elimination of coupling condesers extends the limit of amplification to sufficiently low frequencies and therefore, by such amplifiers, direct voltage and voltages of extremely low frequencies can be amplified. Further, due to the absence of reactive coupling elements response of amplifier is almost free from the frequency distortion and phase shift upto those frequencies at which the internal tube and stray shunt capacities become effective.

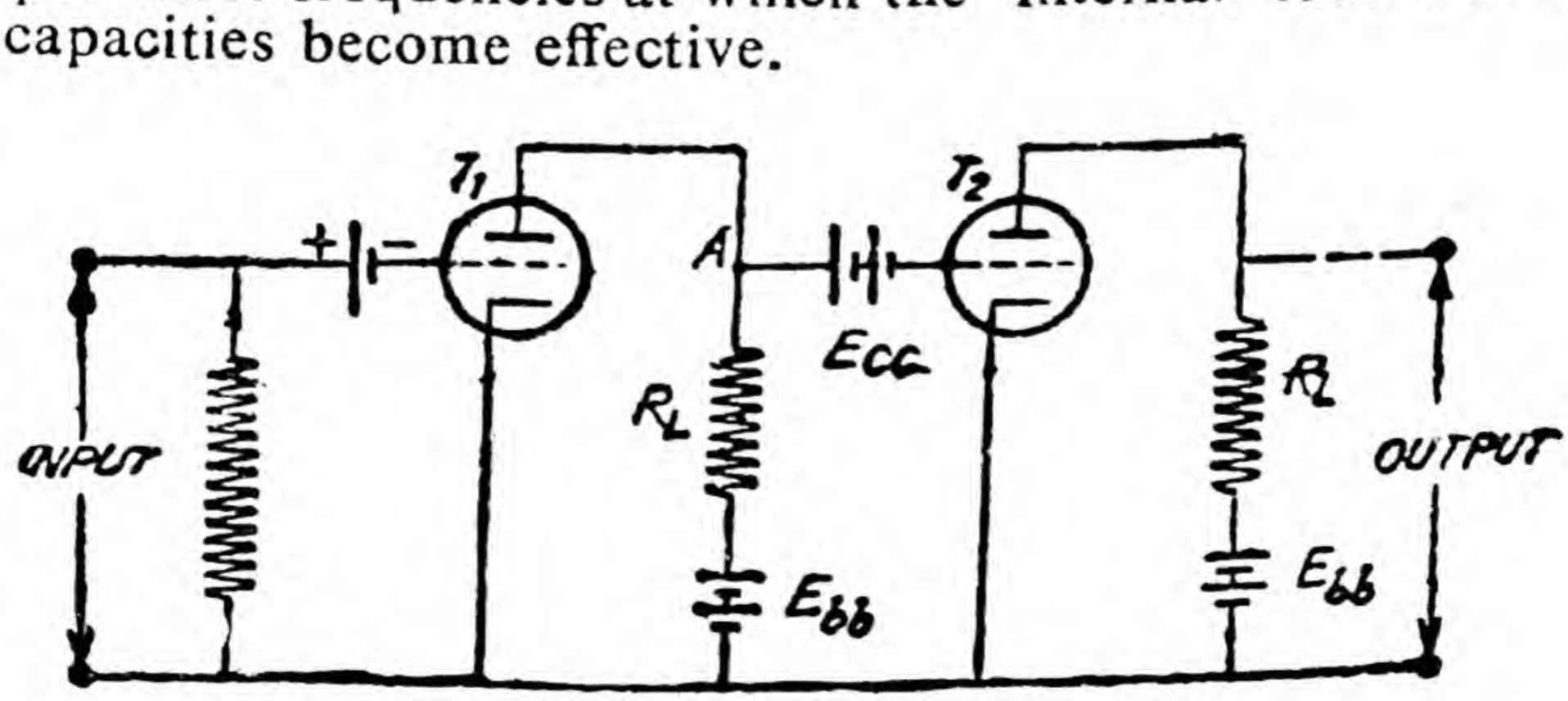


Fig. 12. Two stage direct-coupled amplifier.

Magnitude of bias battery E_{cc} through which the plate of tube T_1 is connected to the grid of tube T_2 can be estimated quite easily. Suppose $E_{bb}=250$ volts, average plate current through R_L is $1{\cdot}6mA$ and R_L is 10^5 ohms. then drop across R_L will be 160 volts that means potential of point A is $+90$ volts with respect to the ground. Now if the operating point with respect to the ground is to be at -10 volts (*i.e.* the grid potential of tube T_2 for amplifying action) E_{cc} must be -100 volts which is a large battery as compared to those commonly applied for biasing purposes. Since the circuit requires three such batteries of high voltage, Loftin-White suggested the use of single *dc* source as shown in figure 13 where different voltages are tapped on the bleeder resistor.

In fig. 13, tube T_1 is biased to -10 volts, voltage of T_1 is $150-10=+140$ volts. Drop across R_L is 70 volts so that point A is at $+70$ volts. But cathode of T_2 is kept at $+80$ volts, giving a bias of -10 volts to tube T_2. Net voltage across tube T_2 is $250-80=170$ volts. If the voltage across R_L is 70 volts then point B is at $+100$ volts. Output is then

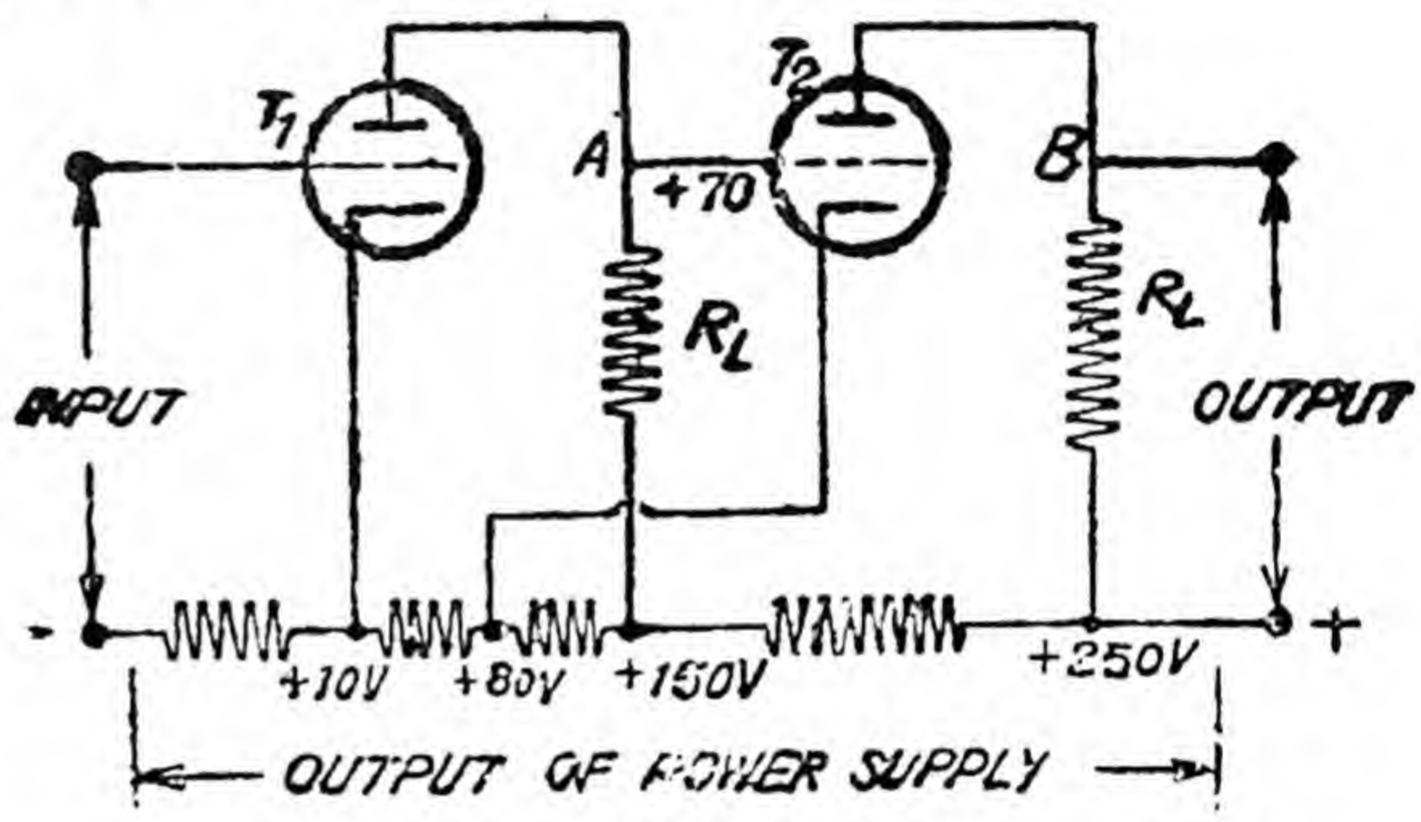


Fig. 13. Loftin-white form d c. amplifier.

$$=+100-(+250)$$
$$=-150 \text{ volts.}$$

Such amplifiers suffer from the following drawbacks :

(*i*) Output voltage changes due to the change in heater, grid bias or plate voltage and thus reflects the inherent instability of d.c. amplifiers, being called as drift in d.c. amplifiers. If supply voltage changes, bias to the tubes will change and hence a change in the operating point or in other words shift in operating point will be the same as caused by the application of d.c. signal at the input and therefore one cannot tell whether a change in output is due to a change in bias or a change in signal.

(*ii*) Since in circuit of fig. 12 E_{cc}, grid bias battery is ungrounded, a high shunt capacity between the grid bias battery and ground also appears.

(*iii*) Neither terminals of R_L is at ground potential which develops the problem of reference or zero voltage when no signal is input to the amplifier and the latter is being used to operate a meter for the measurement of d.c. voltages.

(*iv*) In Loftin White circuit, common coupling and feed back through the tapped supply resistor affects the performance of the amplifier. To overcome these draw-backs of Loftin-White circuit, the following two simple forms of d.c. amplifiers are used.

(1) Resistance Coupled d.c. Amplifier : The circuit is shown in fig. 14. Power supplies negative to the ground (E_{cc}) and other positive to the ground (E_{bb}) are required. Though points of E_{bb} for both tubes are tapped at one point on the supply bleeder resistor but for the sake of analysis we have shown them separately. Similar is the case of E_{cc} in practice. Potential of B depends upon the potential at point A which in turn is a function of bias voltage. Thus by changing R_k, which alters the grid bias, B is brought to ground or midway between the two power supply voltages for zero input signal and output signal, and the output is zero.

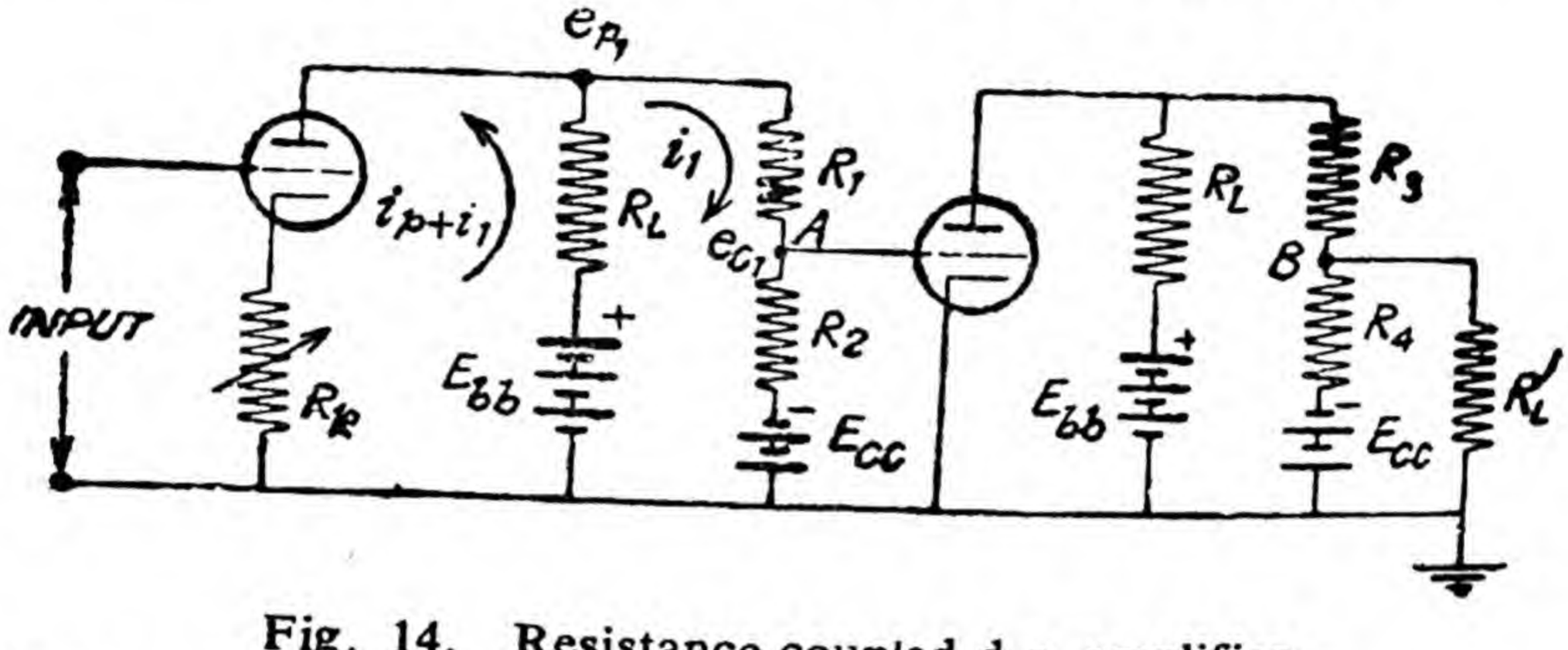


Fig. 14. Resistance coupled d.c. amplifier.

Suppose e_{p_1} and e_{c_1} are the potential levels above the ground as reference for the plate of first stage and the grid of second stage respectively, then we write the voltage relations as

$$(i_1+i_p)\ R_L=E_{bb}-e_{p_1} \qquad ...(1)$$

$$e_{p_1}-e_{c_1}=iR_1 \qquad ...(2)$$

and

$$e_{c_1}-E_{cc}=i_1R_2. \qquad ...(3)$$

Putting for i_1 in equation (1) from equation (2),

$$\left(\frac{e_{p_1}-e_{c_1}}{R_1}+i_p\right) R_L=E_{bb}-ep_1$$

$$\frac{(e_{p_1}-e_{c_1})}{R_1}.R_L=E_{bb}-e_{p_1}-i_pR_L$$

$$\frac{R_1}{R_L}=\frac{e_{p_1}+e_c}{E_{bb}-e_{p_1}-i_pR_L}. \qquad ...(4)$$

Similarly from equation (1) and (3),

$$\frac{R_2}{R_L}=\frac{e_{c_1}-E_{cc}}{E_{bb}-e_{p_1}-i_pR_L}. \qquad ...(5)$$

From equations (4) and (5), values of R_1 and R_2 can be obtained. A suitable value of R_1 and R_2 will provide the correct bias to the tube T_2. Similarly an appropriate choice of R_3 and R_4, will enable to mark zero output across R_L in the absence of any input signal.

(2) Cathode Coupled d.c. Amplifier : Circuit with its equivalent is shown in fig. 15. Two triodes are required and the commpn plate supplies are employed. The circuit amplifies the

signal with considerable gain, an expression for which can be deduced as follows :

For equivalent circuit, we write

$$e_{g_1}=e_i-(i_{b_1}+i_{b_2})\,R_K \qquad \ldots(1)$$

$$e_{g_2}=-(i_{b_1}+i_{b_2})\,R_K \qquad \ldots(2)$$

$$\mu\, e_{g_1}=i_{b_1}\,(r_p+R_K)+i_{b_2}\,R_K \qquad \ldots(3)$$

$$\mu\, e_{g_2}=i_{b_2}\,(r_p+R_K)+R_L\,(i_{b_1}+i_{b_2}) \qquad \ldots(4)$$

Putting e_{g_1} from equation (1) in equation (3),

$$\mu\, e_i-\mu\,(i_{b_1}+i_{b_2})\,R_K=i_{b_1}(r_p+R_K)+i_{b_2}\,R_K$$

$$\mu\, e_i=i_{b_1}\,(\mu\,R_K+R_K+r_p)+i_{b_2}\,(\mu\,R_K+R_K)$$

$$=i_{b_1}\,[(\mu+1)\,R_K+r_p]+i_{b_2}\,R_K\,(\mu+1). \qquad \ldots(5)$$

Putting e_{g_2} from equation (2) in equation (4),

$$-\mu\,(i_{b_1}+i_{b_2})\,R_K=i_{b_2}\,(r_p+R_L+R_K)+i_{b_1}\,R_K$$

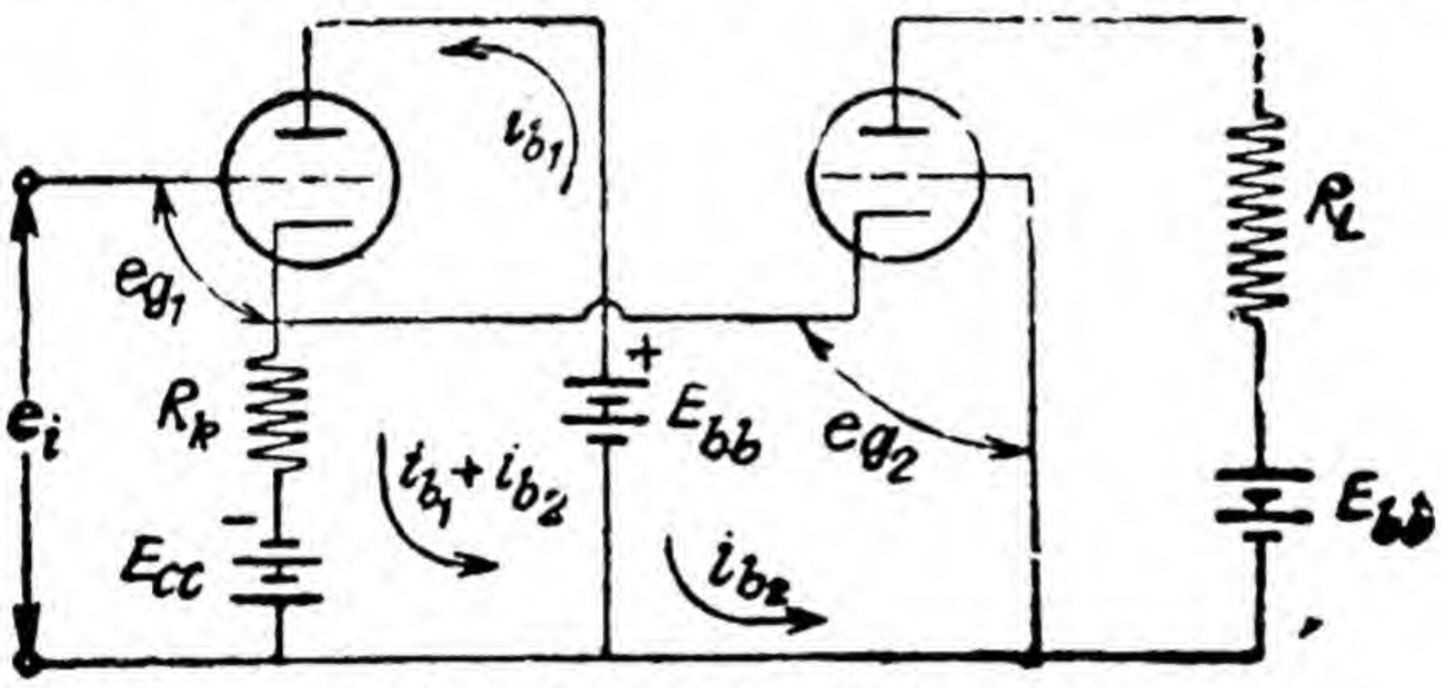


Fig. 15. Cathode coupled d.c. amplifier.

or
$$i_{b_1}=-i_{b_2}\cdot\frac{r_p+R_L+(\mu+1)\,R_K}{(\mu+1)\,R_K} \qquad \ldots 6)$$

Putting i_{b_2} from equation (6) in equation (5), we get

$$\mu e_i=-i_{b_2}\left(\frac{r_p+R_L+(\mu+1)\,R_K}{(\mu+1)\,R_K}\right)[(\mu+1)\,R_K+r_p]$$

$$+i_{b_2}\,(\mu+1)\,R_K$$

or
$$\mu\,(\mu+1)\,R_K\,e_i=-i_{b_2}\,[\{r_p+(\mu+1)\,R_K\}\,\{r_p+R_L+(\mu+1)R_K\}$$

$$-(\mu+1)^2\,R_K{}^2]. \qquad \ldots(7)$$

Therefore gain of the circuit is

$$A=\frac{e_0}{e_i}=\frac{-i_{b_2}R_L}{e_i}$$

$$=\frac{\mu(\mu+1)R_K R_L}{\{r_p+(\mu+1)R_K\}(r_p+R_L+(\mu+1)R_K\}-(\mu+1)^2 R_K^2}$$

$$A=\frac{\mu R_L}{\left(1+\frac{r_p}{(\mu+1)R_K}\right)\{r_p+R_L+(\mu+1)R_K\}-(\mu+1)R_K}$$

$$=\frac{\mu R_L}{\left[\frac{r_p+R_L}{(\mu+1)}\right]\frac{r_p}{R_K}+2r_p+R_L}.$$

Neglecting the first term, as it is small for usual circuit and tube parameters, in the denominator we find gain

$$A=\frac{\mu R_L}{2r_p+R_L}.$$

14·4. DRIFT IN D.C. AMPLIFIERS :

Drift can be minimised by regulating the voltages applied to the amplifier. Regulated power supplies are employed to provide the regulated voltage to the plate and bias electrodes. Heater voltage is not regulated and thus drifts the output voltage with its own variation. The following two circuits to reduce the drift are in most common use :

(1) Balanced push pull amplifier : In this circuit two identical tubes T_1 and T_2 are balanced against each other. With such tubes, having matched characteristics, charges in the plate supply voltage have no effect on the output of the amplifier. In fig. 16, M is a meter to read the output of the amplifier.

The grid voltage can be written as

$$e_{g_1}=e_i-(i_1+i_2)R_K, \quad ...(1)$$

$$e_{g_2}=-(i_1+i_2)R_K. \quad ...(2)$$

For plate circuit equations, we write

$$\mu e_{g_1}=i_1(R_L+r_p+R_K)+i_2 R_K, \quad ...(3)$$

$$\mu e_{g_2}=i_2(R_L+r_p+R_K)+i_1 R_K,$$

From equations (1) and (3),

$$\mu e_i=i_1[R_L+r_p+(\mu+1)R_K]+(\mu+1)R_K i_2. \quad ...(5)$$

From equations (2) and (4),

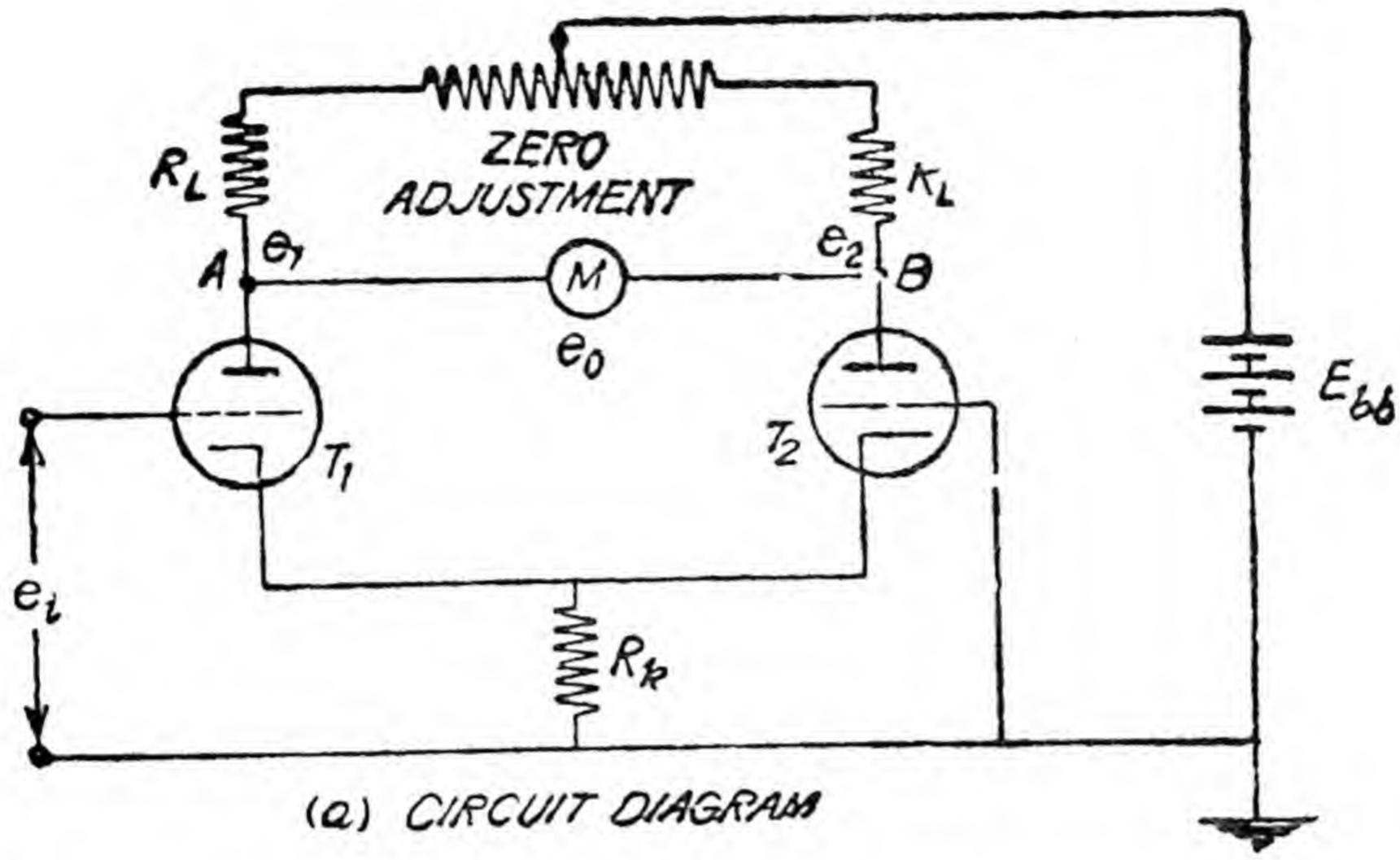
$$-(\mu+1)R_K i_1=i_2[R_L+r_p+(\mu+1)R_K]. \quad ...(6)$$

From equations (5) and (6),

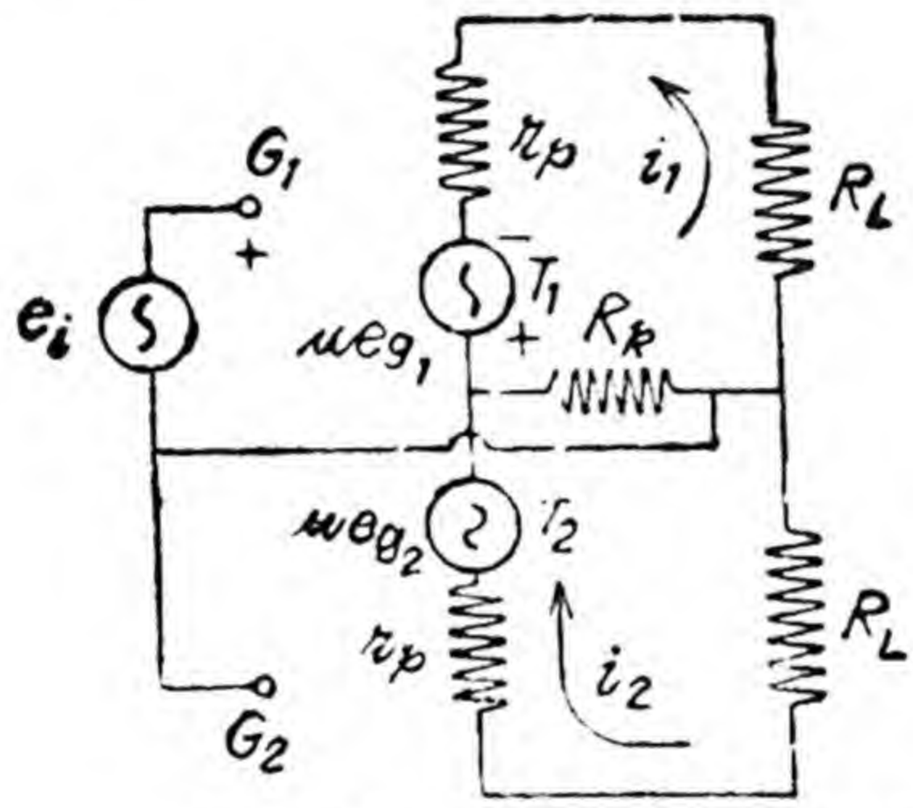
$$i_1=\frac{\mu e_i[R_L+r_p+(\mu+1)R_K]}{2(R_L+r_p)[1+(\mu+1)R_K]}$$

and

$$i_2=\frac{-\mu(\mu+1)e_i R_K}{2(R_L+r_p)[1+(\mu+1)R_K]}$$



(a) CIRCUIT DIAGRAM



(b) EQUIVALENT CIRCUIT

Fig. 16. Balanced pushpull amplifier.

The potentials at A and B are

$$e_2 = E_{bb} - i_2 R_L,$$
$$e_1 = E_{bb} - i_1 R_L,$$

and therefore reading of meter M is

$$e_M = e_2 - e_1$$
$$= (i_1 - i_2) R_L$$
$$= \frac{\mu e_i R_L [R_L + r_p + 2(\mu+1) R_K]}{2(R_L + r_p)[1 + (\mu+1) R_K]},$$

on using current equations.

With proper circuit element and reasonable value of μ, $(R_L + r_p)$ can be made larger than $(\mu+1) R_K$, such as so to give the meter reading

$$e_M = \frac{\mu e_i R_L}{R_L + r_p}$$

and is unaffected by the supply voltage changes because the matched tubes are employed in the circuit.

(2) **Cathode drift compensation :** A compensating arrangement for eliminating the effect of heater voltage changes is shown

in the figure 17. Due to the change in heater voltage, random

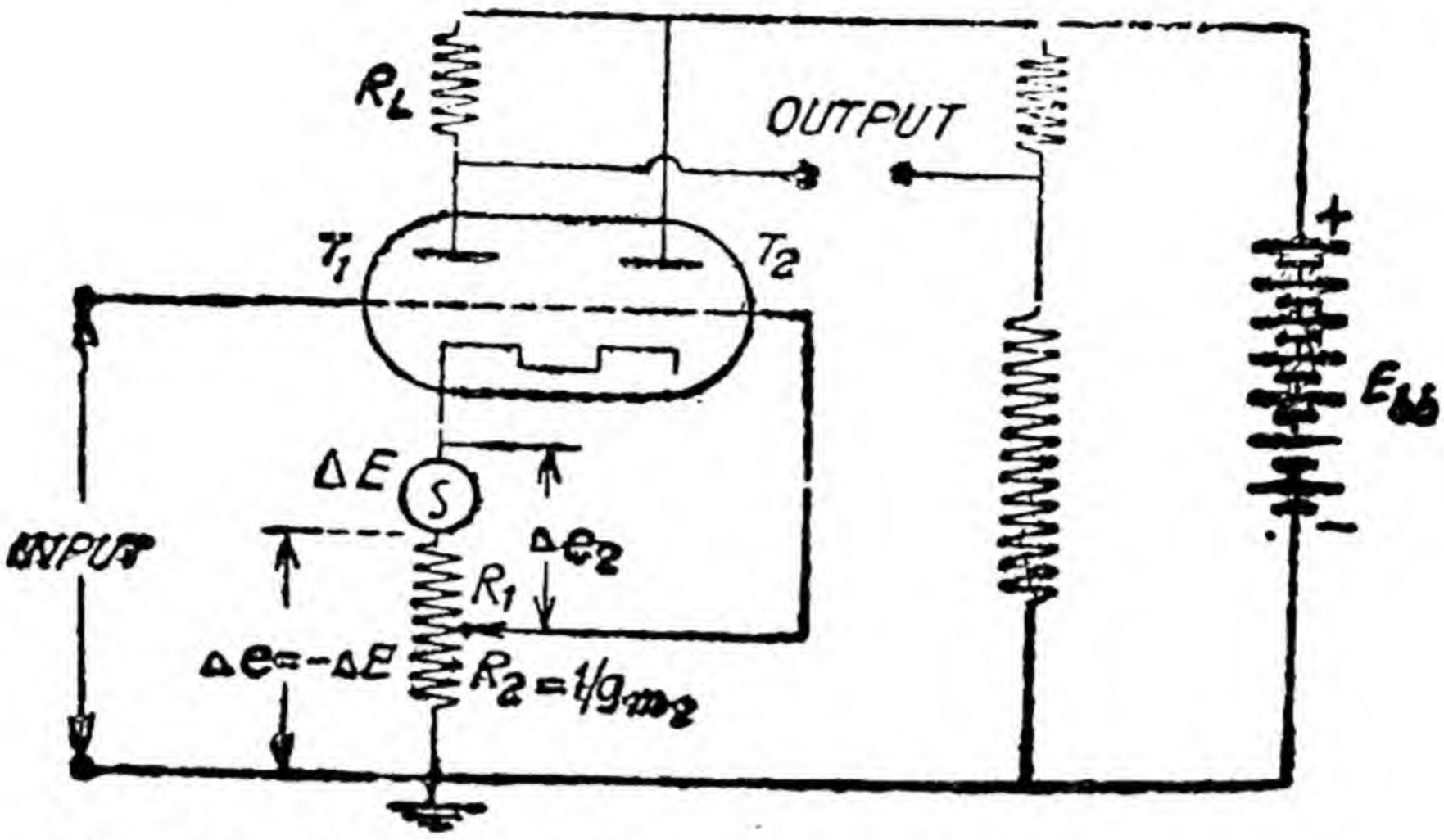


Fig. 17. Cathode drift compensated d.c. amplifier.

changes in cathode emission occur which inturn changes the output voltage.

Suppose the heater voltage changes by an amount ΔE, then to eliminate its effect on the cathode emission, the plate current of T_2 must decrease by an amount ΔI_2 such that drop across (R_1+R_2), *i.e.*, $\Delta I_2\,(R_1+R_2)$ be equal and opposite to ΔE, *i.e.*

$$\Delta E = -\Delta I_2\,(R_1+R_2) \qquad \ldots(1)$$

Now if change in grid voltage of T_2 is Δe_2, then

$$-\frac{\Delta I_2}{\Delta e_2} = g_{m_2},$$

giving

$$\Delta I_2 = = \Delta e_2\, g_{m_2},$$

so that rom equation (1),

$$\Delta E = \Delta e_2 g_{m_2}\,(R_1+R_2). \qquad \ldots(2)$$

Further

$$\Delta e_2 = \Delta E - \text{voltage across } R_1$$

$$= \Delta E - \frac{\Delta E}{R_1+R_2} \times R_1$$

$$= \Delta E\left[1 - \frac{R_1}{R_1+R_2}\right]$$

$$= \frac{R_2 \Delta E}{R_1+R_2} \qquad \ldots(3)$$

From equations (2) and (3),

$$\Delta E = \Delta E.\ \frac{R_2}{R_1+R_2}.\ g_{m_2}\,(R_1+R_2)$$

$$= \Delta E.R_2\, g_{m_2},$$

giving

$$R_2 = \frac{1}{g_{m_2}},$$

so that ΔE be equal and opposite to the voltage drop across (R_1+R_2) due to the flow of current ΔI_2. Thus change ΔE is cancelled and heater voltage remains effectively the same to keep the cathode emission unaltered.

EXERCISES

1. What do you mean by wideband amplifier ? Describe some uses of the wideband amplifier.
2. Draw the circuit of uncompensated wideband amplifier. Describe its performance in low, medium and high frequency ranges.
3. Describe the series peaked high frequency compensation of a wide-band amplifier.
4. Draw the circuit of a wideband amplifier with both low and high frequency compensation Explain the compensation at both the frequencies.
5. Describe the Loftin-White d.c. amplifiers. What are the drawbacks of D.C. amplifiers.
6. How the drawbacks of d.c. amplifiers are overcome in (i) resistance coupled d.c. amplifiers and (ii) cathode coupled d.c. amplifier.
7. What do you mean by draft in d.c. amplifiers ? Describe balanced push pull amplifier.

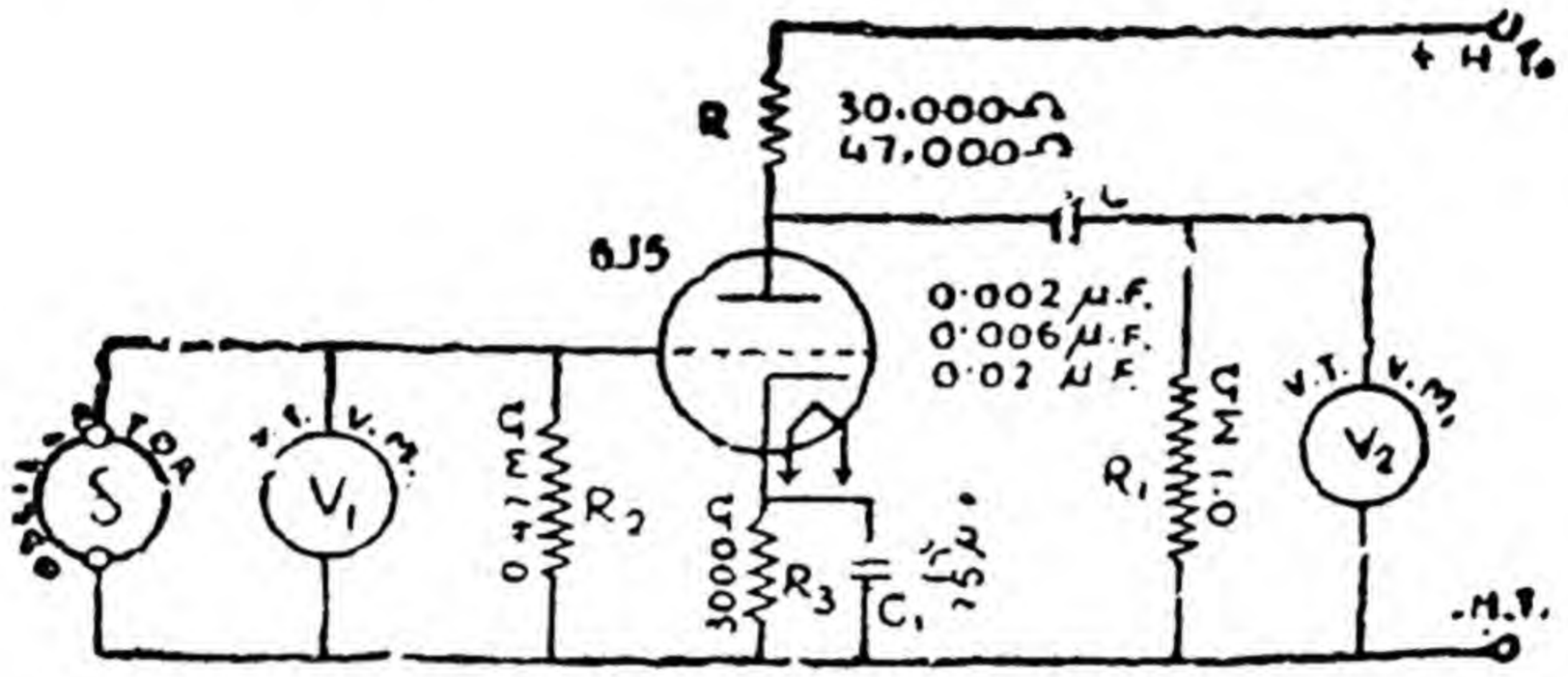
EXPERIMENT NO. 1

To study the frequency response of a single stage audio frequency amplifier with different types of coupling :

(*a*) Resistance-capacitance,
(*b*) Inductance capacitance,
(*c*) Transformer coupling.

(a) Resistance-Capacitance coupling :

Procedure : A triode valve 6*J*5 is connected as shown in the figure. The heater is connected to a six volts battery.



Measurement of gain of resistance-capacitance coupled amplifier.

The plate is connected to positive of H.T. through a load resistance R. The cathode is connected to negative of H.T. through a cathode resistor R_3 of 3000 ohms, having a condenser C_1 of $25\mu F$ across it. These two give automatic negative bias to the grid. An a.f. oscillator is connected across R_2 of $0{\cdot}47\ M\Omega$. The input and output voltages are measured by means of *V.T.V.M.*, V_1 and V_2 respectively. R_1 is the grid leak resistance of $0{\cdot}1\ M\Omega$ and C is the coupling capacitor.

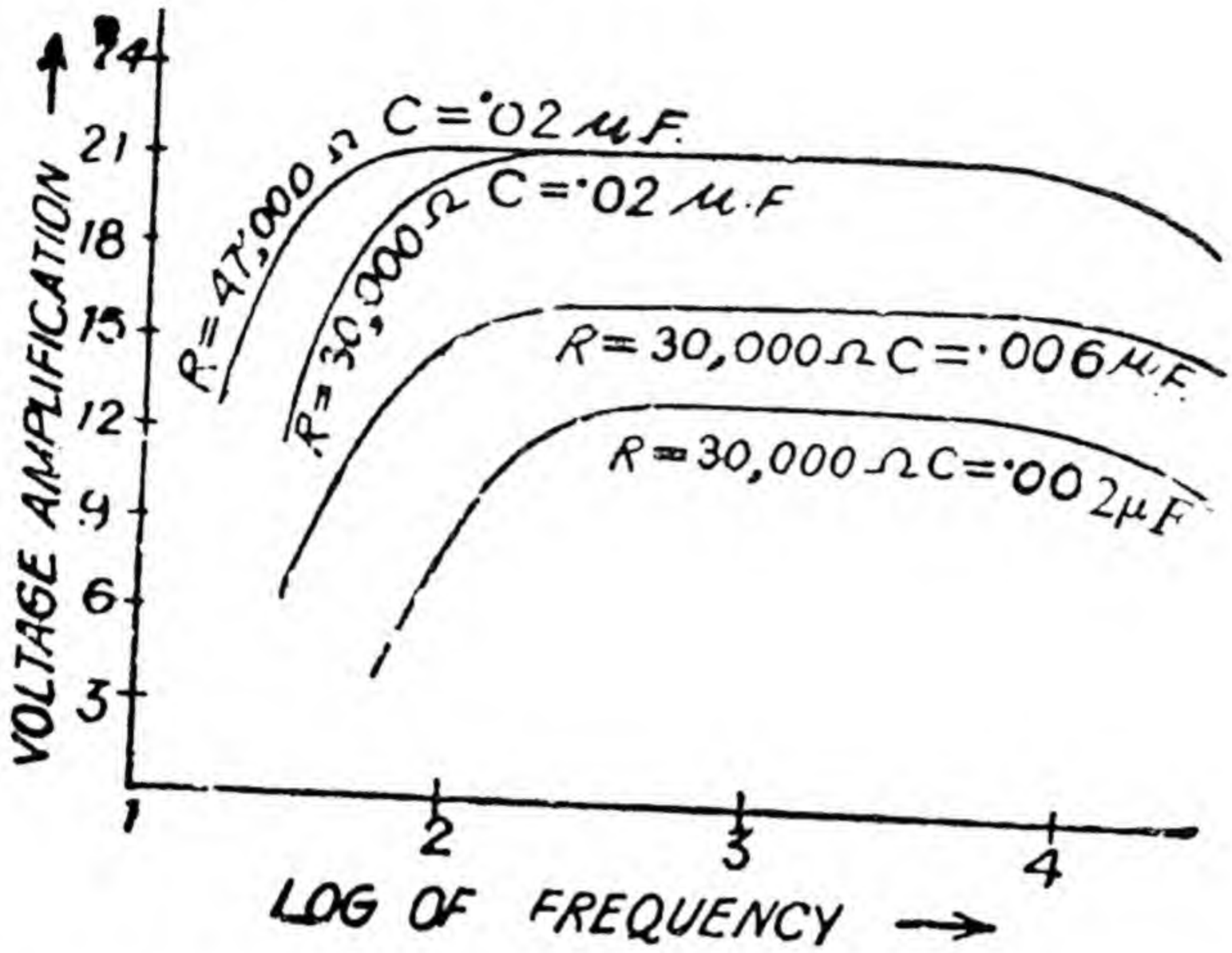
(i) Set R to a value $30K\Omega$ and C to somewhat low value $0{\cdot}002\mu F$. Keeping the input voltage constant, say at 1 volt, now vary the oscillator frequency over the audio range and note the output voltage for corresponding frequencies.

(*ii*) Keeping R fixed, repeat the above observations for two higher value of capacitances C, *viz*. $0{\cdot}006\mu F$ and $0{\cdot}02\mu F$.

(*iii*) Using the highest value of C employed in the preceding case, *i.e* , $0{\cdot}02\mu F$., repeat the above observations with a higher value of load resistance $R=47K\Omega$.

(*iv*) For each case plot a curve of voltage amplification (output/input) to a logarithmic base of frequency.

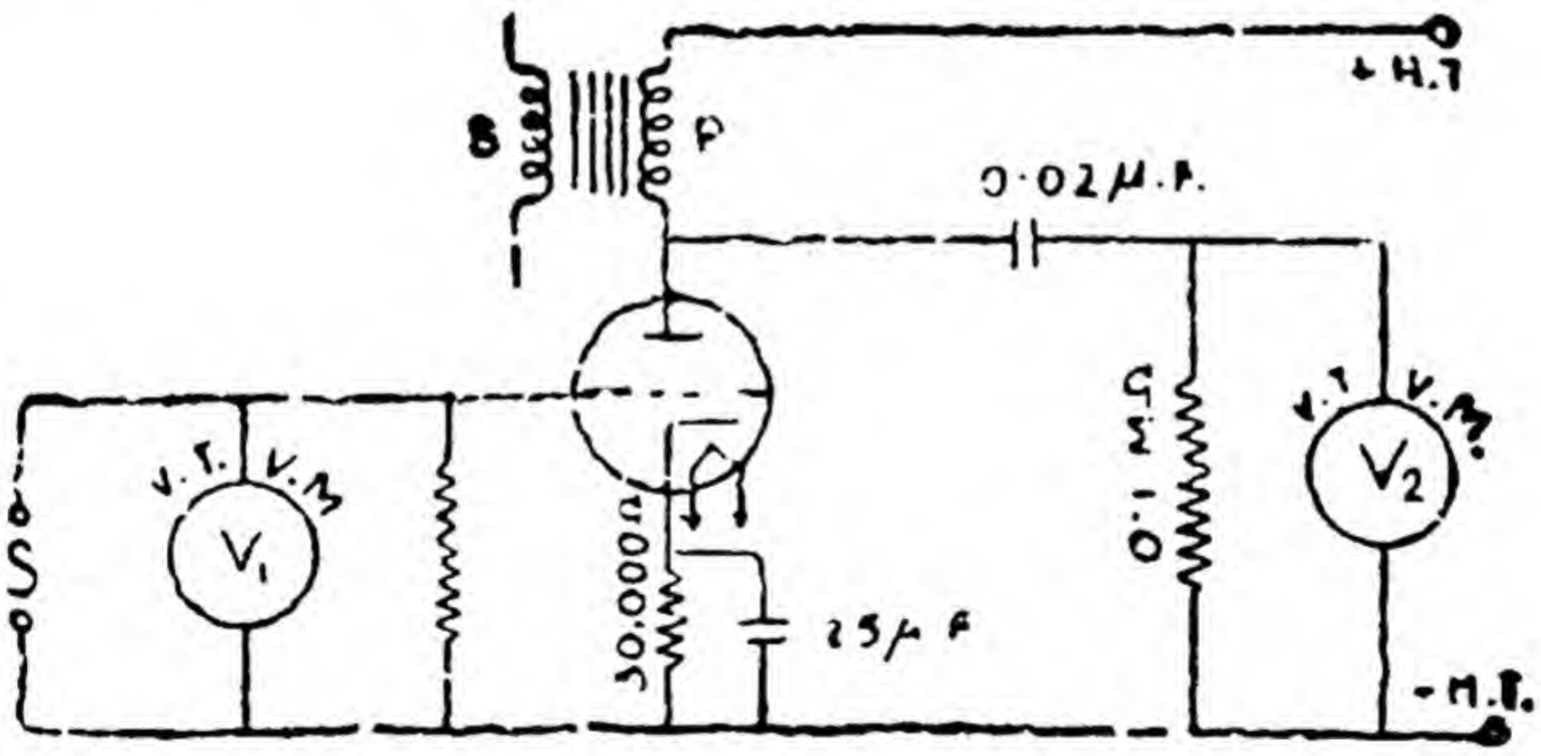
Results : Following are the curves obtained in case of resistance-capacitance coupled amplifier.



Voltage gain of a resistance-capacitance coupled amplifier.

Conclusion obtained from these graphs have been already discussed in the theory of resistance-capacitance coupled amplifier.

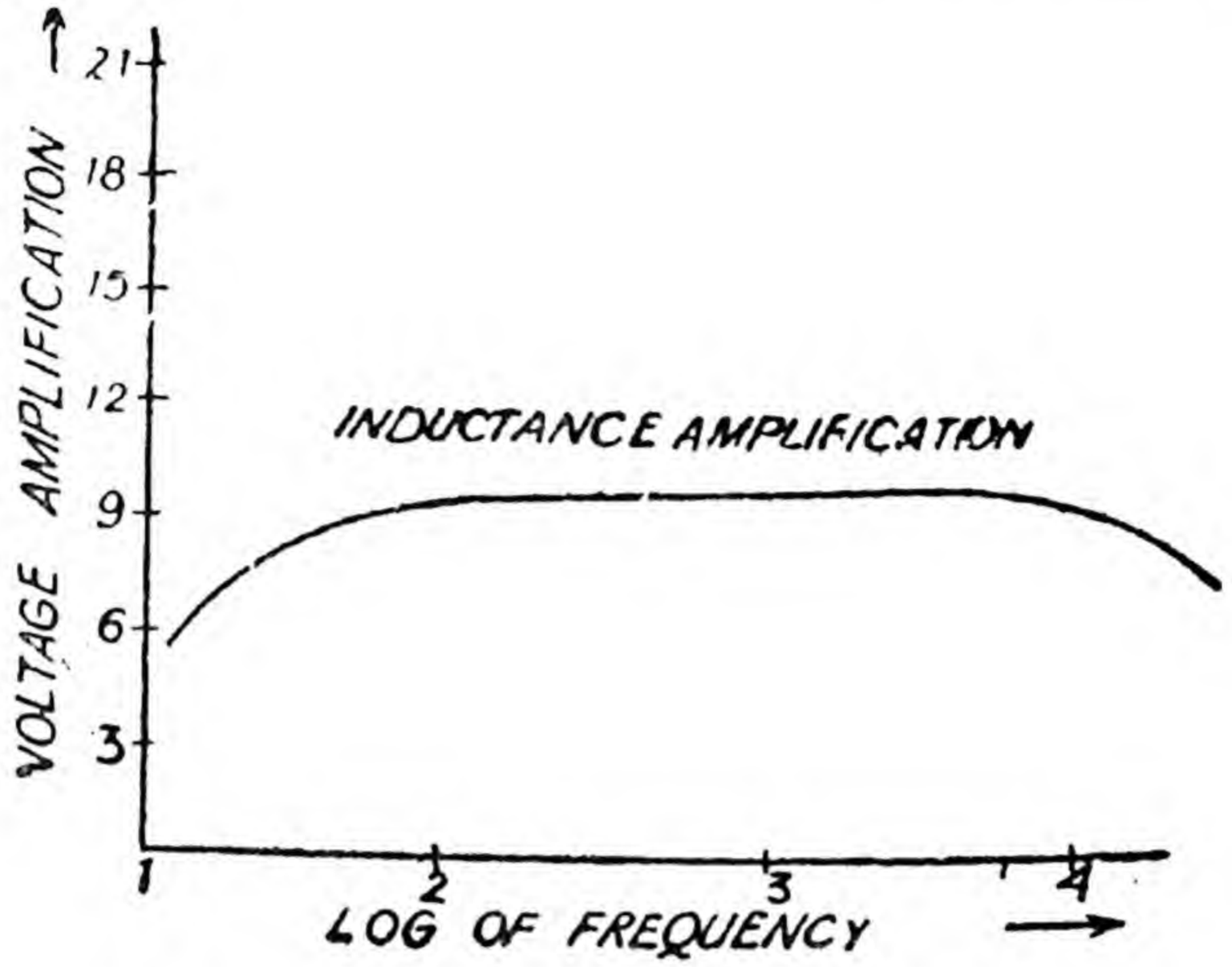
(b) Capacitance-Inductance coupled amplifier :



Measurement of gain of an inductance coupled amplifier.

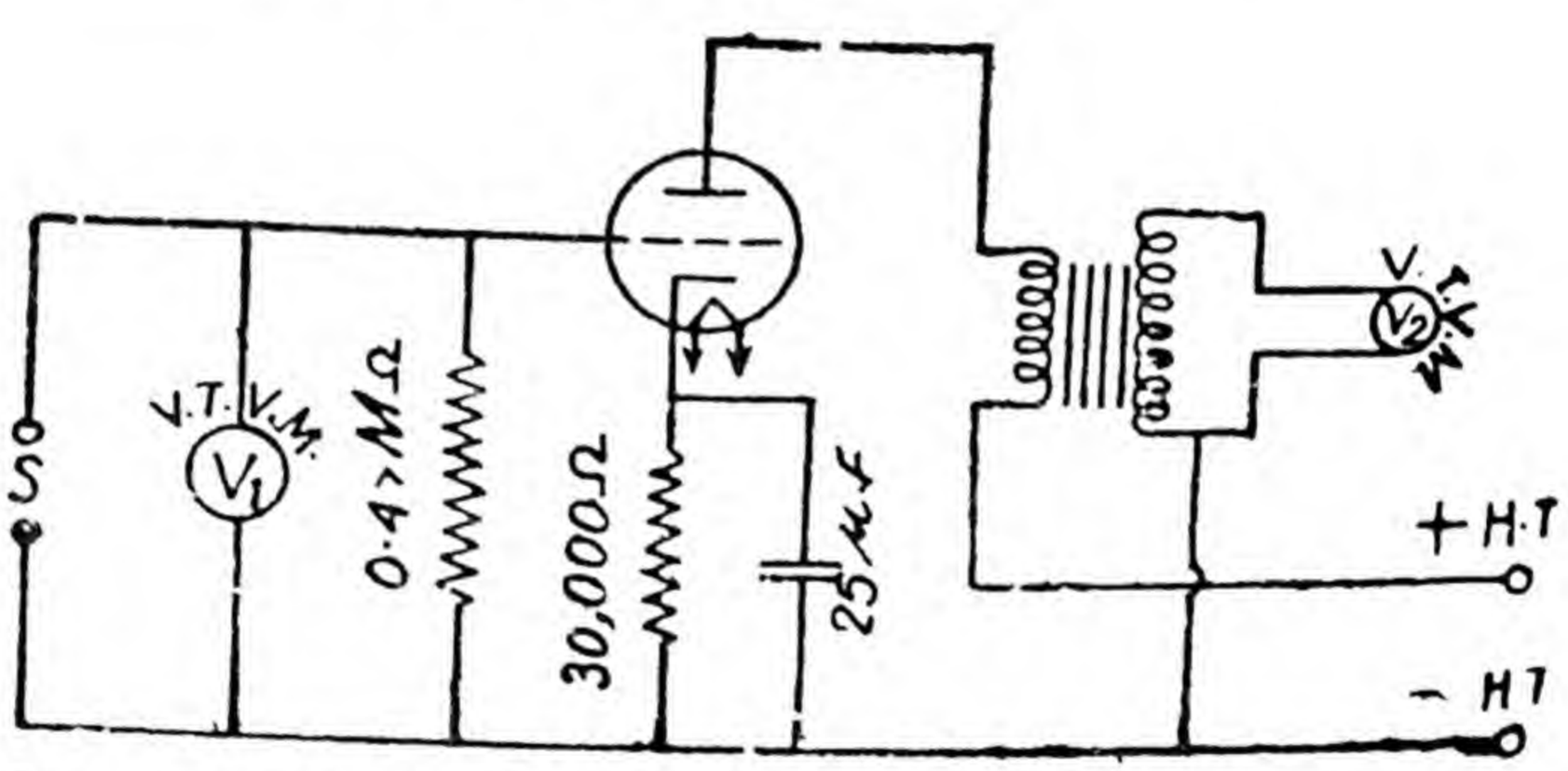
Procedure : The *LC* coupled amplifier differs from resistance-capacitance coupled amplifier in that *R* has been replaced by the primary of an a.f. transformer. Vary the oscillator frequency over the audio range and measure the output voltage. Plot a curve of voltage amplification to a logarithmic base of frequency.

For conclusion refer to the theory of *LC* coupled amplifier.

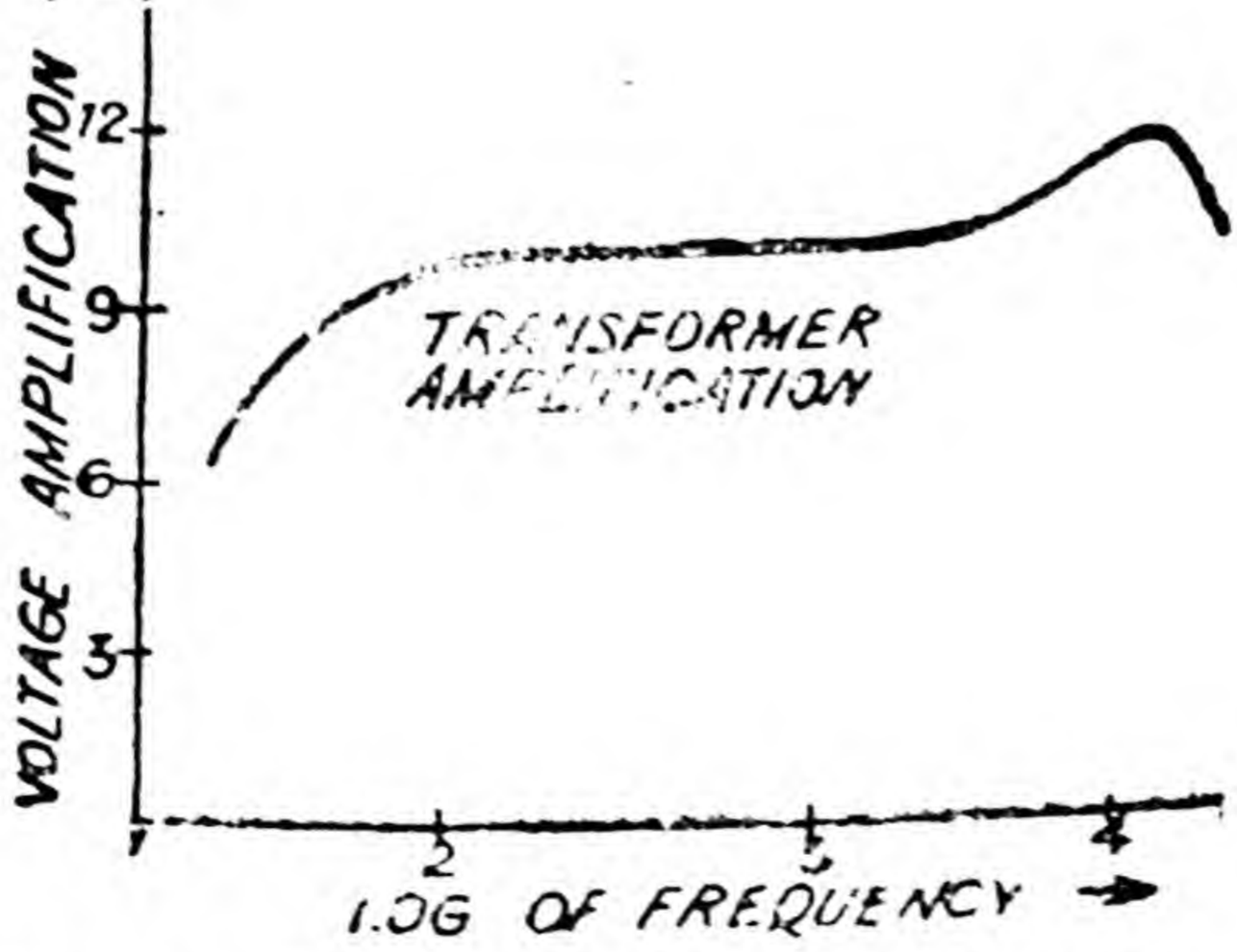


Voltage gain of *LC* coupled amplifier.

(c) Transformer coupled amplifier :



Measurement of gain of transformer-coupled amplifier.



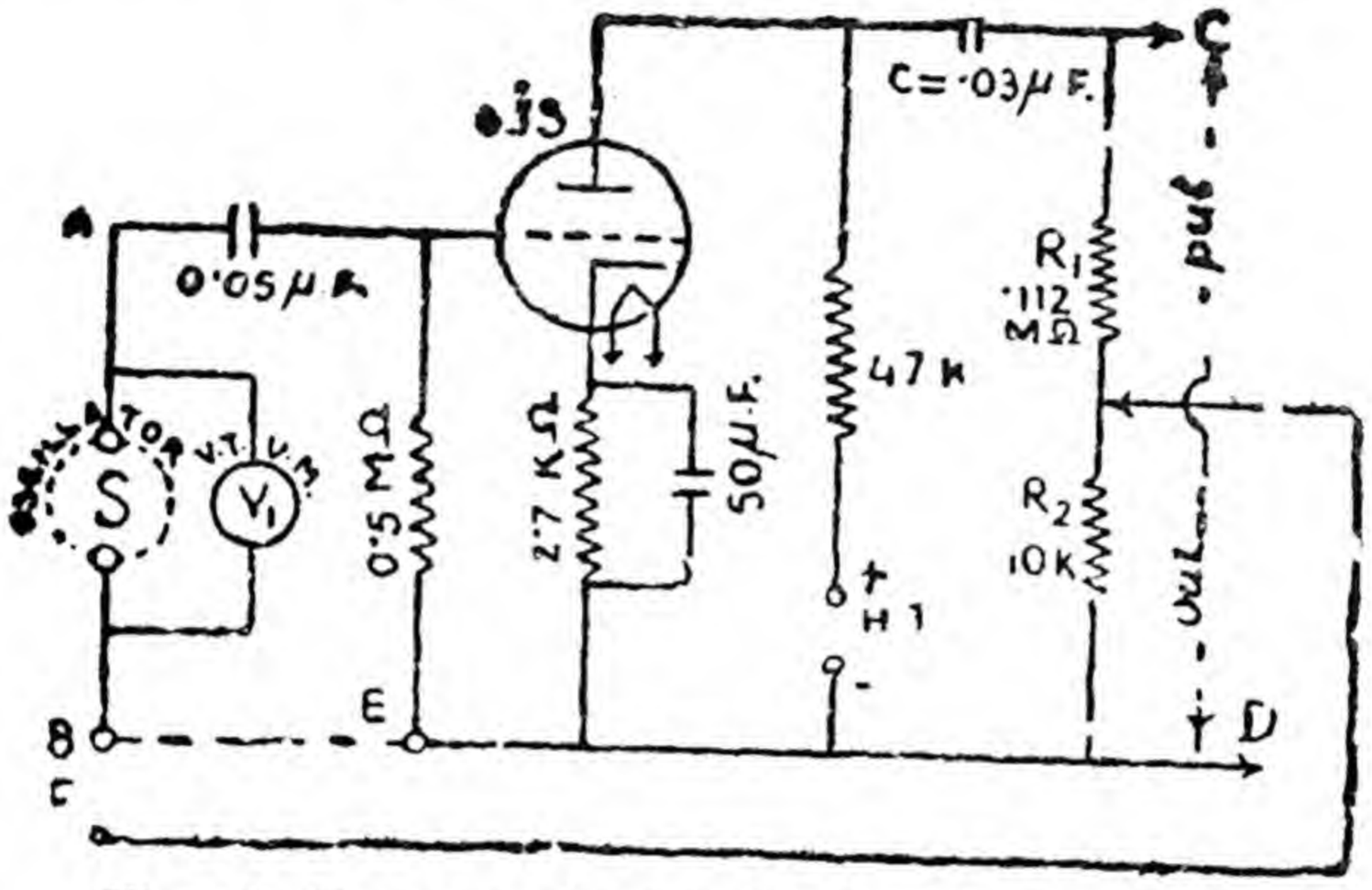
Voltage gain of transformer-coupled amplifier.

Procedure : Disconnect the *R.C.* coupling and repeat the above test with V. T. V. M. connected across the secondary of the transformer. Plot a curve of voltage amplification to a logarithmic base of frequency.

State the conclusions drawn from the curve.

EXPERIMENT No. 2

To study the effect of negative feed back on the performance of an audio frequency amplifier.

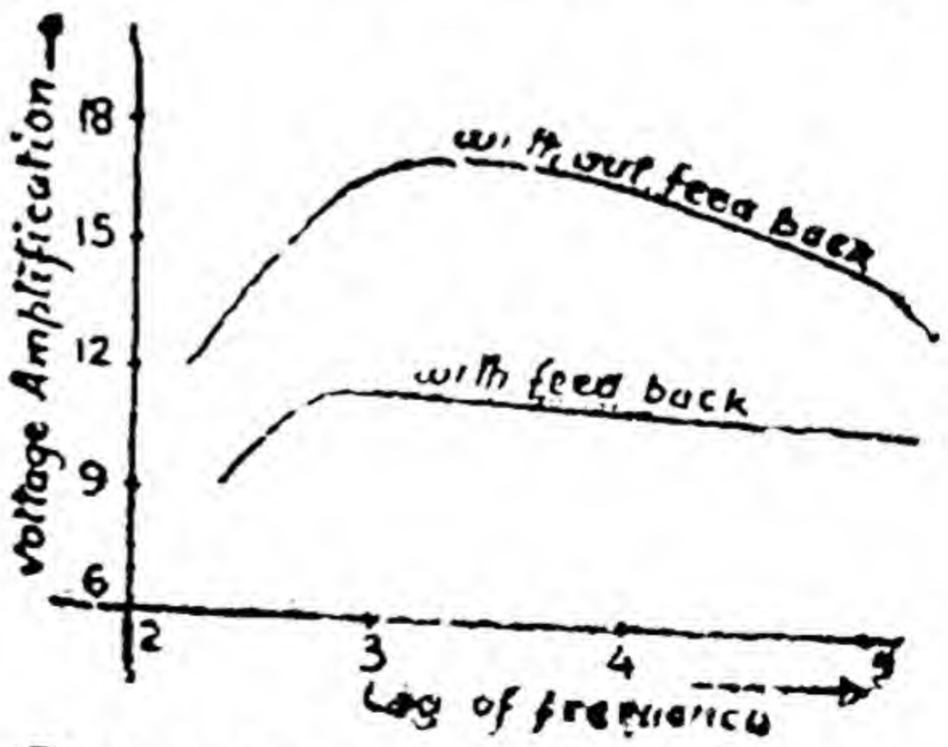


Circuit diagram of negative feed back amplifier.

Procedure : A resistance-capacitance coupled amplifier is shown in figure. *C* is the coupling condenser. Across the output are two series connected resistances R_1 and R_2. They should be high enough to impose little damping on the output and of such values that βA is large compared with unity, where A represents the amplification of the single stage and β is the ratio

$$\frac{R_2}{R_1 + R_2}.$$

In order to perform the experiment, connect a.f. oscillator across *AB*. First perform the experiment without feed-back. For this connect the terminal *EB*. Apply one volt as input and measure the output voltage developed across *CD* by means of a V. T. V.M. Now change the frequency of the oscillator and check the input voltage as it should be one volt and again measure the output. Now connect *C* to *B* and perform the experiment with feed back. Take various readings and draw a graph between log of frequency and voltage gain in both cases.

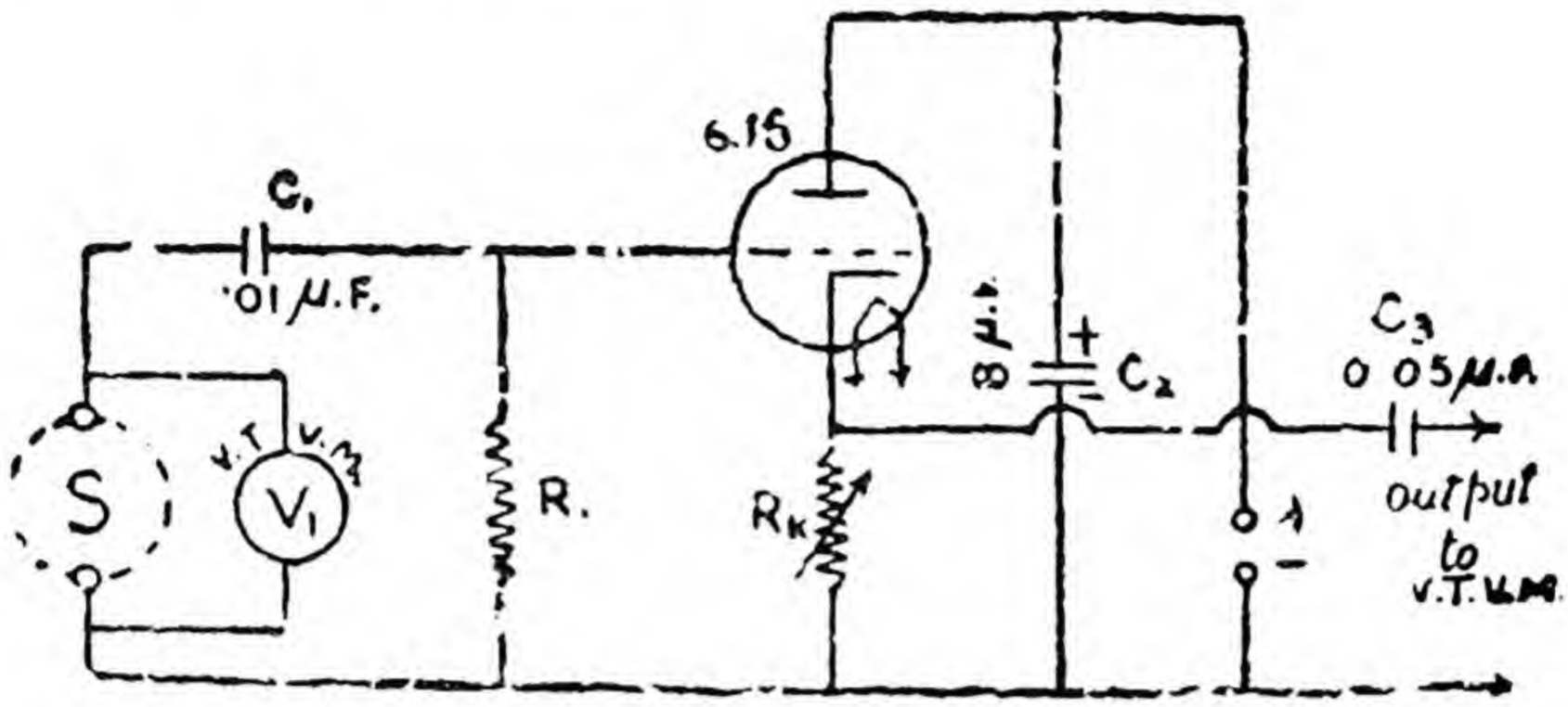


Response curves of a negative feed-back amplifier.

Result : The graph shows the effect of negative feed-back on a.f. amplifier. It is clear from the graph that the voltage gain with feed back is independent o frequency.

EXPERIMENT No. 3.

To study the cathode follower amplifier.



Cathode Follower.

Procedure : (*i*) A triode is connected in a cathode follower circuit as shown in the figure. Knowing the values of μ and r_p for the valve, calculate the theoretical values of voltage amplification for different values of R_k by the formula

$$\text{voltage amplification} = \frac{\mu R_k}{r_p + R_k (1+\mu)}$$

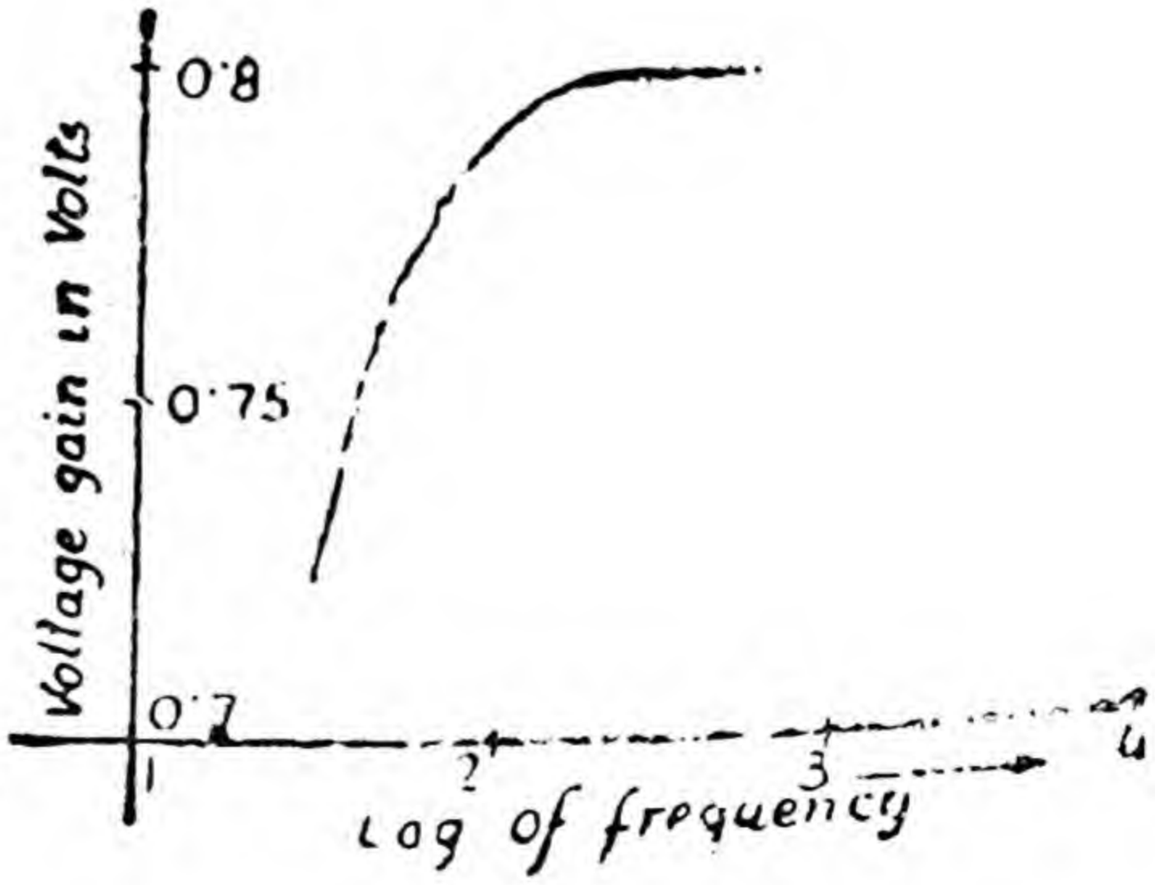
Draw a graph between the theoretical values of voltage amplification and different resistances R_k.

(*ii*) To draw a frequency response curve of cathode follower, various frequencies from a.f. oscillator are applied with constant input and R_k is also kept fixed. Check up input at each frequency. Measure output and draw a graph between log of frequency and voltage gain.

Result : Following are the graphs obtained in case of cathode follower :



(*i*) Graph between resistance and theoretical voltage amplification.



(*ii*) Graph between log of frequency and voltage gain.

Conclusion :

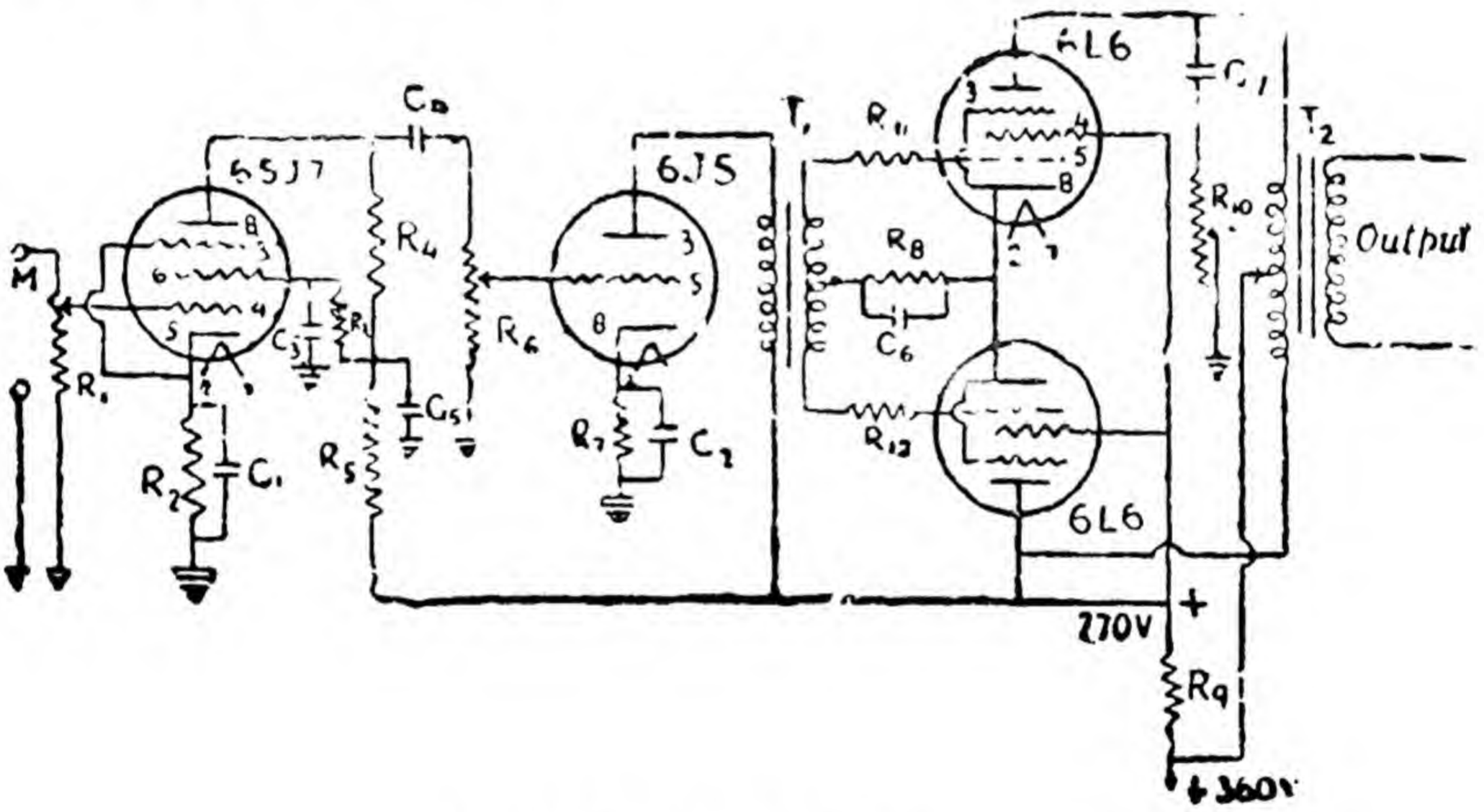
(*i*) The graph between resistance and theoretical voltage amplification shows that as the value of R_k increases, the voltage amplification also increases. For high values of R_k, we get approximately the constant output voltage.

(*ii*) The frequency response curve shows that voltage gain increases upto a certain frequency and then it becomes constant.

EXPERIMENT No. 4.

To design and study the gain and phase shift of a speech amplifier.

Procedure : A simple circuit for the design of speech amplifier is shown in figure. The choice of circuit elements is based on the function which it has to perform. The value and function of each component is given below.



Speech amplifier circuit.

Function and value of each component

S.No.	Elements	Function	Value
1.	Resistor R_1	The input of microphone is applied to the grid of rentode valve 6 *SJ7* through resistor R_1.	4·7 megohms, $\frac{1}{2}$ watt dissipation power.
2.	Combination of R_1 and capacitor C_1.	For cathode biassing.	R_2 is 1·4 kilohms, $\frac{1}{2}$ watt dissipation power, C_1 is of $20\mu F$ 50 volts.
3.	C_3, electrolytic condenser.	Screen current bypass condenser.	$0{\cdot}1\mu F$, 200 volts.

4.	Resistor R_3	Gives proper potential to the screen grid.	1·5 megohms ½ watt.
5.	Capacitor C_4 Resistor R_4	R_4 functioning as the load resistor of first valve and C_4 is a coupling condenser. The latter prevents direct voltage application to the grid of 6J5.	0·01 μF, 400 volts; 0·22 megohms, ½ watt.
6.	Resistor R_6	It is grid leak resistor for triode. Its function is explained in theory This also controls the gain of the amplifier.	1 megohm, ½ watt.
7.	Combination of R_7 and C_2	For cathode biasing.	Same values as those of R_2 and C_1.
8.	Combination of R_8 and C_6	Provides cathode bias.	250 ohms. 10 watt, 50 μF, 50 volts electrolytic.
9.	Resistors R_{12} and R_{11}	These resistors damp the parasitic oscillations which often start in centre-tapped coils	Each of 10 kilohms. ½ watt dissipation power.
10.	Resistor R_9	It is bleeder resistor which stabilises the voltage and improves the working of the tube.	2000 ohms, 10 watts.
11.	Combination of C_7 and R_{10}	Functioning as tone control.	·01 μF, 450 volts 2·5 kilohms, 1 watt.
12.	Combination of R_5 and C_5	This is decoupling arrangement which prevents the positive feed back.	8 μF, 450 volts; 4·7 kilohms, ½ watt.
13.	Transformer T_1 and T_2	T_2 is output transformer of 6L6 and matches the impedance of loudspeaker. T_1 is used for coupling between stages.	

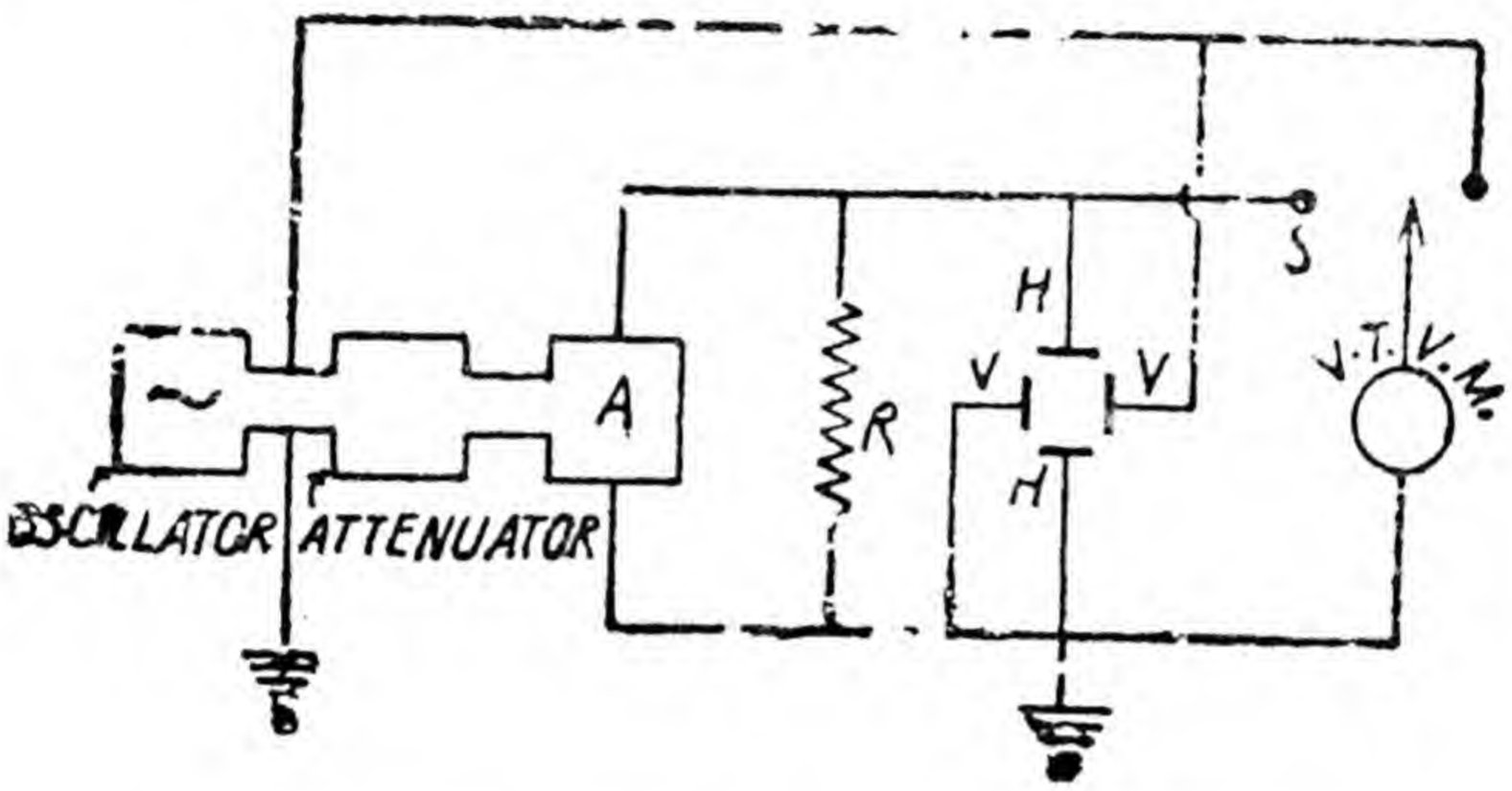
To measure the gain and phase shift of the amplifier, it is connected in the circuit as shown in figure on page 623

The input terminals of the amplifier are connected to a variable frequency oscillator through a calibrated variable attenuator. A V.T.V.M. is connected to a switch by which the output of amplifier

and input of attenuator can be measured. The horizontal and vertical plates of C.R.O. are connected to amplifier output and oscillator output respectively. We proceed as follows :

(*i*) First set the oscillator frequency and output to a low value. Volume control is kept at 'maximum' while the positions of other controls are noted.

(*ii*) Secondly, adjust the attenuator so that V.T.V.M. records the same readings with both positions of switch *S* for which the amplifier gain will be equal to attenuator setting. This is noted.

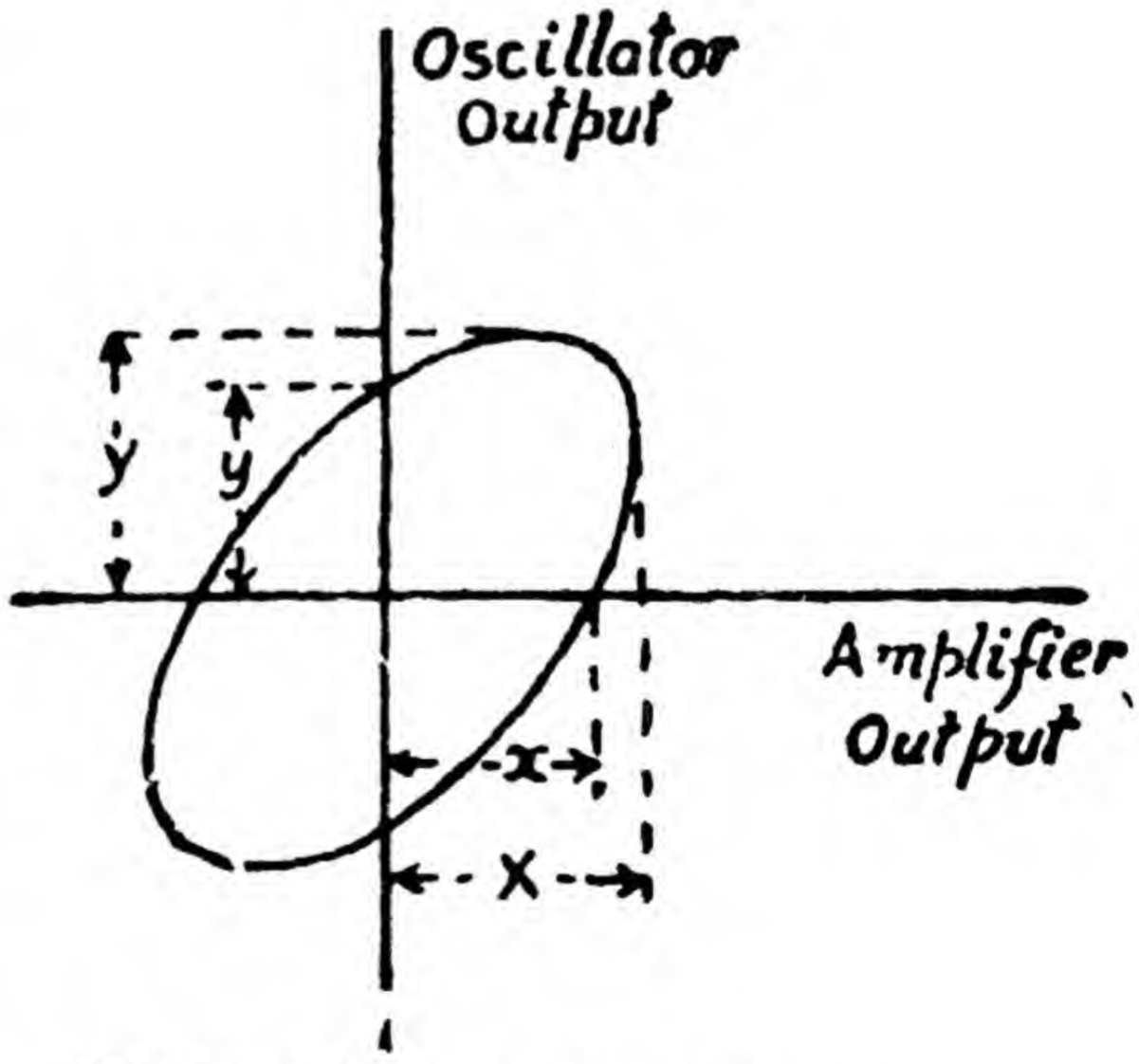


Circuit for measuring gain and phase shift of an amplifier.

(*iii*) Now calculate the phase shift δ from the oscillograph by using the relation

$$\sin \delta = \pm \frac{x}{X} = \pm \frac{y}{Y}.$$

If there is any necessity for the change of direction, the phase of amplifier voltage can be altered by connecting a capacitor in series at amplifier output.



Phase shift measurement.

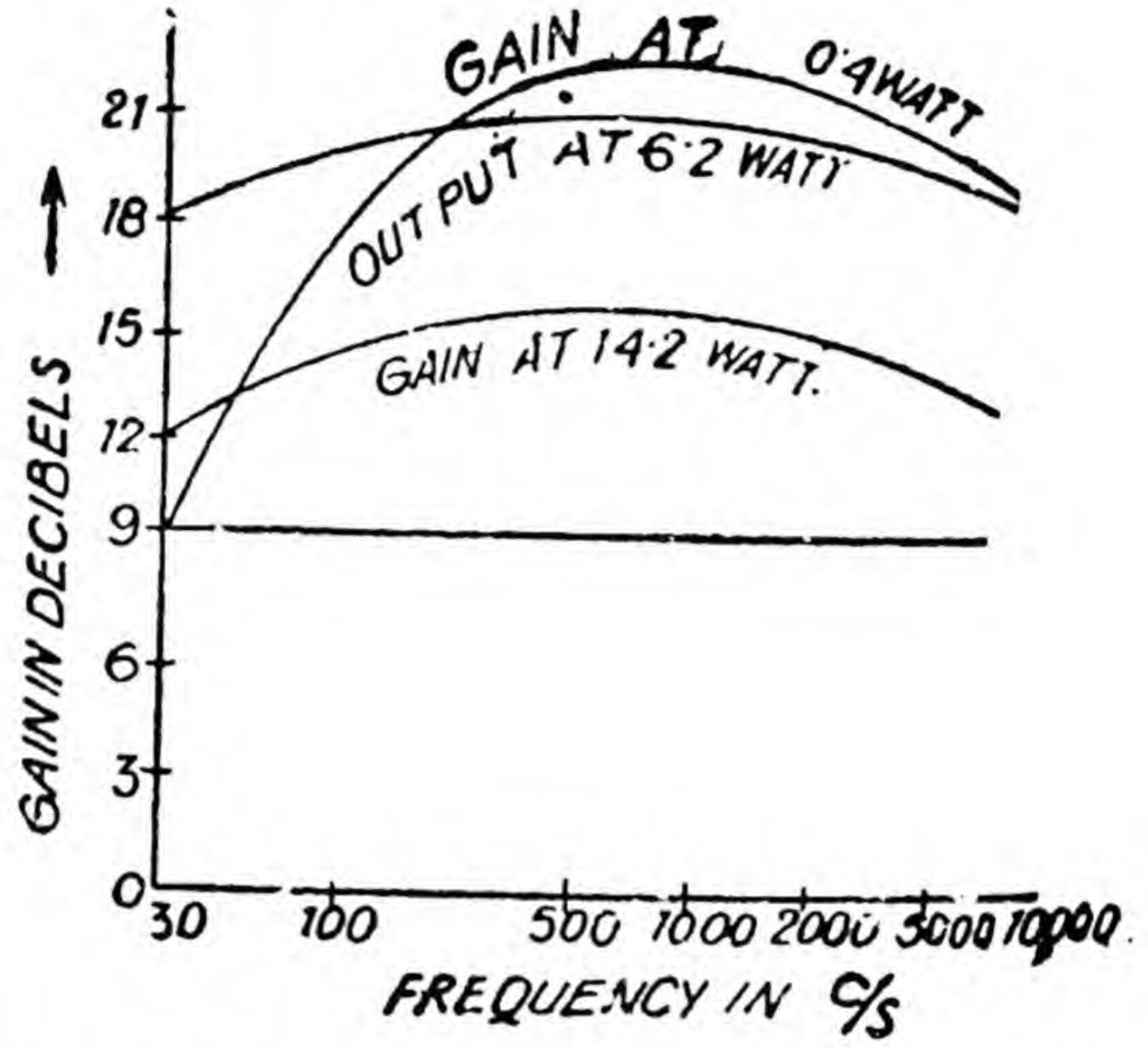
(*iv*) Keeping the output constant, the oscillator frequency is increased in step to a value for which the amplifier is designed

At each step the readings of phase shift and gain are noted.

(*v*) The whole procedure stated in above four points is repeated for higher values of output voltage which is varied upto the maximum rated value of the amplifier. The positions of amplifier and attenuator can be reversed if amplifier output is greater than the available oscillater output.

(*vi*) The whole test may be repeated for various settings of volume control and tone control.

The curves of gain and phase shift to a base of frequency for two or three output values at various setting of available control are plotted as shown in figure.



Amplifier gain and phase shift characterististics.

15 OSCILLATORS

For transmission of speech or music, we want continuous oscillations. The vacuum tube with proper combination of circuit elements may be used to give alternating current of any desired frequency. The vacuum tube does not create any energy and also it is not capable of producing oscillations by itself, but the main function of tne tube is to convert d.c. into alternating current. The vacuum tube alongwith associated circuit is called a vacuum tube oscillator.

Oscillator function are similar to a vacuum tube amplifier. In either case, d.c. power is supplied from the plate supply source and this results in alternating power across the output terminals. In amplifiers, however, the frequency, waveform and magnitude of the generated alternating power is governed by the controlling a c. voltage from an external source applied to the control grid of the amplifier tube. In oscillators, on the other hand, frequency, waveform and magnitude of the a.c. voltage generated depends only on the tube and associated circuit, and *no external controlling voltage is required.*

The principle of oscillations is that the electrons acquire kinetic energy from the plate supply source and the tube with its associated circuit converts their kinetic energies into the alternating field energy at the output terminals. In this way the tube in oscillator converts the d.c. power into a.c. power. Here we shall discuss a simplified theory of oscillations by considering a circuit of pure inductance and capacitance as shown in figure 1 (a).

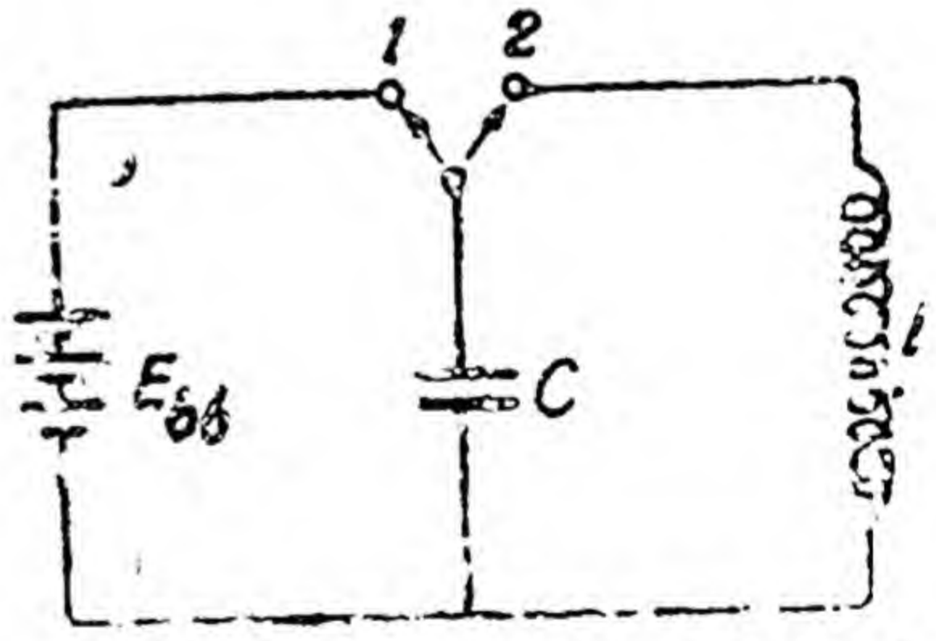


Fig. 1 (a) Showing principle of oscillations in *LC* circuit.

When the switch is connected to terminal 1, the capacitor is charged. The polarity of upper plate is positive and the electrostatic energy stored in it is $\frac{1}{2} CE_{bb}^2$. This is shown in fig. 1(b) at $t=0$. Now let the switch be thrown to terminal 2, *i.e.* the capacitor is connected to the coil L and the current flows through it. The voltage across the capacitor reduces to zero and the

current in the coil becomes I_m. The electromagnetic energy in the coil is $\frac{1}{2}LI_m^2$. Thus electrostatic field of capacitor is converted into electromagnetic field in the coil. This situation is shown at $t=t_1$. After $t=t_1$, the current flows in the *same direction* and the capacitor begins to charge with lower plate positive till $t=t_2$. Thus the whole energy in the magnetic field is changed again into electrostatic field energy. Now the current in L becomes zero. After $t=t_2$, the capacitor begins to discharge through L. The current flows in the upward direction till $t=t_3$. The capacitor is completely discharged. In this way electrostatic energy is converted into electromagnetic energy. After $t=t_3$, the current flows in the upward direction till the upper plate of the capacitor is charged as at $t=0$.

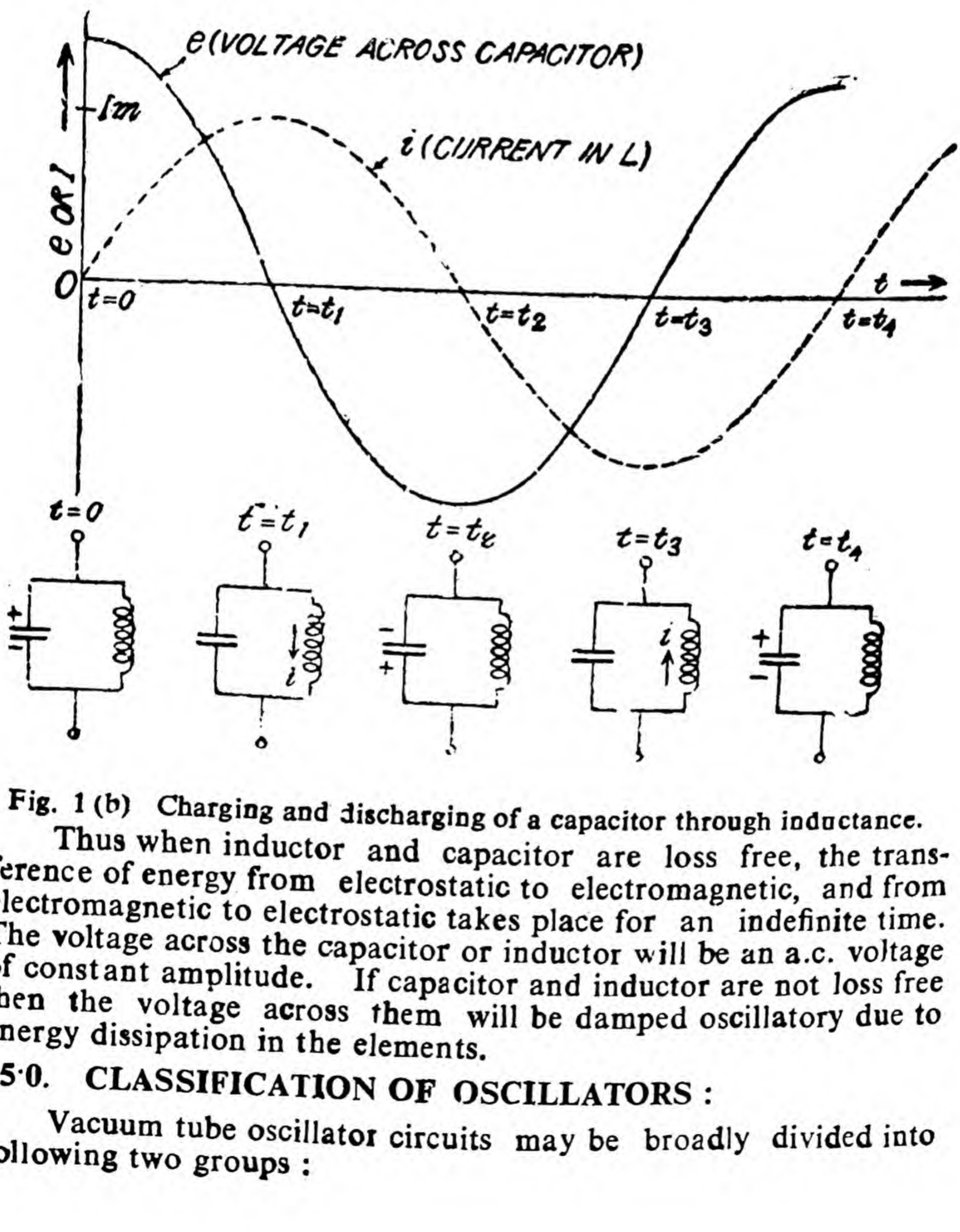


Fig. 1 (b) Charging and discharging of a capacitor through inductance.

Thus when inductor and capacitor are loss free, the transference of energy from electrostatic to electromagnetic, and from electromagnetic to electrostatic takes place for an indefinite time. The voltage across the capacitor or inductor will be an a.c. voltage of constant amplitude. If capacitor and inductor are not loss free then the voltage across them will be damped oscillatory due to energy dissipation in the elements.

15·0. CLASSIFICATION OF OSCILLATORS :

Vacuum tube oscillator circuits may be broadly divided into following two groups :

(A) Those circuits which produce non-sinusoidal waves :

Circuits which produce non-sinusoidal waves are known as *relaxation oscillators.* They are classified as :

(i) Vender pol (ii) Multi-vibrator (iii) Glow tube discharge (iv) Arc tube discharge (v) Saw tooth wave generator (vi) Square wave generator

(B) Those circuits which produce sinusoidal waves :

Sinusoidal waves can be produced by the following types of oscillators :

(1 *Negative resistance osciliators* : If the circuit behaves in such a way that an increase in voltage decreases the current and *vice-versa*, the circuit is said to have negative resistance. Any circuit having a negative resistance characteristic can be used as an oscillator. The four types of negative resistance oscillator circuits most commonly used are :

(*i*) Dynatron,
(*ii*) Transitron,
(*iii*) Push-pull circuit,
(*iv*) Resistance-capacitance circuit.

(2) *Feed-back oscillators* : If a portion of the output is combined with input signal, feed back is said to exist. If the net effect of feed-back is to increase the input signal, feed-back is called positive or direct feed-back. Positive feed-back is generally used in feed-back oscillators. They may classified as follows :

(*i*) Tuned plate,
(*ii*) Tuned grid,
(*iii*) Colpitt,
(*iv*) Complex types using more than one tuned circuits are :
  (*a*) tuned grid-tuned plate,
  (*b*) Meissner,
  (*c*) electron coupled,
(*v*) Resistance-capacitance tuned.

(3) *Crystal oscillators* : Crystal controlled oscillators are used when the frequency of oscillations is to be maintained at a fixed value. The crystal is not used to produce oscillations but controls the output frequency of the oscillator with which it is used.

(4) *Magnetostriction oscillators* : The magnetostriction oscillator circuit is based on the principle that a change in magnetization will cause a magnetic material to expand or contract and, conversely, that a contraction or expansion of a magnetic material will cause a change in magnetization.

15·1. BARKHAUSEN CRITERION FOR OSCILLATIONS :

Figure 2 (*a*) and 2 (*b*) show the block diagram and equivalent circuit to ascertain the conditions of sustained oscillations in a vacuum tube circuit.

Suppose a potential E_g is impressed on the grid of the tube without regard to the source of this potential. The output current

is given by $g_m E_g$ for an assumed linear operation of the circuit. Now because of the current I to the input of the coupling network,

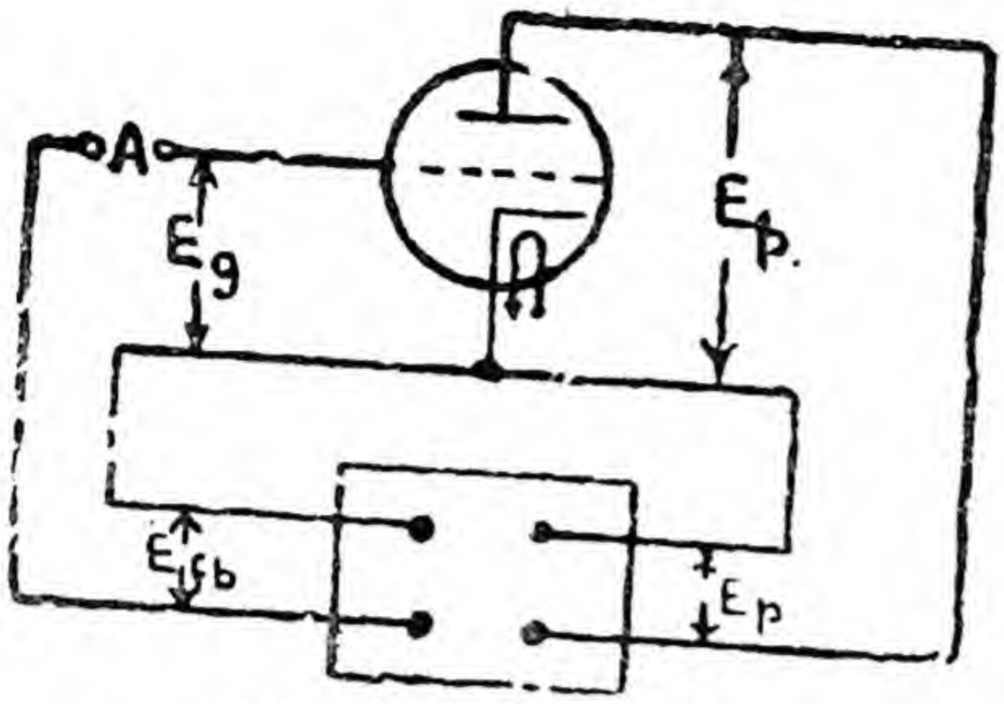


Fig. 2 (*a*) Block diagram.

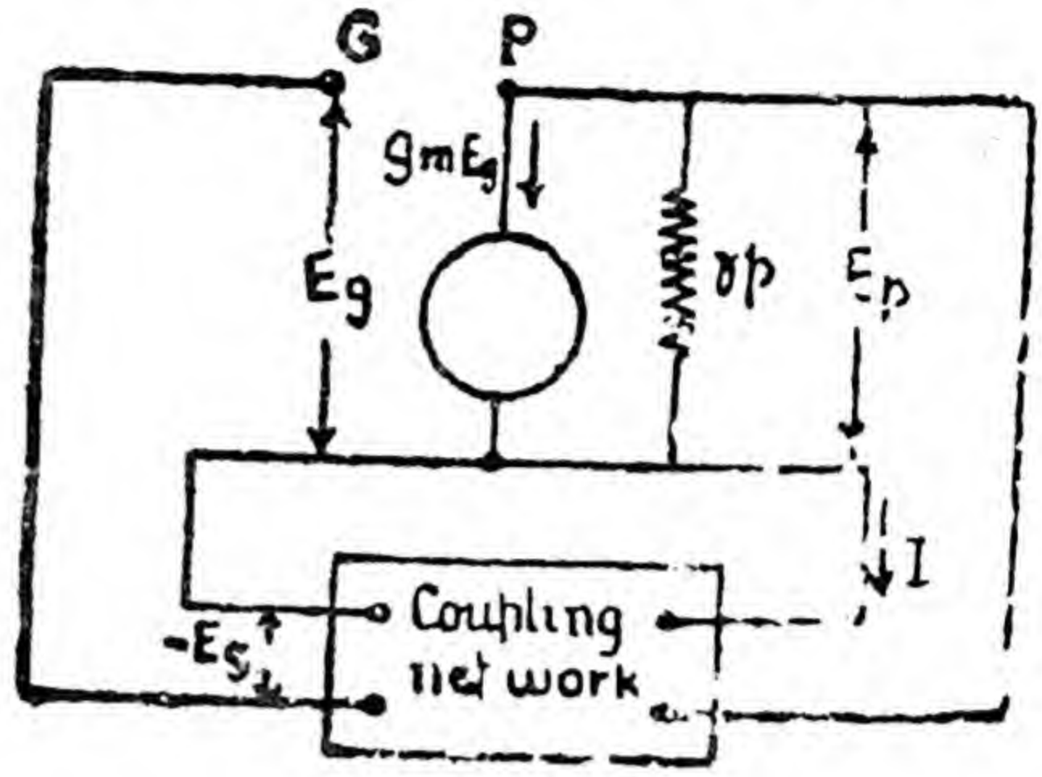


Fig. 2 (*b*) Equivalent circuit.

a certain potential will appear across the output terminals of this coupling network. If this return potential is equal in magnitude and phase to the original potential E_g, the circuit may be connected at point A and the system will continue to operate.

To examine the criterion for oscillations analytically, consider first the transfer impedance of the coupling network, which is denoted by Z_T, and which is defined as the ratio of the output potential to the input current. From the diagram

$$Z_T = -\frac{E_g}{I}.$$

Negative sign is due to phase shift of 180° in the coupling network. Thus

$$Z_T . I = -E_g,$$

$$Z_T \left(g_m E_g + \frac{E_p}{r_p} \right) = -E_g,$$

$$Z_T g_m E_g \left(1 + \frac{E_p}{r_p g_m E_g} \right) = -E_g.$$

$$Z_T g_m \left(1 + \frac{K}{\mu} \right) = -1,$$

where K is complex voltage amplification equal to $\frac{E_p}{E_g}$, and

$$r_p . g_m = \mu.$$

If $\mathbf{Z}$ denotes the impedance as viewed from plate side to coupling network.

$$Z = -\frac{E_p}{I} \quad \text{or} \quad ZI = -E_p,$$

$$Z \left(g_m E_g + \frac{E_p}{r_p} \right) = -E_p,$$

$$Z g_m E_g \left(1 + \frac{E_p}{r_p g_m E_g} \right) = -E_p,$$

$$Z g_m \left[1 + \frac{K}{\mu} \right] = -\frac{E_p}{E_g} = -K,$$

$$\frac{1}{\mu}+\frac{1}{K}=-\frac{1}{Zg_m},$$

$$\frac{1}{\mu}+\frac{1}{Zg_m}=-\frac{1}{K} \qquad \ldots(1)$$

From the coupling network figure, complex voltage amplification is

$$R=\frac{E_p}{E_g}$$

and feed-back factor $\alpha=\frac{E_{fb}}{E_p}$.

But coupling network has the property $E_{fb}=E_g$. Hence $K\alpha=1$. Putting this value in equation (1),

$$\frac{1}{\mu}+\frac{1}{Zg_m}=-\alpha,$$

$$\frac{1}{\mu}+\alpha=-\frac{1}{Zg_m} \qquad \text{or} \qquad \frac{1+\mu\alpha}{\mu}=-\frac{1}{Zg_m},$$

$$-g_m=\frac{\mu}{1+\alpha\mu} \qquad \text{or} \qquad -Zg_m=\frac{r_p \cdot g_m}{1+\alpha\mu},$$

$$Z+\frac{r_p}{1+\alpha\mu}=0.$$

This is another form of the Barkhausen criterion for oscillations.

15·2. ANALYSIS OF THE FUNDAMENTAL OSCILLATOR CIRCUIT :

Before examing particular types of oscillator circuit, it is useful to examine the action of each part of such a circuit. However, these actions are not independent of each other but are closely related, all occurring almost instantaneously. Fig. (3*a*) shows the fundamental oscillatory circuit.

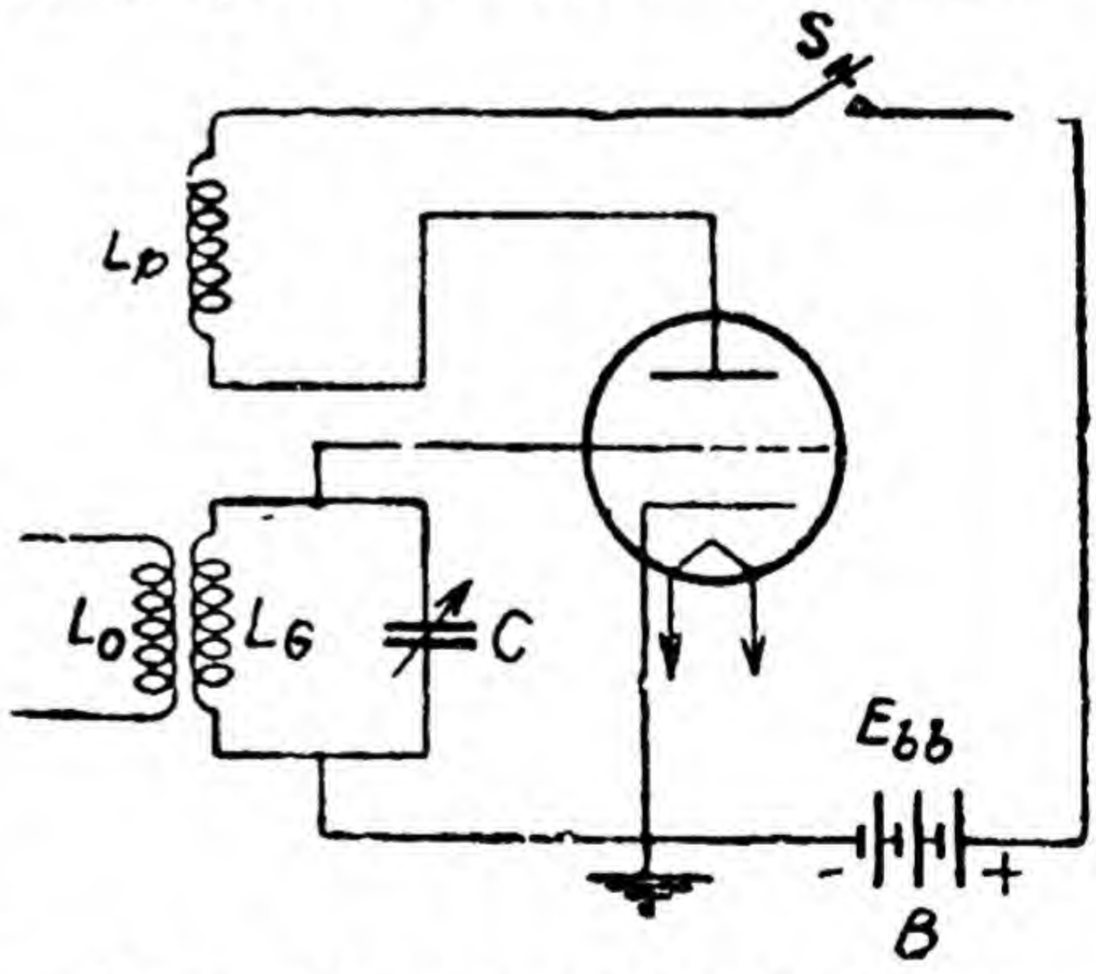


Fig. 3 (a) A fundamental oscillator circuit.

At the instant of closing the switch S, the electrons emitted from the cathode will be drawn to the plate. This causes a current to flow from the cathode to the plate, through the plate coil L_p, the E_{bb} power supply and back to cathode. As the current through L_P increases, a magnetic field builds up around the plate coil and by mutual inductance, a voltage increasing in magnitude is included in the grid coil L_G. Due to this fact the grid becomes positive thereby

increasing the amount of plate current, and the capacitor *C* in the tank circuit becomes charged. The increasing plate current, thus causes a greater voltage to be induced in the grid coil, thereby increasing the plate current still further until a point of saturation, depending upon the resistance of the circuit, plate supply voltage, and the characteristic of the tube, is reached.

As soon as the plate current ceases to increase, the field about the plate coil ceases to expand and no longer induces a voltage in the grid coil. The tank capacitor *C* starts to discharge through the grid coil; hence it makes the grid less positive, which in turn causes the plate current to decrease and the magnetic field about the plate coil to collapse. A voltage will again be induced in the

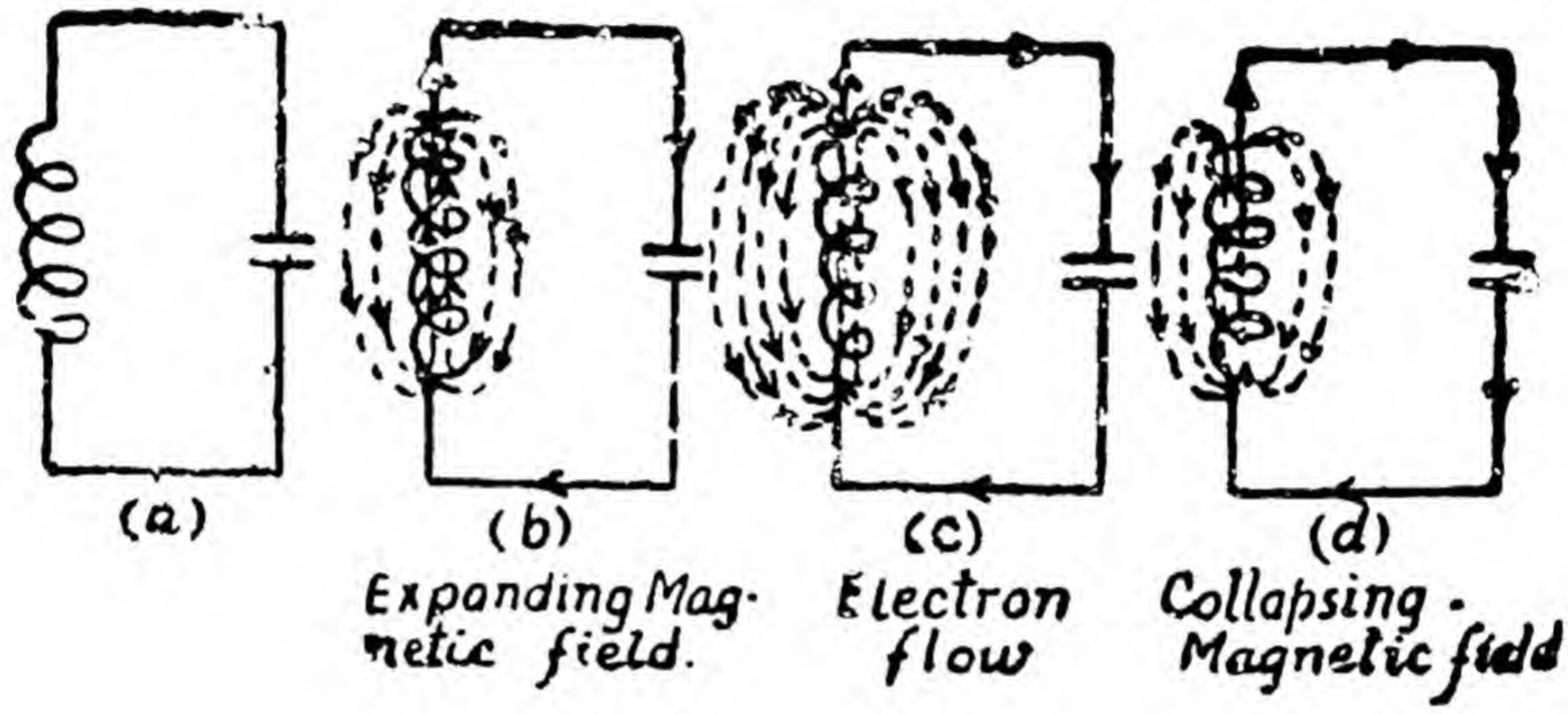


Fig. 3. (*b*)

grid coil but opposite in direction to that produced by an expanding field. The voltage on grid thus becomes negative, thereby decreasing the plate current still further until zero plate current flows.

As the induced voltage in L_G is zero when cut-off is reached, the grid capacitor which has been fully charged in the meanwhile,

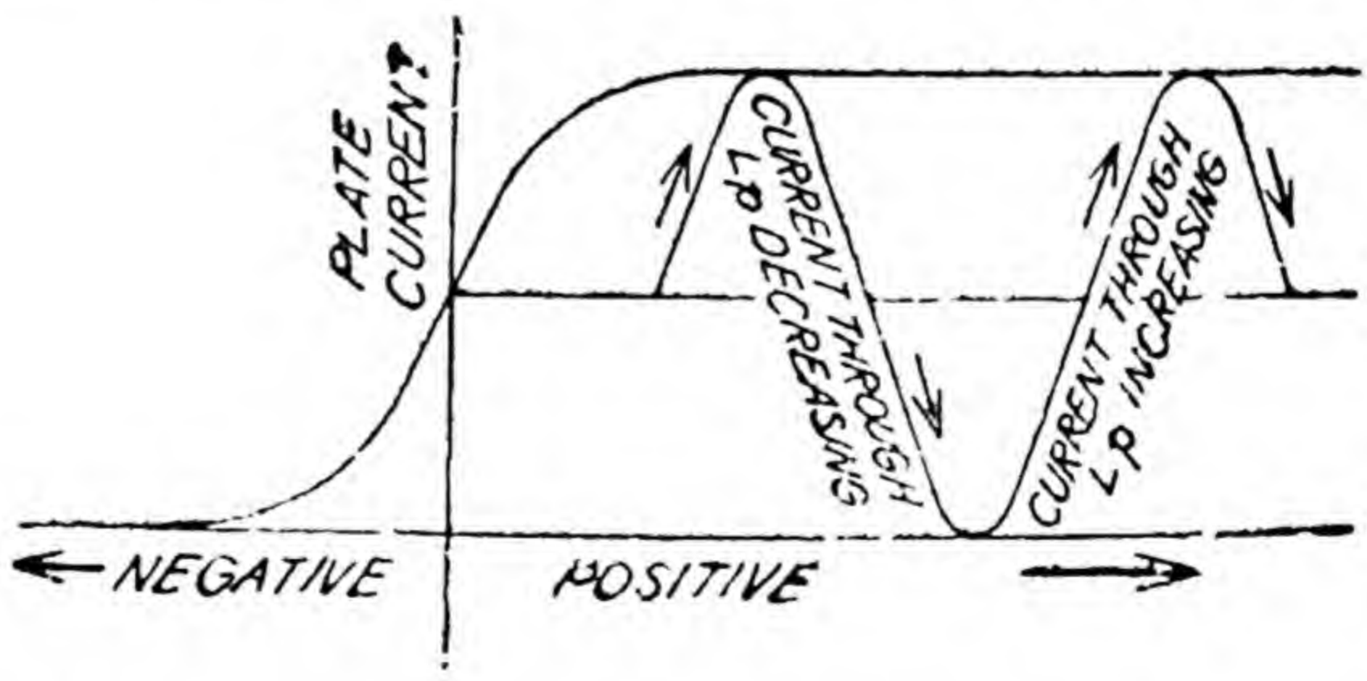


Fig. 4. Variation of plate current in the oscillator circuit.

will now start to discharge through the grid coil. This decrease in voltage across the capacitor makes the voltage on the grid less negative thus causing the plate current to increase and the magnetic field about the grid will make the grid more positive thereby increasing the plate current still further until saturation

is reached. The cycle of operation repeats itself continuously as long as energy is supplied to overcome the losses in the circuit. Oscillating energy from the grid circuit can thus be transferred to the output circuit.

15·3 TUNED PLATE OSCILLATOR :

Figure 5 (*a*) shows the circuit diagram of the tuned plate oscillator. In this oscillator, a parallel tuned circuit is connected directly in the plate circuit. Another coil inductively coupled to the coil of the tuned circuit provides the grid excitation. Here it is assumed that the grid current is zero. The circuit is just similar to tuned class *C* amplifier, but with the difference that this circuit provides its own grid excitation. This combination of grid leak resistor R_g and a grid capacitor C_g allows the self starting feature.

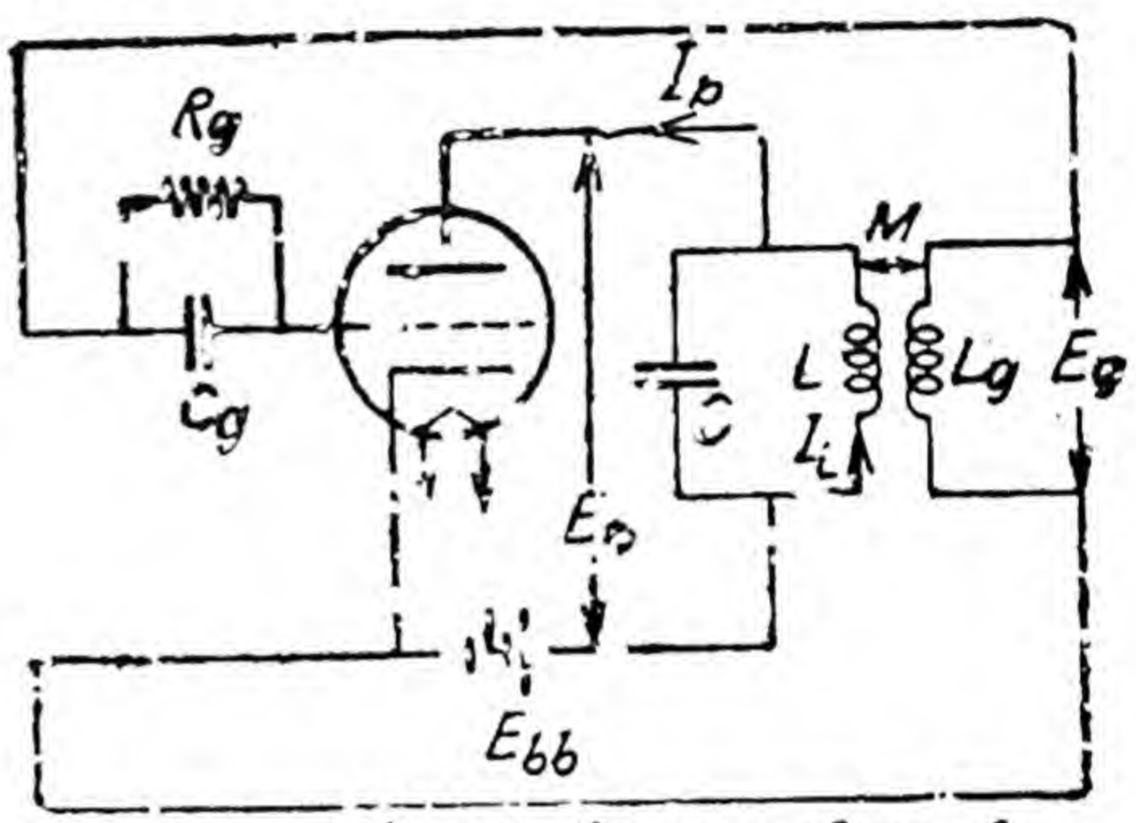


Fig. 5 (*a*). **Circuit diagram of tuned plate oscillator.**

The principle of operation is as follows :

For the time being, we neglect the R_g, C_g combination which simply provides the d.c. grid bias. When the circuit is closed, electrons flow from cathode to the plate and plate current flows through the capacitor *C* thus charges it. Here it should be remembered that virtually the capacitor acts as short circuit and inductance as open circuit. In the charging process of the capacitor, the lower plate remains positive. After a short duration, when the charge across the capacitor builds up, the plate current diminishes because of a fall in plate voltage which is equal to the difference of the plate voltage and the voltage across the capacitor The plate current becomes practically zero when the capacitor is fully charged. The charge of the capacitor now flows through inductance and it is transferred to a stored magnetic field of the coil. The voltage across the capacitor is now reversed. When the magnetic field of the coil is decreasing, the voltage induced in the grid coil is such that the grid becomes more negative than the bias voltage and thus prevents the flow of plate current. When the magnetic field of the coil has completely collapsed, the capacitor is charged fully with upper plate positive. Now the currents flows in the reverse direction *i.e.*, from the upper portion of the coil and thus increases the magnetic field in coil in the opposite direction. The current continues to flow with decreasing magnetic field and the capacitor will begin to charge with lower plate as positive. In this case, the voltage induced in the grid coil is such that the grid becomes positive. The plate current again flows and charges the capacitor *C*,

replenishing energy lost in resistance of the circuit during the previous part of oscillating cycle. The cycle is again repeated.

Let an a.c. voltage E_g be applied at the input of oscillator. This voltage is amplified by the tube and a current I_L flows through the inductance L in the plate circuit. Now due to mutual inductance M, a voltage $j\omega MI_L$ is induced between grid cathode of the tube. If this induced voltage is in the same phase and greater than or equal to the applied voltage E_g, sustained oscillations will be maintained in the circuit. We know that the tube makes a phase shift of 180° and again the transformer makes another 180° phase shift, hence the feed back voltage is applied in the same phase as the original voltage E_g. The oscillations in the circuit are developed due to positive feed back

Figure 5(*b*) shows the equivalent circuit of tuned plate oscillator. Now we shall write the differential equations involving instantaneous voltages and current in the circuit and solve them to obtain the conditions under which oscillations take place.

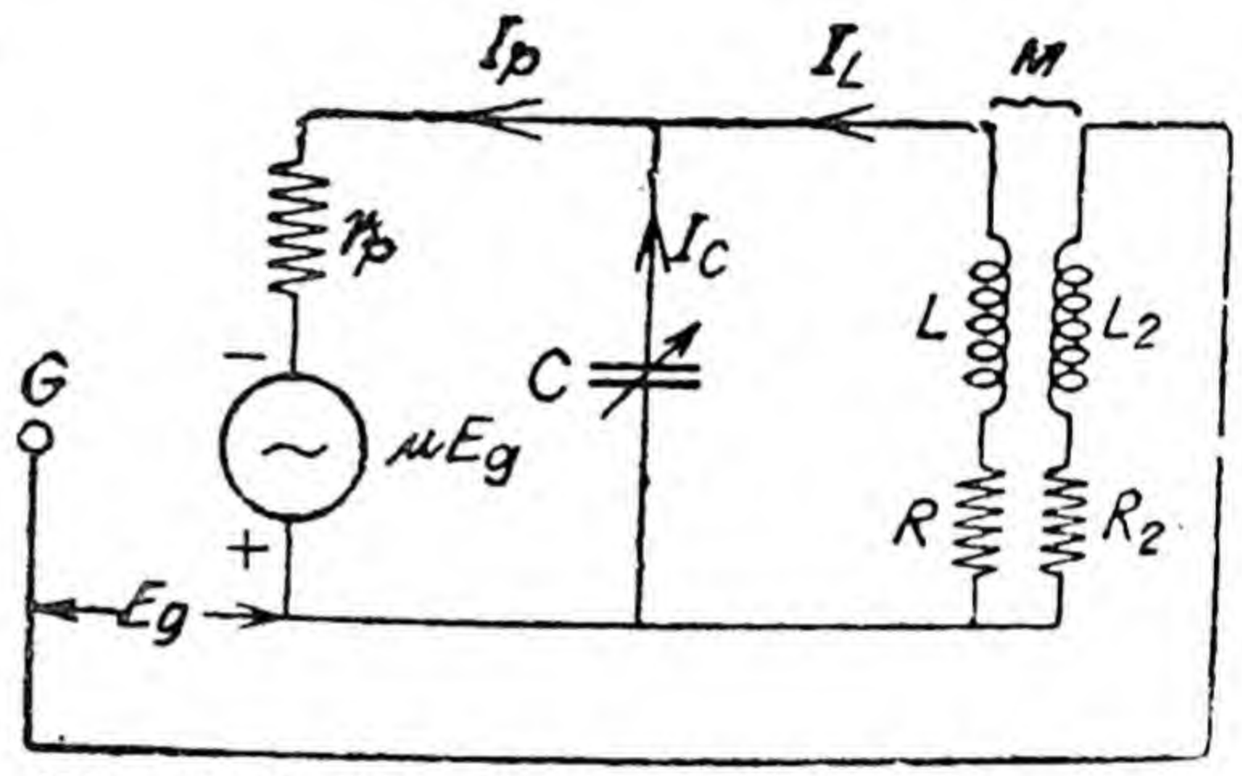


Fig. 5 (*b*) Equivalent circuit of tuned plate oscillator.

The loop equation of the plate circuit is

$$\mu E_g = r_p I_p + I_L (R + j\omega L). \quad \ldots(1)$$

For the tank circuit, we have

$$\frac{I_C}{j\omega C} = I_L (R + j\omega L). \quad \ldots(2)$$

and

$$E_g = j\omega M I_L. \quad \ldots(3)$$

$$I_p = I_L + I_C. \quad \ldots(4)$$

Substituting the values of E_g and I_p from equations (3) and (4) in equation (1), we have

$$\mu j\omega M I_L = r_p (I_L + I_C) + I_L (R + j\omega L)$$
$$= r_p I_L + r_p I_C + I_L (R + j\omega L)$$
$$= r_p I_C + I_L (r_p + R + j\omega L) \quad \ldots(5)$$

From equation (2)

$$I_C = I_L . j\omega C (R + j\omega L). \quad \ldots(6)$$

Substituting the value of I_C from equation (6) in equation (5), we get

$$\mu j\omega M I_L = r_p I_L j\omega C (R + j\omega L) + I_L (r_p + R + j\omega L)$$
$$j\omega M\mu = j\omega C . r_p . (R + j\omega L) + (r_p + R + j\omega L) \quad \ldots(7)$$

Equating the real part in equation (7), we have

$$-\omega^2 LCr_p + r_p + R = 0$$

or $$\omega^2=\frac{r_p+R}{r_pLC}=\frac{1}{LC}\left(1+\frac{R}{r_p}\right).$$

$$\therefore \quad \omega=\frac{1}{\sqrt{(LC)}}\left(1+\frac{R}{r_p}\right)^{1/2}$$

$$=\omega_0\left[1+\frac{R}{2r_p}\right], \text{ since } R \ll r_p.$$

where ω_0 is the constant angular frequency of the tuned circuit. If R is very small,

$$f=f_0\left(1+\frac{R}{2r_p}\right) \qquad ...(8)$$

So the frequency of oscillations is slightly greater than the resonant frequency of the tuned circuit.

Equating the imaginary part in equation (7), we have

$$\omega M\mu=\omega Cr_pR+\omega L$$

$$M=\frac{L+CRr_p}{\mu} \qquad ...(9)$$

This is the minimum value of the mutual inductance between tuned circuit and grid bias for maintained oscillations.

Again $$M=\frac{L}{\mu}+\frac{CRr_p}{\mu}$$

$$=\frac{L}{\mu}+\frac{CR}{g_m},$$

$$M-\frac{L}{\mu}=\frac{CR}{g_m}$$

or $$g_m=\frac{\mu CR}{M\mu-L}. \qquad ...(10)$$

Equation (10) gives the value of g_m for sustained oscillations in a tuned plate oscillator. If g_m is less than this value, oscillations die out and if g_m is more than this value, the oscillations continue to build up and then with R_gC_g bias, the point of operation shifts, until such a position is reached where g_m has the desired value.

Building up of grid bias voltage in the grid leak and capacitor circuit.

The self starting feature and a grid bias voltage of magnitude equal to or greater than the cut-off value are both realized with the help of grid leak and capacitor combination. Before the supply is first connected, the capacitor C_g is uncharged and the grid bias is zero. When the supply is connected, a surge of plate current passes which induces a sinusoidal voltage of small amplitude in grid coil. This voltage is applied to C_gR_g combination and grid cathode structure.

Drawing the positive half cycle, a pulse of grid current occurs and the capacitor C_g receives a charge. The action is similar to a half wave rectifier circuit with a resistive load and a smoothing capacitor. The capacitance of the capacitor should be large enough to have neglible reactance at the operating frequency incomparison with the grid leak resistance R_g. That is,

$$\frac{1}{\omega C_g} \ll R_g. \qquad \ldots(11)$$

Hence the time constant of the biasing circuit must satisfy the relation

$$R_g C_g \gg \frac{1}{2\pi f}.$$

$$\therefore\ R_g C_g \gg \frac{T}{2\pi}, \qquad \ldots(12)$$

where T is the period of oscillations.

During the other half cycle, the grid is negative with respec. to cathode, no grid current flows and the only path by which C_g may discharge is through the resistance R_g. Thus the charge of C_g leaks off through R_g under the condition 2. As the rate of leakage is small, some charge is accumulated on C_g during the cycle. This process is cumulative, the charge on the capacitor builds up, the negative grid bias voltage increases. and the grid current pluse decreases until a steady state is reached when the total charge received by the capacitor is equal to the total charge lost by it through R_g.

15·4. TUNED GRID OSCILLATOR :

Figure 6 (*a*) shows the circuit diagram of a tuned grid oscillator, also called the 'Tickler coil oscillator". In this case, the tunep circuit is connected between the grid and cathode of the valve, anp coupled to the anode by mutual inductance.

When the circuit operates, the plate current flows. The increasing plate current induces a current in the grid coil. This

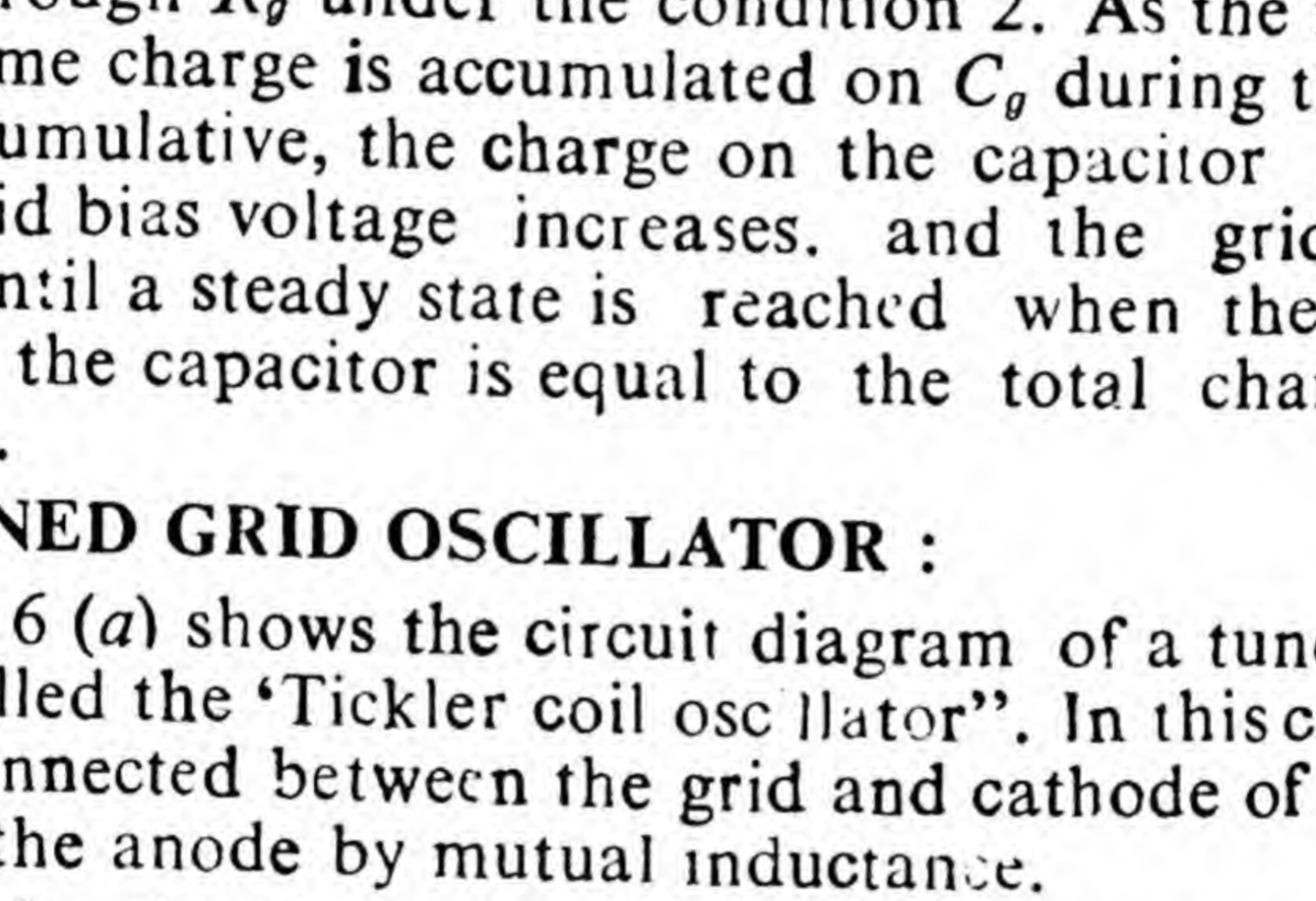


Fig. 6 (a). Tuned grid oscillator.

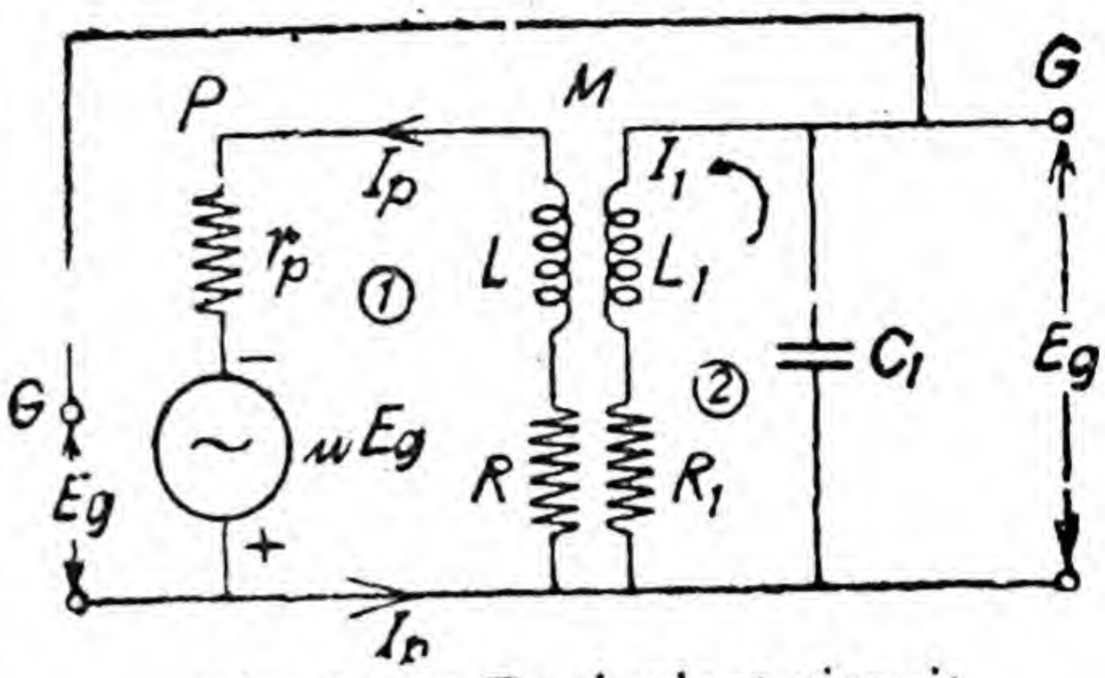


Fig. 6. (b) Equivalent circuit.

current charges the condenser in the tuned circuit in such a direction that the grid becomes more positive, thus increases the plate current further. This process will continue until the plate current reaches a saturation value. As there is no further change in current in plate coil, there is no further change in grid potential. The tuned circuit capacitor, which had charged earlier, would now discharge through tuned circuit coil. When the condenser discharges, a fall of potential takes place which makes the grid less and less positive with the result that plate current starts decreasing The decrease in plate current in the plate coil will induce a current in the tuned circuit coil, such that the tuned circuit capacitor will discharge more rapidly, resulting in grid going more negative and plate current decreasing. Now the plate current reaches the cut off value. The cycle of operation repeats itself at a frequency determined by the frequency of the tuned circuit. Any loss of power in dissipative elements of tuned circuit is made up (in magnitude and right phase) from the valve amplifier and oscillations can be maintained as self sustaining.

For the purpose of calculating the condition of oscillations and frequency generated by the oscillator, we assume that a voltage E_g is applied between the grid and cathode of the tube. Due to this voltage a current I_p flows through the inductance L. Due to mutual inductance a voltage E_g is induced between the grid and cathode. If this induced voltage is greater or equal to the input voltage and is in the same phase. sustained oscillations will be produced in the circuit. The feed back voltage is in the same phase with input voltage because the tube makes a phase shift of 180° and again the coupling transformer makes a phase shift of 180° *i.e.* a total phase shift of 360° takes place.

Applying Kirchhoff's law to meshes (1) ann (2).

$$\mu E_g = I_p (R + r_p + j\omega L) - j\omega M I_1 \qquad \ldots(1)$$

$$0 = I_1 \left[R_1 + j\omega L_1 + \frac{1}{j\omega C_1} \right] - j\omega M I_p. \qquad \ldots(2)$$

Again $$E_g = \frac{I_1}{j\omega C_1}$$

Putting the value of E_g in equation (1),

$$\mu \frac{I_1}{j\omega C_1} = I_p (R + r_p + j\omega L) - j\omega M I_1,$$

$$I_1 \left[\frac{\mu}{j\omega C_1} + j\omega M \right] = I_p (R + r_p + j\omega L). \qquad \ldots(3)$$

Also from equation (2),

$$I_1\left[R_1+j\omega L_1+\frac{1}{j\omega C_1}\right]=j\omega MI_p. \qquad \ldots(4)$$

Dividing equation (3) by (4), *i.e.* eliminating I_1 and I_p, we shall have complex quantity involving frequency and circuit constants, the real part of which will give the condition of oscillations and imaginary part will give the frequency of oscillations. Thus, we have

$$\frac{\frac{\mu}{j\omega C_1}+j\omega M}{R_1+j\omega L_1+\frac{1}{j\omega C_1}}=\frac{R+r_p+j\omega L}{j\omega M}$$

If $R \ll r_p$ and is neglected, we have

$$\frac{\frac{\mu}{j\omega C_1}+j\omega M}{R_1+j\omega L_1+\frac{1}{j\omega C_1}}=\frac{r_p+j\omega L}{j\omega M} \qquad \ldots(5)$$

On cross-multiplication,

$$\frac{\mu M}{C_1}-\omega^2M^2=r_p\left(R_1+j\omega L_1+\frac{1}{j\omega C_1}\right)+j\omega L\left(R_1+j\omega L_1+\frac{1}{j\omega C_1}\right) \ldots(6)$$

Equating the real parts, we have

$$\frac{\mu M}{C_1}-\omega^2M^2=R_1r_p-\omega^2LL_1+\frac{L}{C_1}.$$

$$\mu M-\omega^2M^2C_1=R_1r_pC_1-\omega^2C_1LL_1+L.$$

If the circuit is resonant, so that

$$\omega^2L_1C_1\simeq 1,$$

than $$\mu M\left(1-\frac{M}{L_1\mu}\right)=R_1r_pC_1, \text{ since } \omega^2=\frac{1}{L_1C_1}.$$

In general $\frac{M}{L_1\mu}$ is small compared with unity and so

$$M=\frac{R_1r_pC_1}{\mu}=\frac{C_1R_1r_p}{r_pg_m}.$$

$$M=\frac{C_1R_1}{g_m}.$$

This gives the maximum value of M required for the maintenance of oscillation, *i.e.*

$$g_m \geqslant \frac{C_1R_1}{M}$$

On equating the imaginary part in equation (6),

$$0=\omega L_1r_p-\frac{r_p}{\omega C_1}+R_1\omega L,$$

$$0=\omega^2L_1r_pC_1-r_p+R_1\omega^2LC_1,$$

$$r_p=\omega^2\left(L_1C_1r_p+R_1LC_1\right)$$

$$\omega^2=\frac{r_p}{L_1C_1r_p+R_1LC_1}$$

$$=\frac{1}{L_1C_1+\dfrac{R_1C_1L}{r_p}}$$

$$=\frac{1}{L_1C_1\left[1+\dfrac{R_1L}{L_1r_p}\right]}$$

$$\omega=\frac{1}{\sqrt{(L_1C_1)}\sqrt{\left(1+\dfrac{R_1L}{L_1r_p}\right)}}$$

$$=\frac{1}{\sqrt{(L_1C_1)}}\left[1+\frac{R_1L}{L_1r_p}\right]^{-1/2}$$

$$=\omega_0\left[1-\frac{R_1L}{2L_1r_p}\right]$$

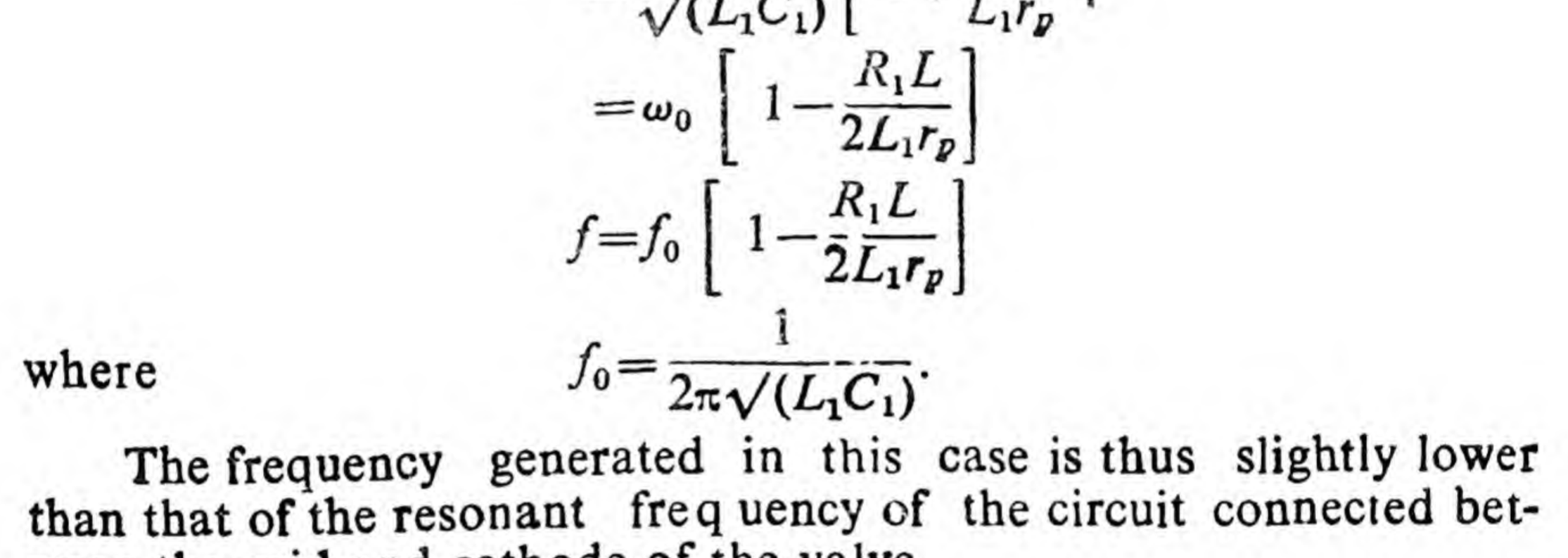
$$f=f_0\left[1-\frac{R_1L}{2L_1r_p}\right]$$

where $$f_0=\frac{1}{2\pi\sqrt{(L_1C_1)}}.$$

The frequency generated in this case is thus slightly lower than that of the resonant freq uency of the circuit connected between the grid and cathode of the valve.

15·5. HARTLEY OSCILLATOR :

The Hartley oscillator is one of the simplest type of oscillator circuits. The circuit arrangement can be shown in two forms, (*i*) series-fed circuit, and (*il*) parallel fed-circuit as shown in fig. 7.



(*a*) A series-fed circuit. (*b*) A parallel-fed circuit.

Fig. 7. Hartley oscillator circuit.

In Hartley oscillator only one coil is used, which is tapped such that a portion L_1 of the coil is in the plate circuit, while L_2 is in the grid circuit. The amplified energy in the plate section is fed back to the grid circuit by means of inductive coupling and amount of coupling will depend upon the number of turns in L_1 and L_2.

In series fed circuit, the plate power supply is connected in series with the plate section. In this case, oscillatory current flows through the high tension source. To remove this, high tension source is shunted by a capacitor C_1.

In parallel fed circuit, the plate circuit is divided into two parallel branches, one to provide a path for direct current and other for alternating current. R.F. choke keeps the alternating current out of the d.c. circuit while C_2 keeps the direct current out of a.c. circuit. In order that most of the r.f. current may flow through the plate section, instead of through the plate power supply, the reactance of C_2 at the resonant frequency should be small in comparison with the reactance of the choke coil at the same frequency. Thus the function of C_2 is to prevent the d.c. component of power supply to reach the tank circuit.

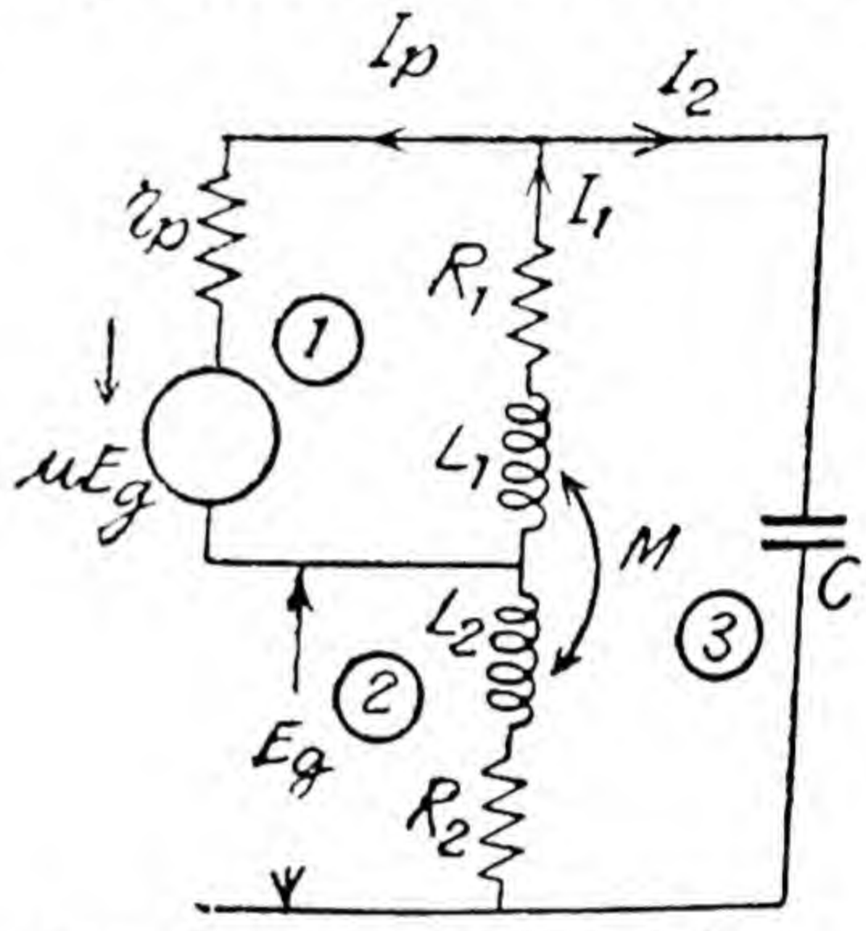


Fig. 8. Equivalent circuit.

If the presence of C_1, C_2 and r. f. choke is neglected, then the equivalent circuit for the two arrangements will be the same as given in fig. (8).

Applying the Kirchoff's law to meshes (1), (2) and (3), we get

$$\mu E_g = I_1 (R_1 + j\omega L_1) + I_2\, j\omega M + I_p r_p \quad \ldots(1)$$

$$E_g = I_2 (R_2 + j\omega L_2) + I_1\, j\omega M \quad \ldots(2)$$

$$0 = I_1 (R_1 + j\omega L_1) + I_2\, j\omega M + I_2 (R_2 + j\omega L_2) + I_1\, j\omega M + \frac{I_2}{j\omega C} \quad \ldots(3)$$

Again $$I_p = I_1 - I_2, \quad \ldots(4)$$

and form equation (2) and equation (4), putting the values of E_g and I_p in equation (1), we have

$$\mu I_2 (R_2 + j\omega L_2) + \mu I_1\, j\omega M = I_1 (R_1 + j\omega L_1) + I_2\, j\omega M + r_p (I_1 - I_2)$$

$$I_2 [(\mu R_2 + r_p) + j\omega (\mu L_2 - M)] = I_1 [R_1 + r_p + j\omega (L_1 - \mu M)]$$

$$\frac{I_2}{I_1} = \frac{R_1 + r_p + j\omega (L_1 - \mu M)}{\mu R_2 + r_p + j\omega (\mu L_2 - M)}. \quad \ldots(5)$$

From equation (3), we have

$$-I_1 [R_1 + j\omega (L_1 + M)] = I_2 \left[R_2 + j\omega (L_2 + M) - \frac{j}{\omega C} \right]$$

$$\frac{I_2}{I_1} = -\frac{R_1 + j\omega (L_1 + M)}{R_2 + j\omega (L_2 + M) - \frac{j}{\omega C}} \quad \ldots(6)$$

Equating equations (5) and (6),

$$\frac{R_1 + r_p + j\omega (L_1 - \mu M)}{\mu R_2 + r_p + j\omega (\mu L_2 - M)} = -\frac{R_1 + j\omega (L_1 + M)}{R_2 + j\omega (L_2 + M) - j/\omega C} \quad \ldots(7)$$

Cross multiplying and equating imaginary parts, we get

$$\omega R_2 (L_1-\mu M)+\omega(L_2+M)(R_1+r_p)-\frac{(R_1+r_p)}{\omega C}$$

$$=-\mu R_2\omega (L_1+M)-r_p\omega (L_1+M)-R_1\omega(\mu L_2-M).$$

$$R_2\omega^2 (L_1-\mu M)+\omega^2 (L_2+M)(R_1+r_p)+\mu R_2\omega^2 (L_1+M)$$

$$+r_p (L_1+M)\omega^2+R_2\omega^2 (\mu L_2-M)=\frac{R_1+r_p}{C},$$

$$\omega^2 [L_1R_2 (1+\mu)+R_1L_2 (1+\mu)+r_p (L_1+L_2+2M)=\frac{R_1+r_p}{C}$$

Now $L_1R_2(1+\mu)$ and $R_1L_2(1+\mu)$ can be neglected in comparison to $r_p(L_1+L_2+2M)$; hence we get

$$\omega^2=\frac{R_1+r_p}{Cr_p (L_1+L_2+2M)}$$

$$\omega=\sqrt{\left(\frac{R_1+r_p}{r_p}\right)}\sqrt{\left(\frac{1}{C(L_1+L_2+2M)}\right)}.$$

But

$$\frac{1}{\sqrt{[C(L_1+L_2+2M)]}}=\text{resonant frequency } f_r=\frac{1}{2\pi\sqrt{(LC)}},$$

because in this case of coupling $L=(L_1+L_2+2M)$.

Hence
$$2\pi f=\sqrt{\left(\frac{R_1+r_p}{r_p}\right)}.2\pi f_r,$$

$$f=\left(1+\frac{R_1}{r_p}\right)^{1/2} f_r,$$

$$f=\left(1+\frac{R_1}{2r_p}\right) f_r. \qquad \ldots(8)$$

Equating the real part in equation (6), we get

$$\mu=\frac{L_1+M}{L_2+M}+\frac{Cr_p (R_1+R_2)(L_1+L_2+2M)}{(L_1+M)(L_2+M)}.$$

If the resistance of the coil is negligible.

$$\mu=\frac{L_1+M}{L_2+M}.$$

The circuit can thus be adjusted for oscillations by the alternation of the tapping point on the coil.

15·6. COLPITT'S OSCILLATOR :

The principle of Colpitt's oscillator is similar to that of Hartley oscillator, but the change being that the cathode is connected to a tapping point on the capacitor instead of the coil.

As there is no conduction path between the grid and cathode of the valve, it is essential to provide a grid leak resistor R_g. In addition, a low reactance capacitor C' is required to block the direct potential of anode from reaching the grid. A high frequency choke is also necessary as the circuit is shunt fed. The grid excitation voltage is obtained from the grid tank capacitor C_2.

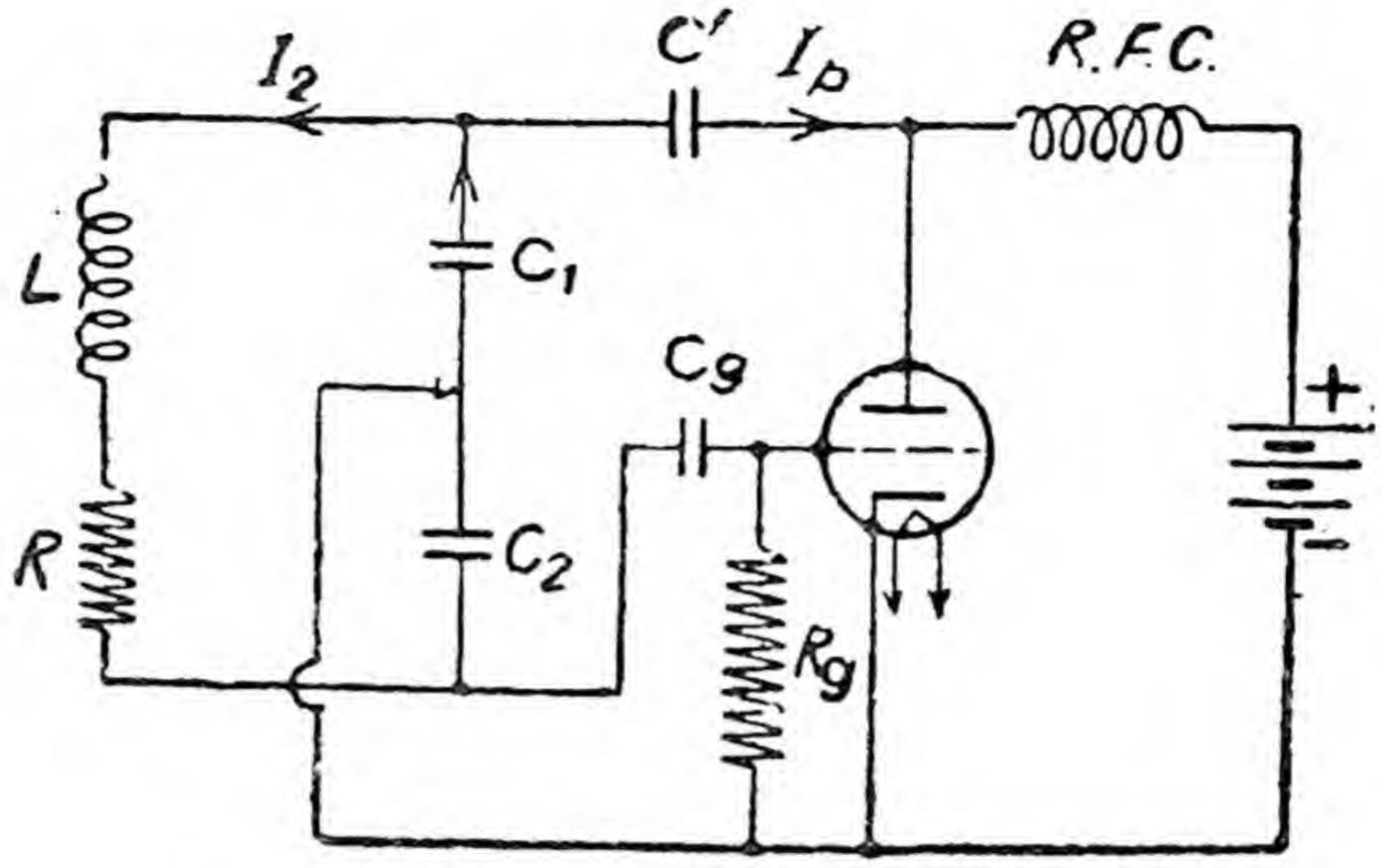


Fig. 9. The Colpitt's oscillator circuit.

In this oscillator, as plate current starts to flow, the plate blocking capacitor C' starts to charge, which in turn also causes the plate tank capacitor C_1 to charge. This increase in voltage is transferred to the grid tank capacitor C_2 through the r.f. choke, causing the terminal connected to the grid to become more negative. When the feed back voltage causes the plate current to decrease, less energy will be stored in the plate tank capacitor, thus reversing the direction of the feed-back voltage. As this action continues, sustained oscillations will be produced.

If the effect of the extra components *i e.* r.f. choke, condenser C' and resistance R_g be neglected, then the equivalent circuit is of the form as shown in fig. 10.

Applying the Kirchoff's second law to the various meshes shown in the fig. 10, we arrive at

$$\mu E_g = I_p r_p + \frac{I_1}{j\omega C_1} \quad \ldots (1)$$

$$E_g = \frac{I_2}{j\omega C_2} \quad \ldots (2)$$

$$0 = \frac{I_1}{j\omega C_1} + I_2 \left(R + j\omega L + \frac{1}{j\omega C_2} \right) \quad \ldots (3)$$

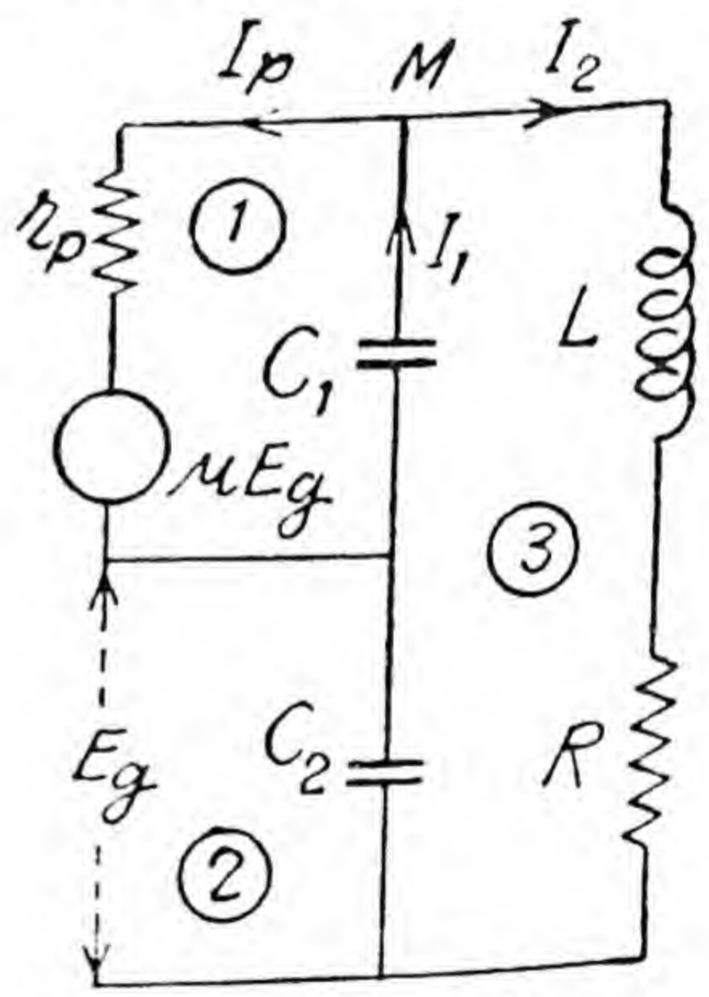


Fig. 10. Equivalent circuit.

Applying first law at the point M

$$I_p = I_1 - I_2. \quad \ldots (4)$$

Putting the value of E_g from eq. (2) and I_p from eq. (4) in eq. (1), we have

$$\mu \frac{I_2}{j\omega C_2} = (I_1 - I_2)\, r_p + \frac{I_1}{j\omega C_1},$$

$$I_2 \left(\frac{\mu}{j\omega C_2} + r_p \right) = I_1 \left(r_p + \frac{1}{j\omega C_1} \right).$$

$$\frac{I_2}{I_1}=\frac{r_p+\frac{1}{j\omega C_1}}{r_p+\frac{\mu}{j\omega C_2}} \qquad \ldots(5)$$

From equation (3), we have

$$-\frac{I_1}{j\omega C_1}=I_2\left(R+j\omega L+\frac{1}{j\omega C_2}\right),$$

$$\frac{I_2}{I_1}=-\frac{1}{j\omega C_1\left(R+j\omega L+\frac{1}{j\omega C_2}\right)} \qquad \ldots(6)$$

Equating equations (5) and (6), we have

$$\frac{r_p+\frac{1}{j\omega C_1}}{r_p+\frac{\mu}{j\omega C_2}}=-\frac{1}{j\omega C_1\left(R+j\omega L+\frac{1}{j\omega C_2}\right)}.$$

$$r_p\, j\omega C_1\left(R+j\omega L+\frac{1}{j\omega C_2}\right)+\left(R+j\omega L+\frac{1}{j\omega C_2}\right)=-\left(r_p+\frac{\mu}{j\omega C_2}\right).$$

$$-\omega^2 C_1C_2r_pR-r_p\, j\omega^3C_1C_2L+r_p\, j\omega C_1+j\omega C_2R-\omega^2LC_2+1$$
$$=-r_p\, j\omega C_2-\mu$$

Equating the imaginary parts,

$$-r_p\omega^3C_1C_2L+r_p\omega C_1+\omega C_2R=-r_p\omega C_2,$$
$$r_pC_1+C_2R+r_pC_2=\omega^2r_pC_1C_2L$$

$$\omega^2=\frac{r_pC_1}{r_pC_1C_2L}+\frac{C_2R}{r_pC_1C_2L}+\frac{r_pC_2}{r_pC_1C_2L}$$

$$=\frac{1}{C_2L}+\frac{R}{r_pC_1L}+\frac{1}{C_1L}$$

$$=\frac{C_1+C_2}{C_1C_2L}+\frac{R}{r_pC_1L}$$

$$=\frac{1}{C_1L}\left[\frac{C_1+C_2}{C_2}+\frac{R}{r_p}\right]$$

$$=\frac{1}{C_1L}\left(\frac{C_1+C_2}{C_2}\right)\left[1+\left(\frac{C_2}{C_1+C_2}\right)\frac{R}{r_p}\right]$$

$$\omega=\sqrt{\left(\frac{C_1+C_2}{C_1C_2L}\right)}\sqrt{\left\{1+\left(\frac{C_2}{C_1+C_2}\right)\frac{R}{r_p}\right\}}$$

and

$$f=\frac{1}{2\pi}\sqrt{\left(\frac{C_1+C_2}{C_1C_2L}\right)}\sqrt{\left\{1+\left(\frac{C_2}{C_1+C_2}\right)\frac{R}{r_p}\right\}}$$

$$f=f_0\sqrt{\left\{1+\left(\frac{C_2}{C_1+C_2}\right)\frac{R}{r_p}\right\}},$$

where f_0 is the resonant frequency of the tuned circuit, given by

$$f_0=\frac{1}{2\pi}\sqrt{\left(\frac{C_1+C_2}{LC_1C_2}\right)}.$$

The condition for the circuit to be oscillatory may be obtained by equating the real part. That is

$$(1+\mu)=Rr_p\omega^2C_1C_2+\omega^2LC_2.$$

$$1+\mu=\omega^2\,[Rr_pC_1C_2+C_2L]$$

$$=\frac{C_2+C_1}{C_1C_2L}\left[1+\left(\frac{C_2}{C_1+C_2}\right)\frac{R}{r_p}\right][Rr_pC_1C_2+C_2L]$$

$$=\left[1+\left(\frac{C_2}{C_1+C_2}\right)\frac{R}{r_p}\right]\left[\left(\frac{C_1+C_2}{C_1}\right)+\frac{Rr_p}{L}(C_1+C_2)\right]$$

Neglecting the term $\frac{C_2R_2}{(C_1+C_2)\,r_p}$, being very small, we get

$$1+\mu=\frac{C_1+C_2}{C_1}+\frac{Rr_p}{L}\,(C_1+C_2).$$

$$\mu=\frac{C_2}{C_1}+\frac{Rr_p}{L}\,(C_1+C_2).$$

This is the condition for the circuit to be oscillatory.

15·7. DYNATRON OSCILLATOR :

It makes use of that portion of the plate characteristics of a tetrode which displays negative resistance. Fig. 11 shows the plate characteristics of tetrode for a fixed value of screen grid voltage.

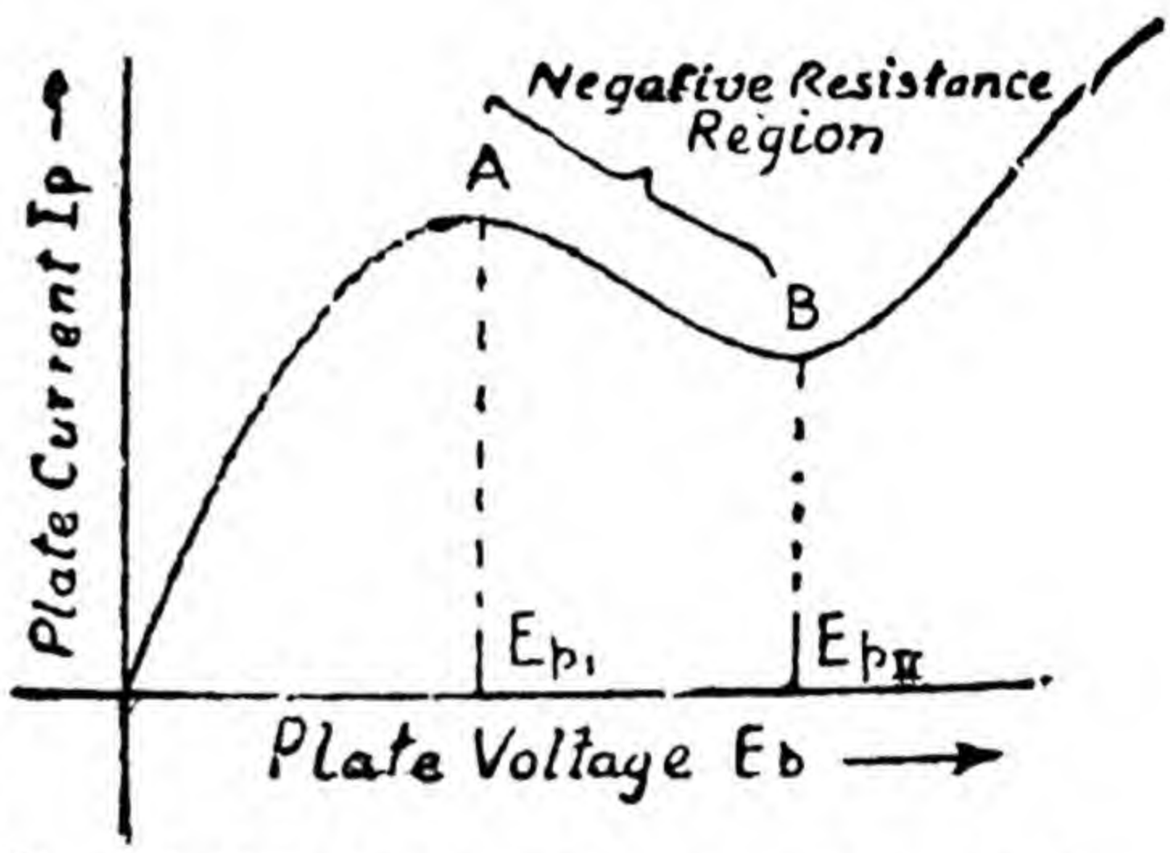


Fig. 11. Negative resistance characteristics of a tetrode.

As the plate voltage is increased from zero, the plate current increases normally. At a particular value of plate voltage E_{P1}, secondary emission from the plate begins and these electrons are

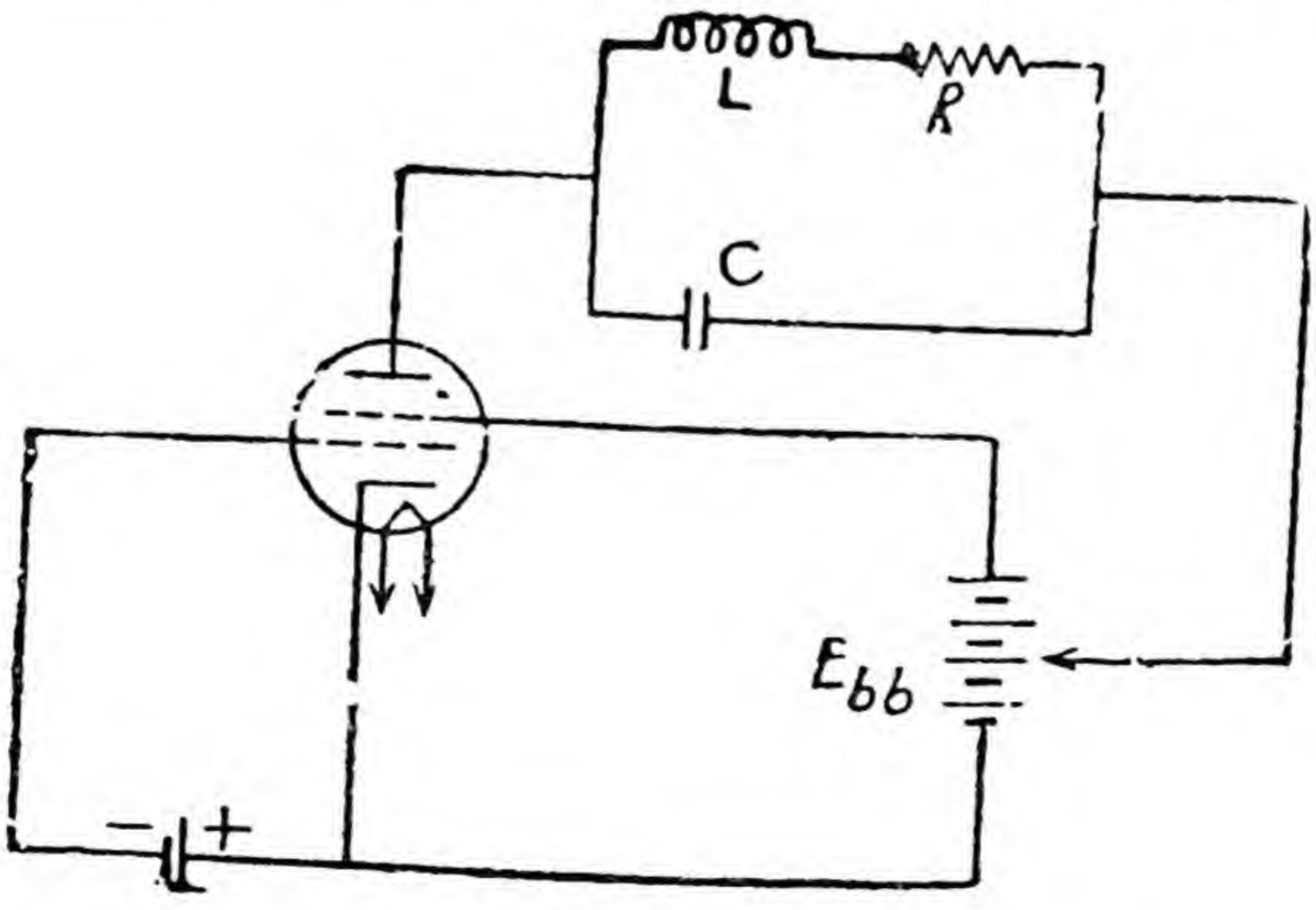


Fig. 12. Circuit of a dynatron oscillator.

collected by the screen grid because screen grid voltage is then much higher than plate voltage. The number of secondary electrons so lost by the plate, may exceed the number of primary electrons so that the net plate current actually decreases with the increase of plate voltage. Beyond plate voltage E_{g1} secondary electrons are drawn back to the plate because this potential is sufficiently large to hold them. The negative resistance portion AB of the tetrode is used in dynatron oscillator, circuit of which is given in fig. 12.

This oscillator operates with a fixed grid and screen voltages. The plate circuit contains a parallel LCR circuit. As there is no alternating current in plate and r_p is negative, equivalent circuit is given in fig. 13.

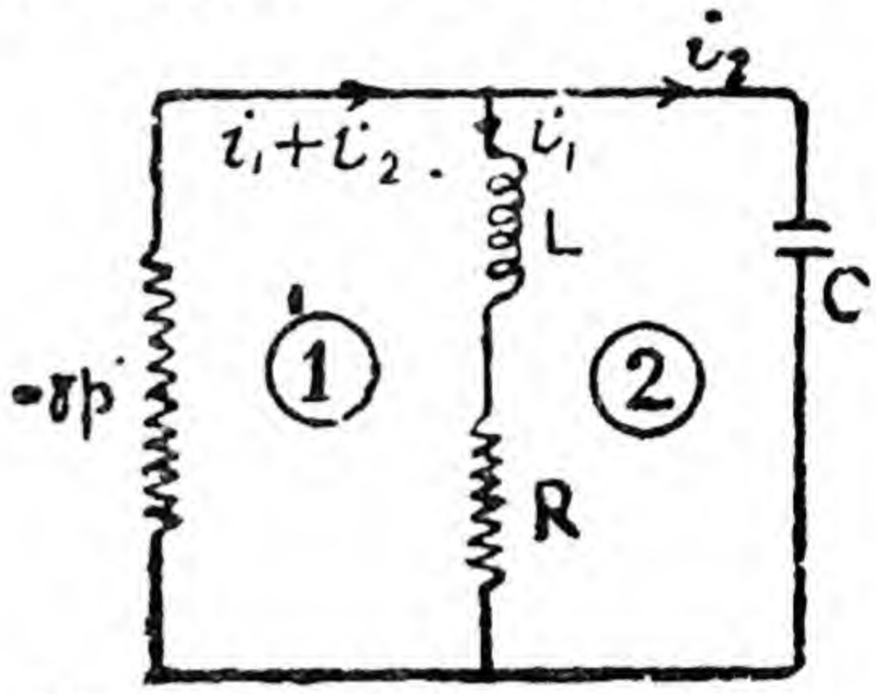


Fig. 13. Equivalent circuit.

Applying Kirchoff's law to the two loop network,

$$-r_p(i_1+i_2)+L\frac{di_1}{dt}+i_1R=0, \quad (1)$$

$$L\frac{di_1}{dt}+i_1R-\frac{1}{C}\int i_2\,dt=0 \quad \ldots(2)$$

On differentiating equation (2), we get

$$L\frac{d^2i_1}{dt^2}+R\frac{di_1}{dt}-\frac{1}{C}i_2=0.$$

Multiplying by C,

$$LC\frac{d^2i_1}{dt^2}+CR\frac{di_1}{dt}=i_2. \quad \ldots(3)$$

Again multiplying by r_p, we have

$$LCr_p\frac{d^2i_1}{dt^2}+CRr_p\frac{di_1}{dt}=i_2r_p.$$

Putting the value of i_2r_p from equation (1), we write

$$LCr_p\frac{d^2i_1}{dt^2}+r_pCR\frac{di_1}{dt}=L\frac{di_1}{dt}+i_1R-i_2r_p$$

$$LCr_p\frac{d^2i_1}{dt^2}+(r_pCR-L)\frac{di_1}{dt}=i_1(R-r_p)$$

$$\frac{d^2i_1}{dt^2}+\frac{r_pCR-L}{LCr_p}\cdot\frac{di_1}{dt}+\frac{(r_p-R)}{LCr_p}i_1=0. \quad \ldots(4)$$

For sustained oscillations, the coefficient of $\frac{di_1}{dt}$ should be zero,

i.e.

$$\frac{r_pCR-L}{LCr_p}=0.$$

$$r_pCR-L=0,$$

$$r_p=\frac{L}{CR} \quad \ldots(5)$$

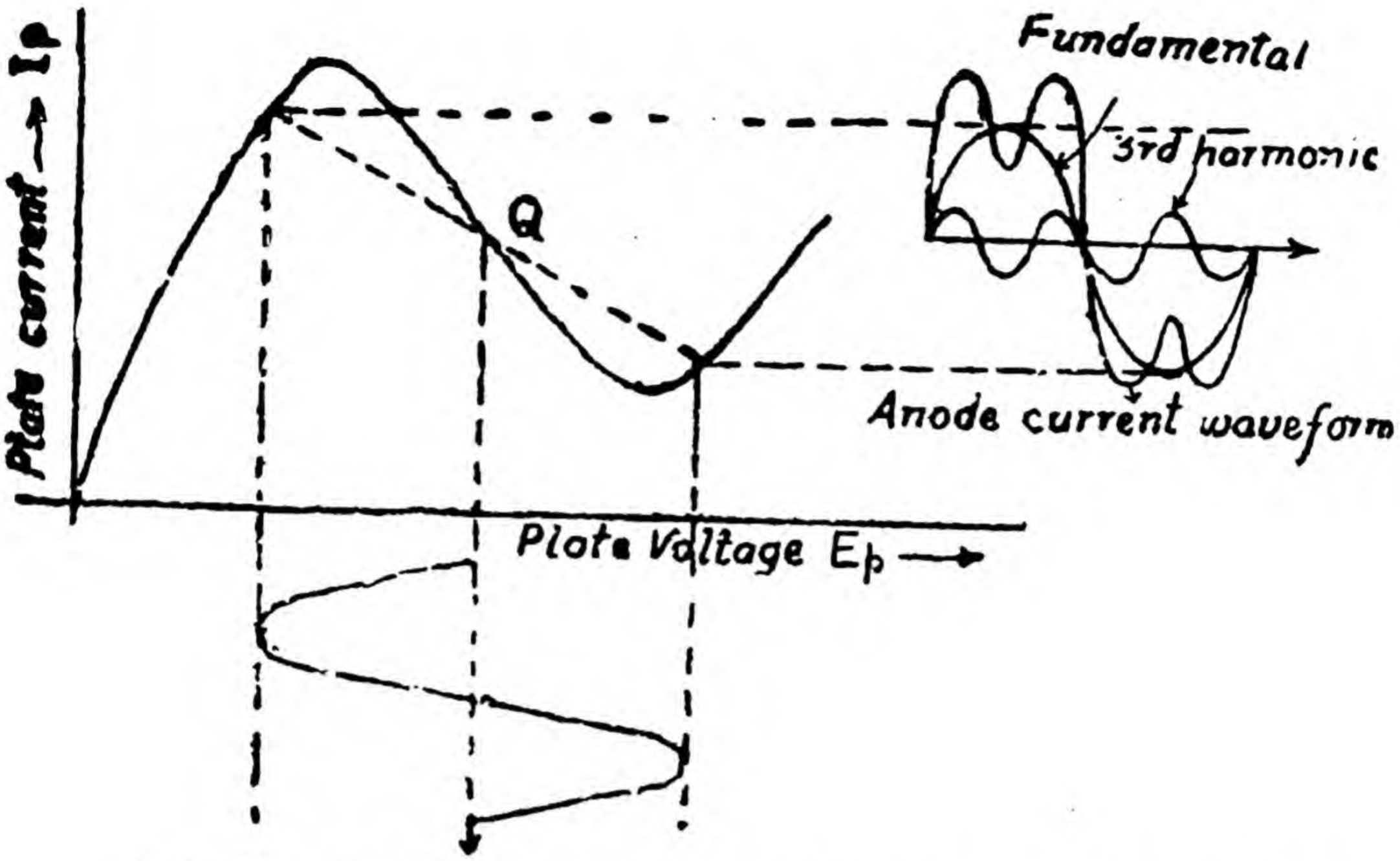


Fig. 14. Current and voltage waveforms of dynatron oscillator.

Figure 14 shows the current and voltage waveforms of a dynatron oscillator. If r_p has a value greater than $\frac{L}{RC}$, the amplitude will decay exponentially and oscillations will not be sustained. If r_p has a value smaller then $\frac{L}{RC}$, the amplitude will increase and finally when $r_p = \frac{L}{RC}$, the oscillations will be sustained. The limiting condition for sustained oscillations is

$$r_p \leqslant \frac{L}{CR}$$

The frequency of oscillations will be

$$\omega^2 = \frac{r_p - R}{r_p LC} \text{ (coefficient of } i_1)$$

$$\omega = \frac{1}{\sqrt{(LC)}} \sqrt{\left(\frac{r_p - R}{r_p}\right)}$$

$$f = \frac{1}{2\pi\sqrt{(LC)}} \sqrt{\left(1 - \frac{R}{r_p}\right)}. \qquad \ldots(6)$$

The condition (5) of sustained oscillations and frequency expression (6) of oscillations can also be derived by applying Kirchoff's law to the meshes (1) and (2).

For mesh (1)

$$-r_p (i_1 + i_2) + i_1 (R + j\omega L) = 0. \qquad \ldots(7)$$

For mesh (2),

$$(R+j\omega L)\, i_1=\frac{i_2}{j\omega C}. \qquad \ldots(8)$$

From equation (7), we have

$$-r_p i_1-r_p i_2+i_1\,(R+j\omega L)=0. \qquad \ldots(9)$$

Substituting the value of i_2 from equation (8) in equation (9) we get

$$-r_p i_1-r_p\,(R+j\omega L)\,(j\omega C)\, i_1+i_1\,(R+j\omega L)=0$$

or

$$-r_p-r_p\,(R+j\omega L)\,(j\omega C)+(R+j\omega L)=0$$

$$-r_p-r_p Rj\omega C+r_p\omega^2 LC+R+j\omega L=0. \qquad \ldots(10)$$

Equating the imaginary part on both the sides of equation (10) we get

$$-r_p R\omega C+\omega L=0$$

or

$$r_p=\frac{L}{CR}.$$

Again equating the real part on both the sides of equation (10) we get

$$-r_p+r_p\omega^2 LC+R=0$$

or

$$\omega^2=\frac{r_p-R}{r_p LC}$$

$$=\frac{1}{LC}\left(1-\frac{R}{r_p}\right)$$

$$\therefore\quad f=\frac{1}{2\pi\sqrt{(LC)}}\left(1-\frac{R}{r_p}\right)^{1/2}$$

The frequency stability of dynatron oscillator is very high because the cathode and plate capacity which can affect the frequency of oscillations is very small. In this case there is no variable coupling factor between the grid and plate. The main advantage of dynatron oscillator is its dependence upon the secondary emission.

15·8. RC OSCILLATORS :

As the name suggests, *RC* oscillators employ only the resistors and capacitors in network ; no inductor is used in such a type of oscillators. Such circuits have an advantage that they can be used for very low frequencies, where *LC* resonant circuits, would be too much bulky and expensive. The basic principle of using an amplifying device and feeding back a positive signal of sufficient amplitude to overcome the losses still applies. There are two types of these oscillators : *(i)* Phase shift oscillator, and *(ii)* Wien bridge oscillator (*RC* tuned oscillator).

(i) **Phase shift oscillator** : This oscillator may be considered as class *A* type amplifier with resistive load, in which the output

developed across the load resistance R_L is fed back to the grid through a phase shift network which introduces a phase shift of 180°. The phase 180° is produced because each L section consisting of an RC combination makes a phase shift of 60° at the oscillating frequency. Since the amplifier itself introduces a phase shift of 180°, total phase shift in the feedback loop is 360°. Oscillations may, therefore, be produced. The feed back network consists of three sections of RC network, each section contains one resistance and one capacitance in series.

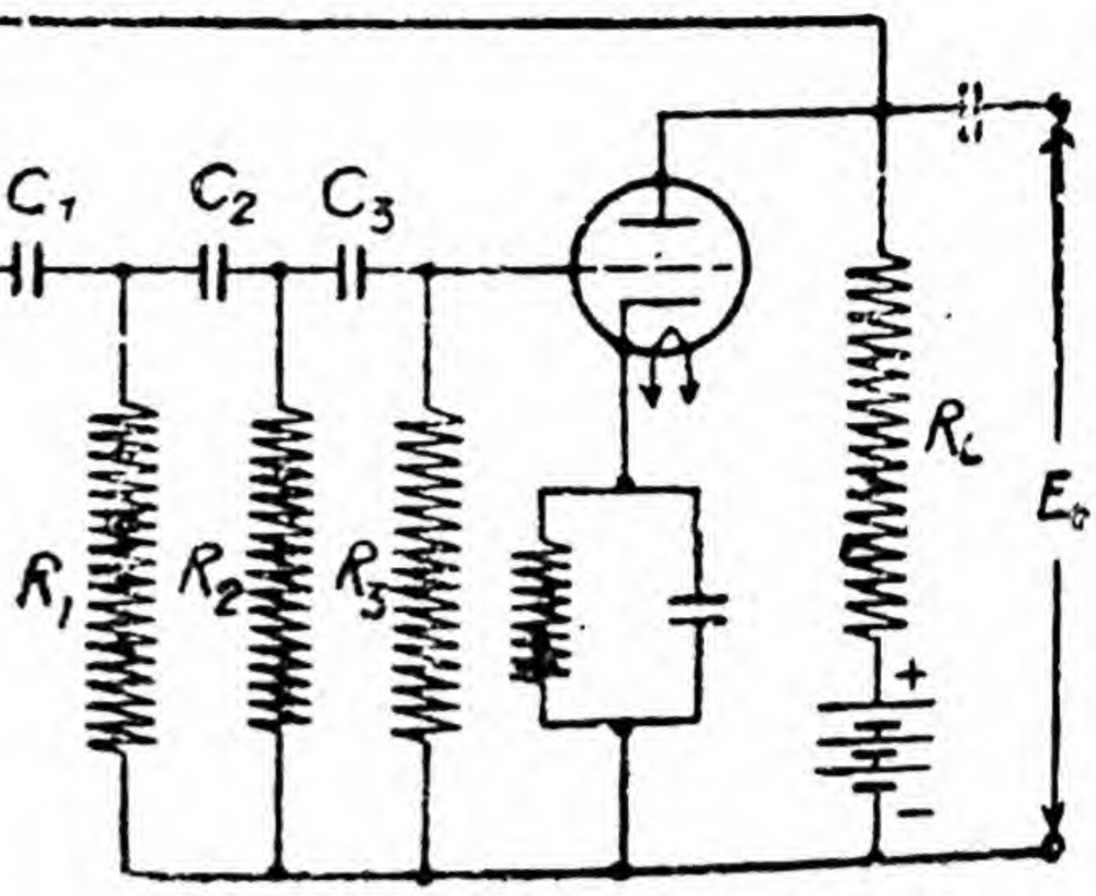


Fig. 15. (*a*) Phase shift oscillator.

For the frequency of oscillations, consider the current distribution in the phase shift network as shown in the figure 15 (*b*). It is assumed that

$$R_1 = R_2 = R_3 = R.$$

and $$C_1 = C_2 = C_3 = C$$

E_1 is potential between plate and cathode, and E_2 the potential between grid and cathode.

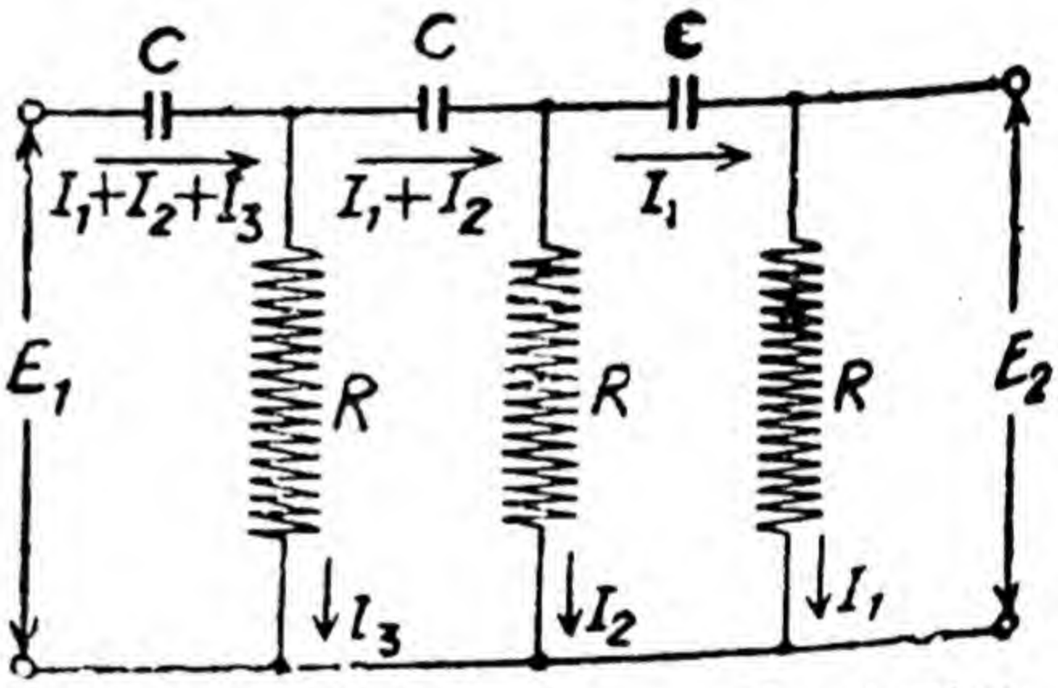


Fig. 15 (*b*) Current distribution in phase shift network.

Then $$I_1 R = E_2. \qquad \ldots(1)$$

$$I_2 R = I_1 \left(R + \frac{1}{j\omega C} \right), \qquad \ldots(2)$$

producing a phase shift of 60°. The frequency of oscillations can be changed by varying C or R values. In some commercial applications the R values are changed in steps ($\times 1$, $\times 10$, $\times 100$), and a variable ganged capacitor is used to select the specific frequency with these step values. Further

$$I_3 R = (I_1 + I_2) \frac{1}{j\omega C} + I_2 R \qquad \ldots(3)$$

and $$E_1 = \frac{(I_1 + I_2 + I_3)}{j\omega C} + I_2 R \qquad \ldots(4)$$

From equations, (1), (2) and (3) substituting the values of I_1, I_2, and I_3 in equation (4), we get

$$E_1=\frac{I_1+I_2+I_3}{j\omega C}+\frac{(I_1+I_2)}{j\omega C}+\frac{I_1}{j\omega C}+E_2$$

$$=\frac{3I_1+2I_2+I_3}{j\omega C}+E_2. \qquad \ldots(A)$$

From (1), $I_1=\frac{E_2}{R}$. ...(5)

From equation (2),

$$I_1=\frac{I_1}{R}\left(R+\frac{1}{j\omega C}\right)=\frac{E_2}{R}+\frac{E_2}{R^2 j\omega C}$$

$$=E_2\left(\frac{1}{R}+\frac{1}{j\omega CR^2}\right) \qquad \ldots(6)$$

$$I_3=I_2+\frac{I_1+I_2}{j\omega CR}=E_2\left[\left\{\frac{1}{R}+\frac{1}{j\omega CR^2}\right\}+\frac{1}{j\omega RC}\left\{\frac{1}{R}+\frac{1}{R}+\frac{1}{j\omega CR^2}\right\}\right],$$

$$I_3=E_2\left[\frac{1}{R}+\frac{3}{j\omega CR^2}+\frac{1}{j^2\omega^2C^2R^3}\right]. \qquad \ldots(7)$$

Substituting the values of I_1, I_2 and I_3 from equations (5), (6) and (7) into equation (A), we get

$$\frac{1}{\alpha}=\frac{E_1}{E_2}$$

$$=1+\frac{\frac{3}{R}+2\left(\frac{1}{R}+\frac{1}{j\omega CR^2}\right)+\left(\frac{1}{R}+\frac{3}{j\omega CR^2}+\frac{1}{j^2\omega^2C^2R^3}\right)}{j\omega C}$$

$$=1-j\,\frac{6}{\omega CR}-\frac{5}{\omega^2C^2R^2}+\frac{j}{\omega^3C^3R^3}$$

$$\frac{1}{\alpha}=\left[1-\frac{5}{\omega^2C^2R^2}\right)+j\left(\frac{1}{\omega^3C^3R^3}-\frac{6}{\omega CR}\right), \qquad \ldots(B)$$

where α is fed-back ratio.

Let A be the voltage gain represented by

$$A=A_r+jA_i, \qquad \ldots(8)$$

where A_r is the real component and A_i is the imaginary component of voltage gain.

For a given circuit, assuming the effect of phase shift network on R_L to be negligible, we have

$$A=-\frac{\mu R_L}{r_p+R_L}$$

Obviousiy, $A_i=0$ and $A_r=\frac{-\mu R_L}{r_p+R_L}$. ...(9)

For oscillations to be sustained,

$$A\alpha=1 \text{ or } \frac{1}{\alpha}=A.$$

Hence $\left[1-\frac{5}{R^2\omega^2C^2}\right]+j\left[\frac{1}{\omega^3C^3R^3}-\frac{6}{\omega CR}\right]=A_v$. ...(10)

Equating imaginery parts in equation (10), we have

$$\frac{1}{\omega^3C^2R^3}-\frac{6}{\omega CR}=0,$$

$$\omega^2=\frac{1}{6R^2C^2},$$

$$\omega=\frac{1}{\sqrt{6}.RC};$$

then frequency $$f=\frac{1}{2\pi\sqrt{6}.RC}.$$

These phase shift oscillations are used only as fixed frequeney oscillators because the variation of resistance R and capacitance C in the phase shift network is troblesome.

Advantages of these oscillators :

(*a*) They may be constructed with simple and cheap components.

(*b*) They can be made to operate at low frequencies.

(*c*) The frequency depends upon the product of R and C. This enables a wide range of frequencies to be obtained.

(*d*) Phase shift oscillators operate in class A conditions.

(ii) Wien Bridge Oscillators : Wien bridge oscillator is a two stage RC coupled amplifier in which a fraction of the output voltage is fed back to the input terminals. Oscillations are produced if the feed back voltage is equal in magnitude and phase with the voltage initially assumed at the input of the first stage. Figure 16 shows the circuit diagram.

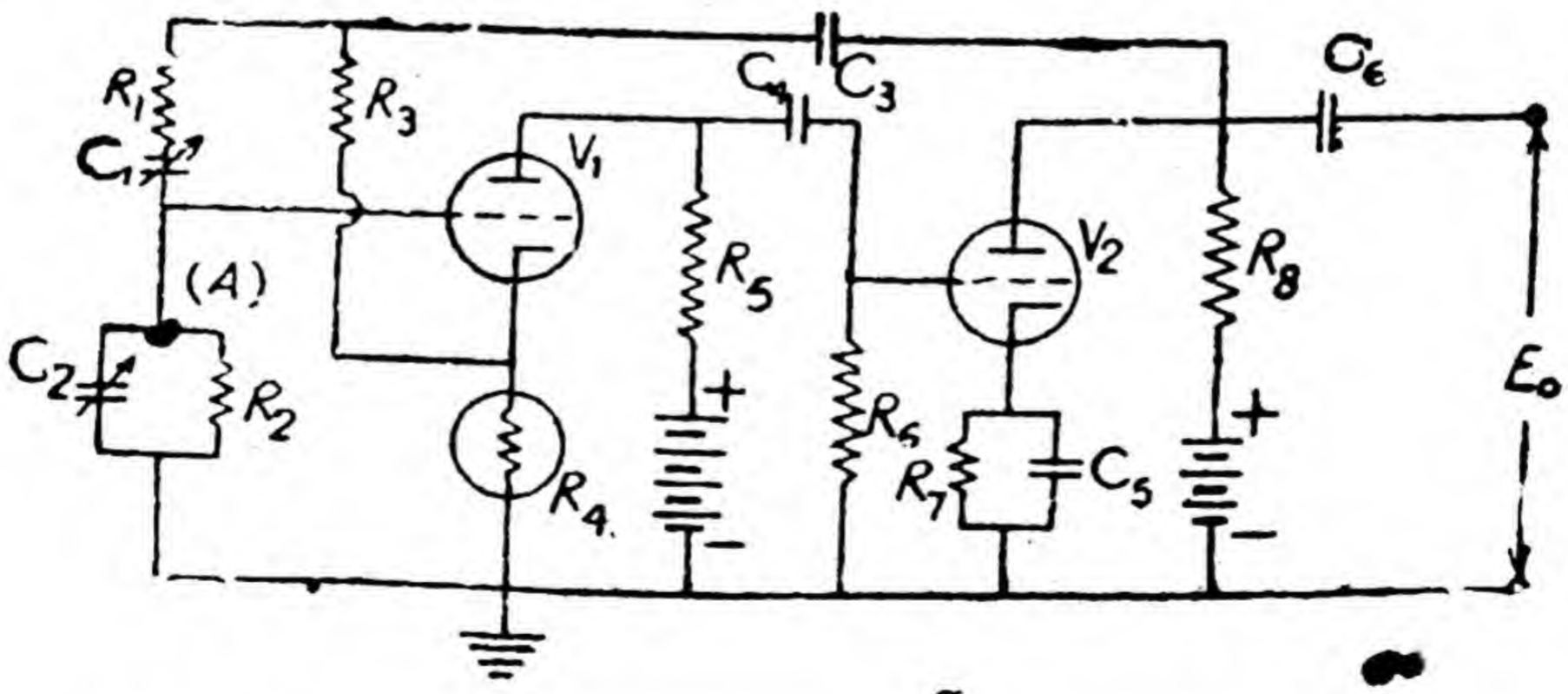


Fig. 16. Circuit diagram of Wien Bridge oscillator.

Resistors R_3 and R_4 are used to stabilize the amplitude of the output. Resistors R_4 is either a tungsten lamp or thermister. In

either case, its resistance increases with current. Feed-back from V_2 is applied through C_3 and R_3 to the cathode of V_1. This constitutes negative feed-back.

The frequency selective circuit consists of the series element R_1C_1 and the parallel element R_2C_2. The voltage developed between their junction (point A) and the ground is applied as input to the two stage amplifier. Depending on these component values, there will be one frequency at which this input voltage is in phase and of sufficient magnitude to sustain oscillations.

Let E_f be a fraction of the output voltage E_0, feed-back to the input of the amplifier so that we have

$$\frac{E_f}{E_0}=\frac{\text{impedance in } E_f \text{ branch}}{\text{impedance in } E_0 \text{ branch}} \qquad \ldots(1)$$

$$\frac{E_1}{E_0}=\frac{R_2/j\omega C_2}{R_2+\frac{1}{j\omega C_2}}-\times\frac{1}{R_1+\frac{1}{j\omega C_1}+\frac{R_2/j\omega C_2}{R_2+\frac{1}{j\omega C_2}}}$$

or
$$\frac{E_0}{E_1}=\left[R_1+\frac{1}{j\omega C_1}+\frac{R_2/j\omega C_2}{R_2+(1/j\omega C_2)}\right]\times\frac{R_2+(1/j\omega C_2)}{R_2/j\omega C} \qquad \ldots(2)$$

$$=\left[R_1+\frac{1}{j\omega C_1}+\frac{R_2}{R_2\,j\omega C_2+1}\right]\times\left[\frac{R_2\,j\omega C_2+1}{R_2}\right]$$

$$=\frac{R_1}{R_2}(R_2\ j\omega C_2+1)+\frac{(R_2\,j\omega C_2+1)}{j\omega C_1R_1}+1$$

$$=j\omega C_2R_1+\frac{R_1}{R_2}+\frac{C_1}{C_2}+\frac{j}{\omega C_1\,R_2}+1.$$

$$=j\omega C_2R_1+\frac{R_1}{R_2}+\frac{C_1}{C_2}-\frac{j}{\omega C_1\,R_2}+1. \qquad \ldots(3)$$

As the amplifier behaves to have zero phase shift, E_0 and E_f will be in the same phase ; hence imaginary part should be zero ;

$$\omega C_2R_2=\frac{1}{\omega C_1R_2};\ \omega^2C_1C_2R_1R_2=1\ ;\ \omega=\frac{1}{\sqrt{(R_1R_2C_1C_2)}} \qquad \ldots(4)$$

then
$$f=\frac{1}{2\pi\sqrt{(R_1R_2C_1C_2)}}. \qquad \ldots(5)$$

Equating real parts, $\frac{E_0}{E_f}=A=\frac{R_1}{R_2}+\frac{C_1}{C_2}+1.$ $\ldots(6)$

where A is voltage amplification.

Advantages. It is used for the production of frequency due to following advantages :

(*i*) it contains no inductor,
(*ii*) low distortion,
(*iii*) good frequency stability and
(*iv*) constant output voltage over a wide frequency range.

15·9. FREQUENCY STABILITY OF OSCILLATORS :

Due to the variation in the operating characteristic either due

to tube or circuit parameters, vacuum tube oscillators suffer changes in frequency. For most of the purposes the frequency of the oscillator must remain constant. Now we shall consider the factors which affect the frequency of the oscillator.

(i) Due to variation of supply voltage : When the supply voltage changes, the frequency of the oscillator also changes. The frequency change can be overcome by using a regulated power supply.

(ii) Due to variation in plate load : When there is any variation in plate load, the frequency of oscillator changes. This can be eliminated by using an amplifier to isolate the load from the oscillator. This system is called *a master oscillator power amplifier.*

(iii) Variation in inductance and capacitance of the tank circuit elements due to change in temperature : Temperature changes affect slightly the inductance and capacitance of the tuned circuit which changes the frequency of the oscillator. The change can be minimised by a careful choice of the inductor and capacitor either with negligible temperature coefficient or with such a temperature variation which is compensated for the elements.

(iv) Due to variation in the tube elements : Due to temperature changes, the element of the tube contract and expand. Due to contraction and expansion of tube elements, the interelectrode capacities vary. As these capacities are the part of the tuned circuit. the frequency of the oscillator will be affected.

(v) Due to the antiresonant frequency of the output circuit : This occurs when the phase shift in the feed back circuit is not exactly 180°. The maximum frequency stability occurs when the phase shift is exactly 180°.

15·10. ELECTRON COUPLED OSCILLATOR :

For an oscillator, high stability of output frequency is usually required. This can be achieved by the electron coupled oscillator. This provides the isolation between the load and tank circuit by using electron coupling between them. Figure 17 shows the electron coupled oscillator using a triode.

The cathode grid, and screen constitute a triode Hartley oscillator with one difference, namely, the radio frequency grounding of screen instead of cathode. The load is connected in the plate circuit. The name electron coupled oscillator is given due to the fact that the oscillator and plate circuits are coupled solely by the stream of electrons between screen grid and plate.

As the screen grid is positive, electrons flow from cathode to screen grid. The oscillations are generated as already discussed. The frequency is determined by the values of C and L in the grid tank circuit. Although a small number of electrons are intercepted by the screen grid, even then they are sufficient to maintain oscillations in the tank circuit. The remaining electrons pass through

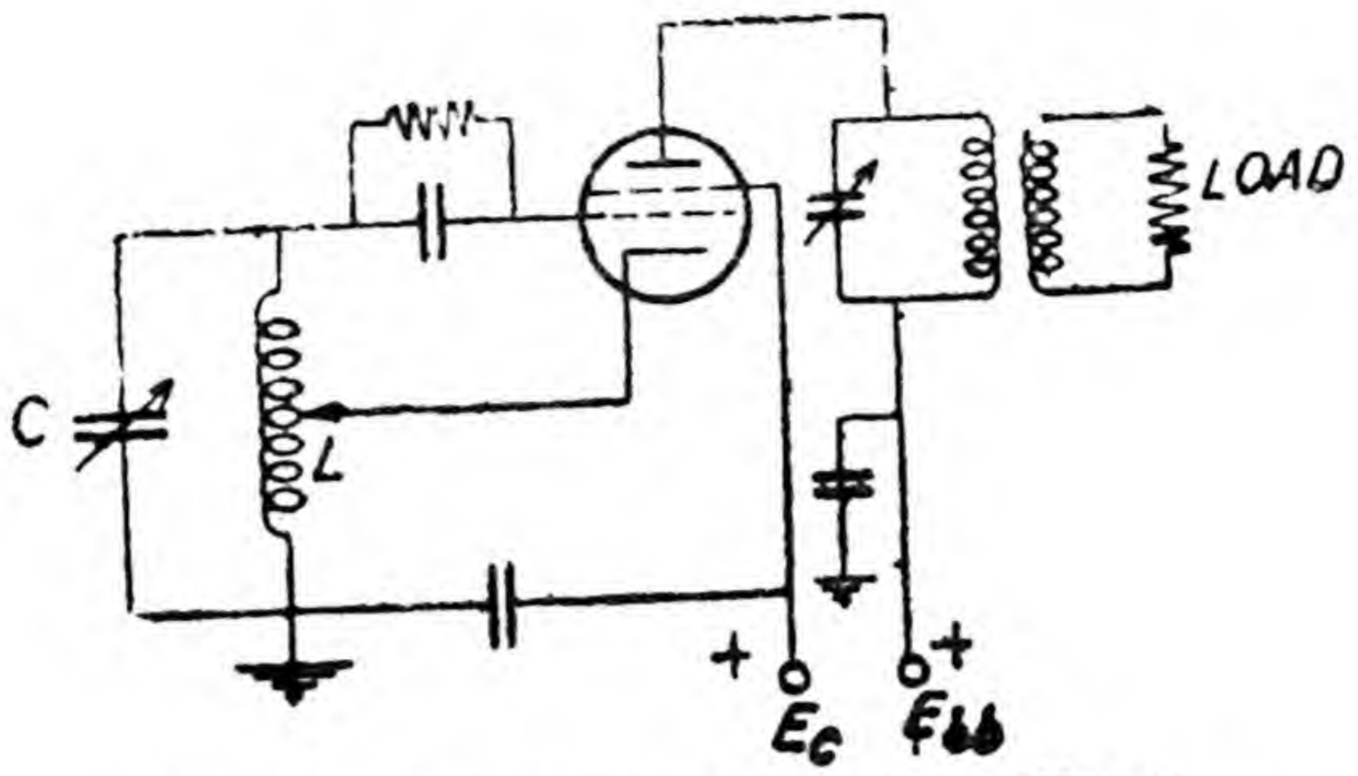


Fig. 17. Electron coupled oscillator.

the screen and reach the plate and set up a voltoge across the tuned circuit. As in case of a tetrode or pentode, the plate current is unaffected by plate voltage, the load circuit changes have negligible effect on the frequency determining grid circuit. In this way the improved frequency stability is obtained.

15·11. CRYSTAL OSCILLATORS :

Quartz crystal and its cuts : In its natural form a quartz crystal has a hexagonal cross-section and pointed ends as shown in the figure 18. The axis joining the two pointed ends of the crystal is called the *z*-axis.

Figure 19 shows a section of the crystal at right angles to the *z*-axis. With reference to the figure 19, the three axes *XX*, *X'X'* and *X"X"* passing through the corners of the hexagon are known as the *X*-axes or 'Electrical axes'. The three axes *YY*, *Y'Y'* and *Y"Y"* which are perpendicular to the faces of the crystal are called the *Y*-axis or 'mechanical axes.

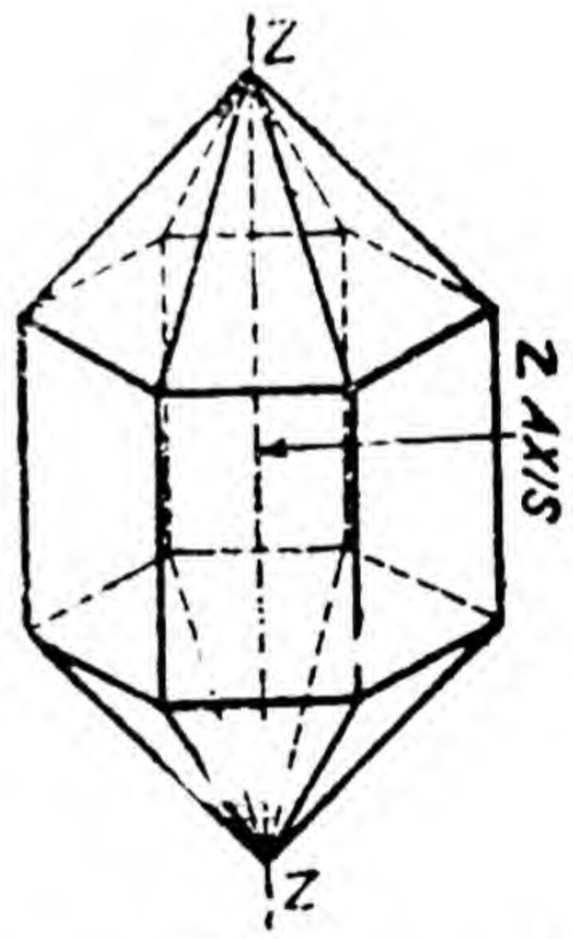


Fig. 18. Quartz crystal in its natural form.

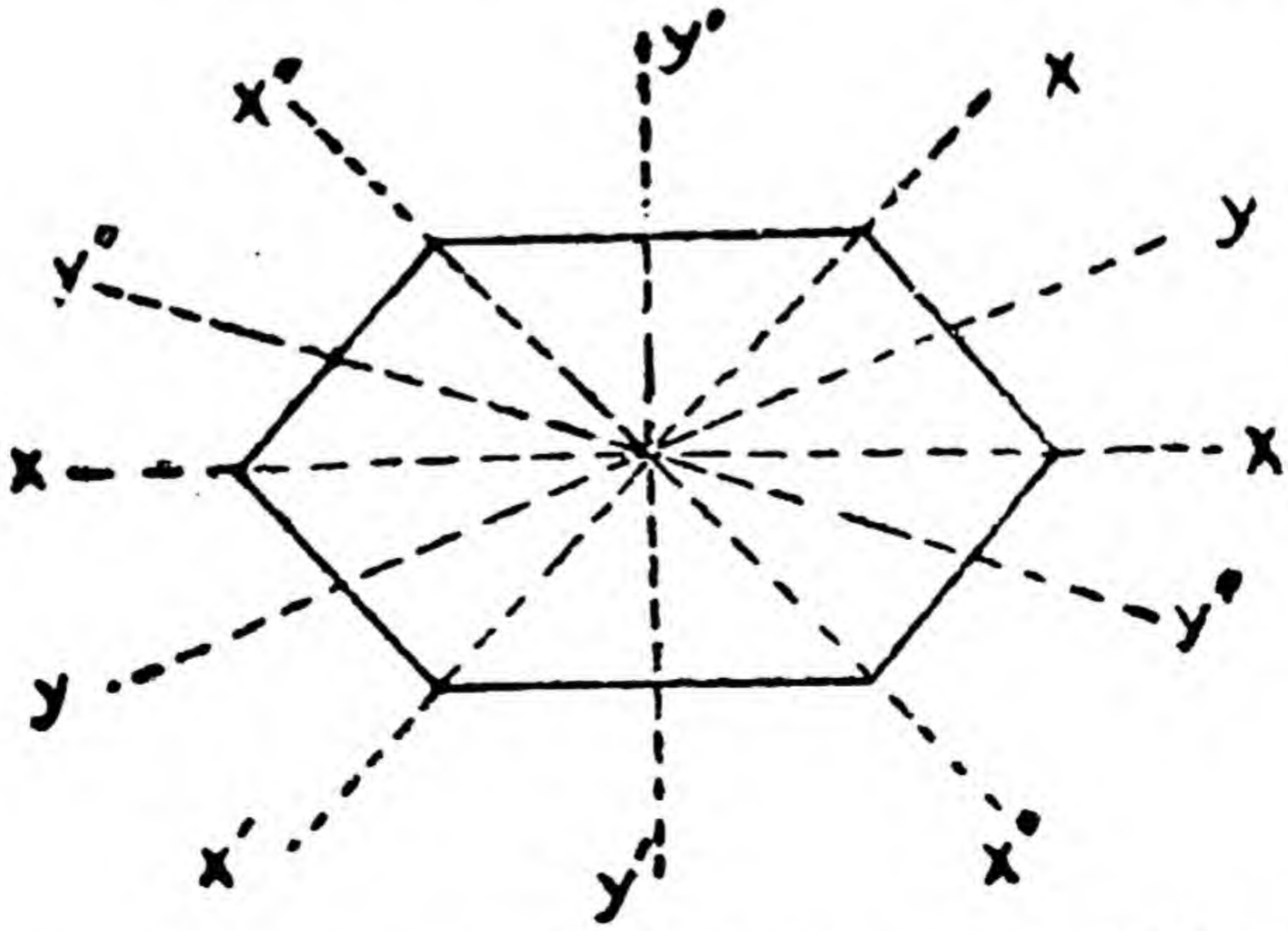


Fig. 19. Section of quartz crystal perpendicular to Z-axis.

The common crystal are (*i*) *X* cut or Curie cut, (*ii*) *Y*-cut or 30° cut and (*iii*) *AT*-cut.

The *X* cut or Curie cut crystal is taken out of a face perpendicular to *X*-axis as shown in the fig. 20.

The *Y*-cut or 60° cut crystal is taken out of a face perpendicular to *Y*-axis as shown in the fig. 21.

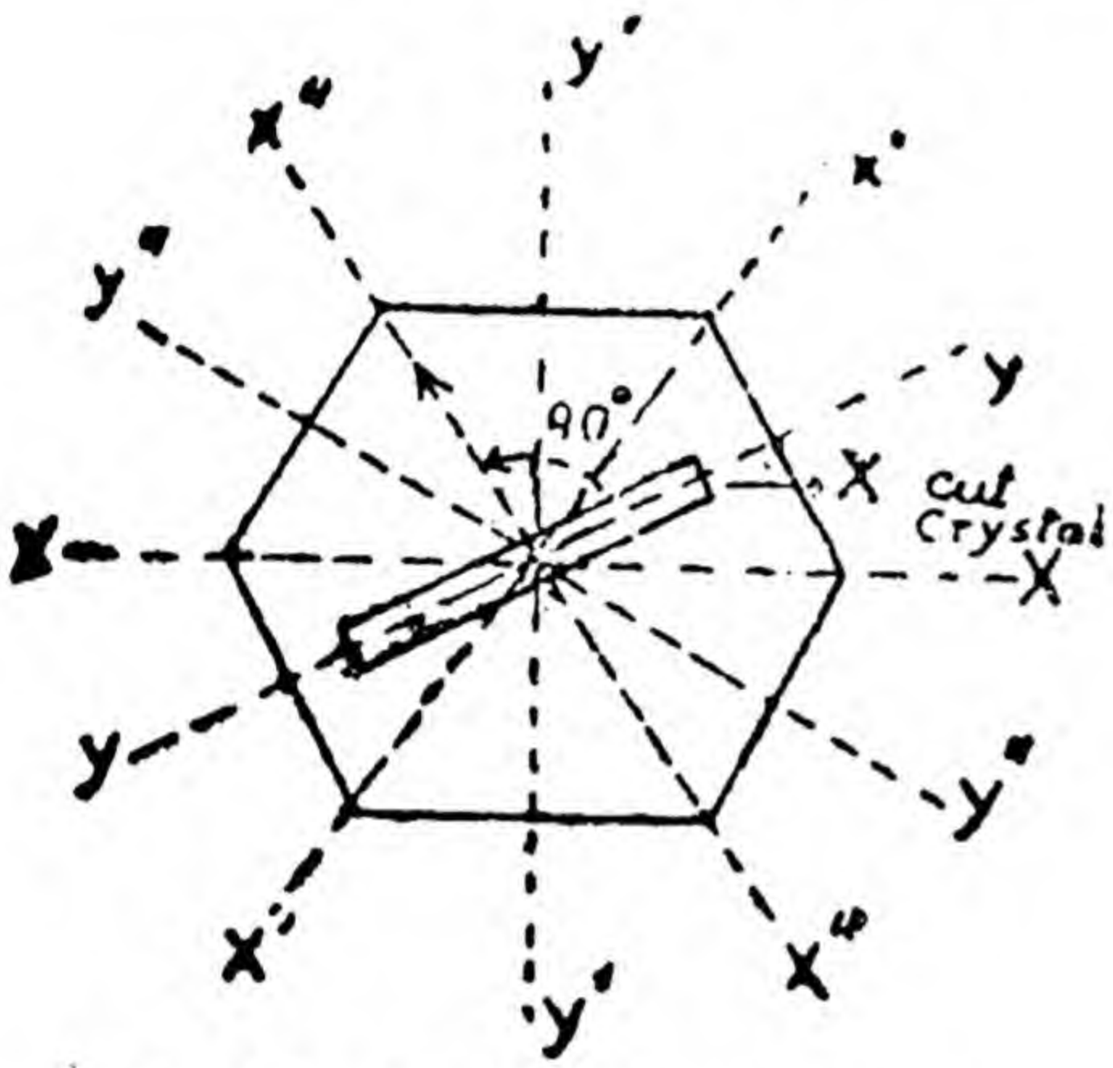


Fig. 20. Orientation of *X*-cut crystal

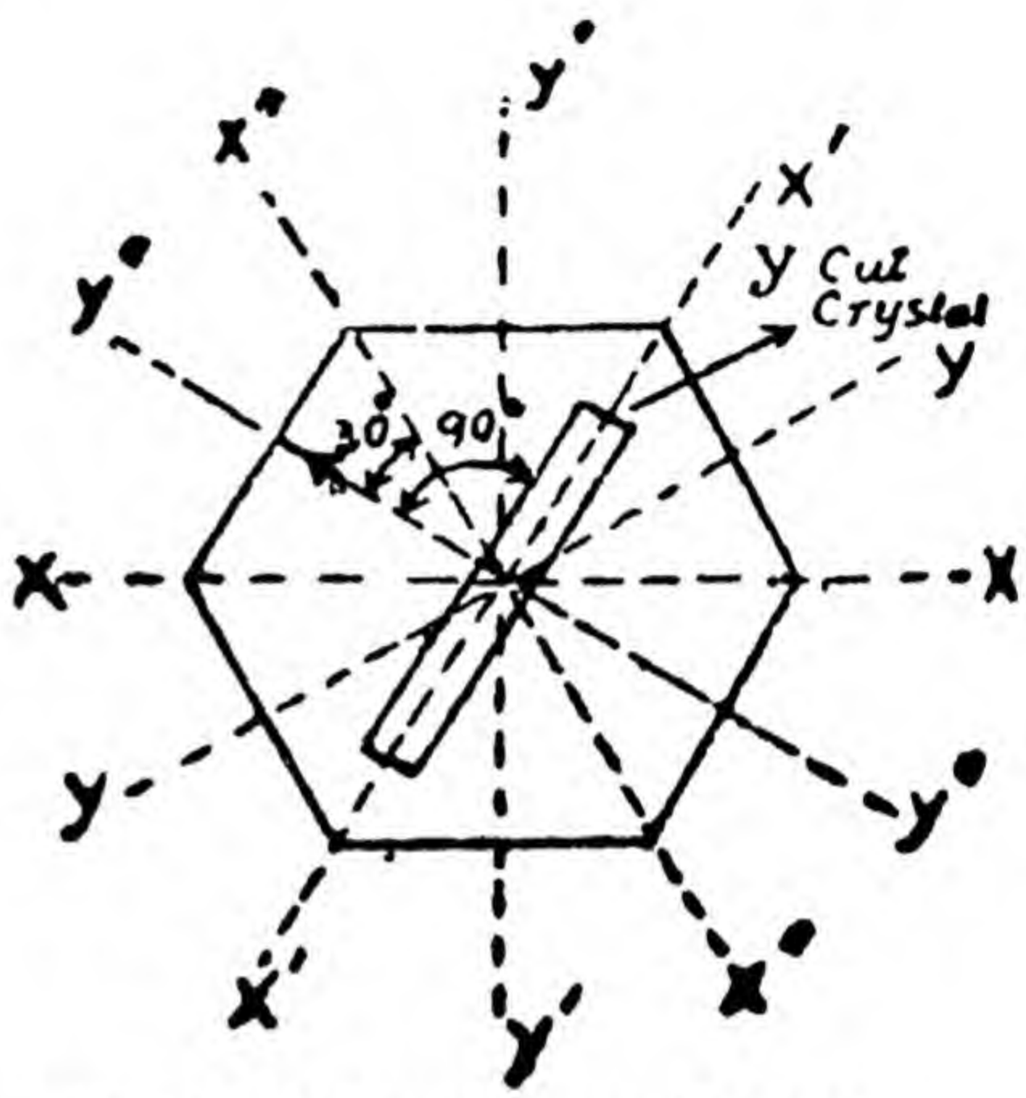


Fig. 21. Orientation of *Y*-cut crystal.

X and *Y* cuts are not used these days because of their large temperaturs coefficients. Also these plates often exhibit discontinuous frequency temperature characteristics. Modern theory has shown that it is possible to have a suitable cut of the crystal for which the temperature coefficient is zero or it is a very small quantity. The cuts which are suitable for generation of frequencies in the frequency range of 50 kc/s to about 10 mc/s. are *AT* and *BT* cuts. *AT* cut crystal plate is cut from a plate rotated about *X*-axis so that the angle θ made with *Z*-axis is approximately 35·5 degrees as shown in fig. 22.

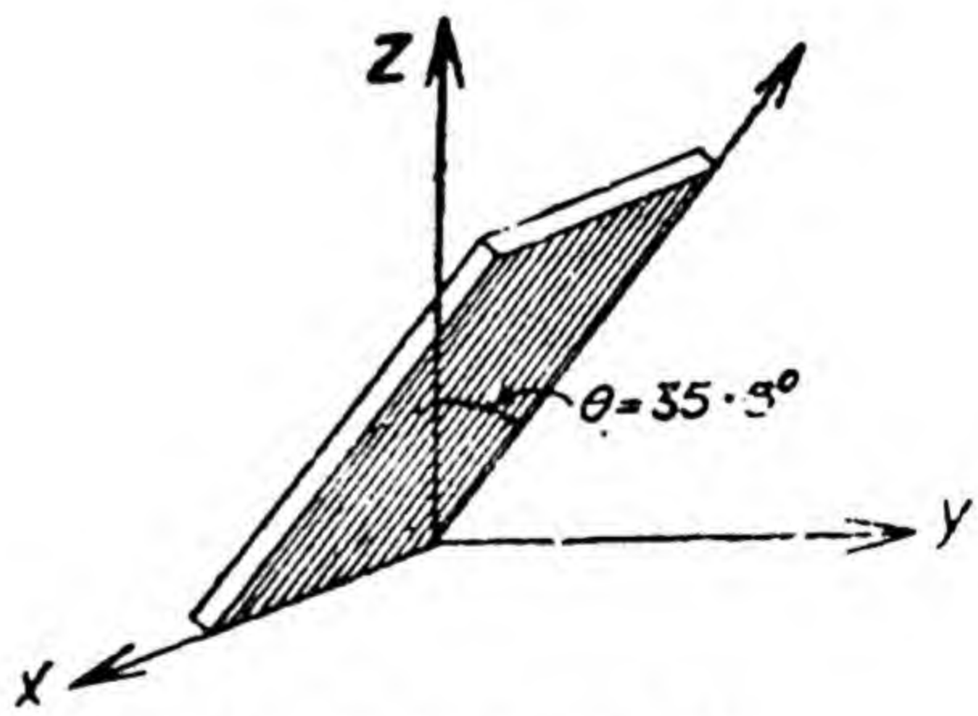


Fig. 22. Orientation of *AT*-cut crystal.

When a crystal, such as quartz crystal, is subjected to mechanical stress along *Y*-axis, electric charges will appear on the faces of the crystal perpendicular to the *X*-axis which is at right angles to the *Y*-axis connected. Conversely, if the electrical changes are placed on flat crystal surface by applying a voltage across these faces, mechanical stress is produced on other faces. This property by which electrical and mechanical properties are interconnected

in a crystal is known as *Piezo electric effect* and has been utilised in crystal oscillator.

Crystal oscillator circuit : In order to make electrical connections to the crystal, it is usually mounted horizontally between two metal plates. As the crystal must vibrate to produce oscillations, it must be mounted between these two plates in such a manner as to allow for the required amount of mechanical vibration. A vibrating crystal can be replaced by an equivalent electrical circuit as shown in the fig. 23 (*b*).

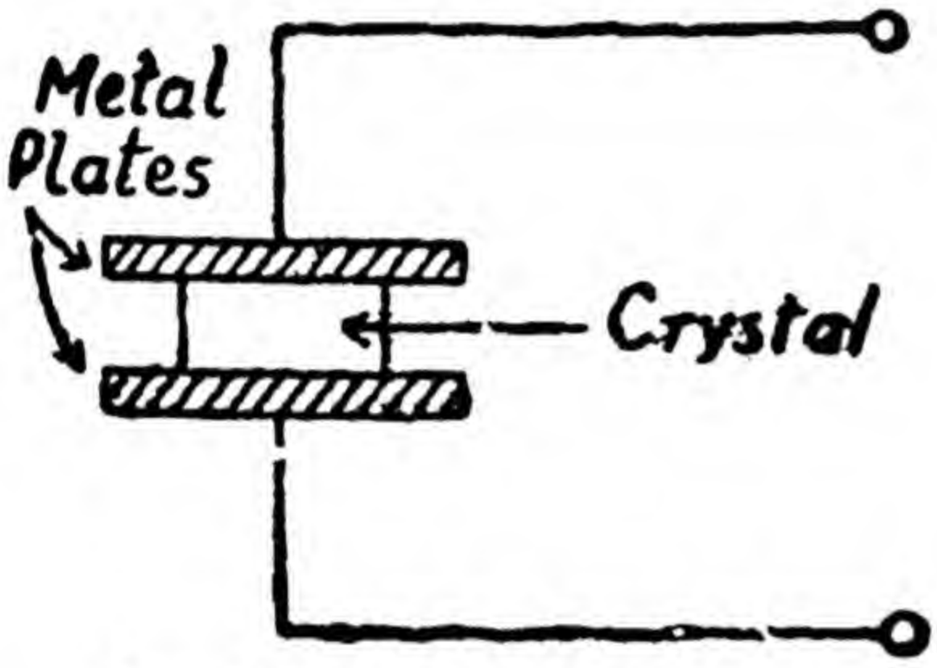


Fig. 23. (*a*) Crystal with mounting electrodes.

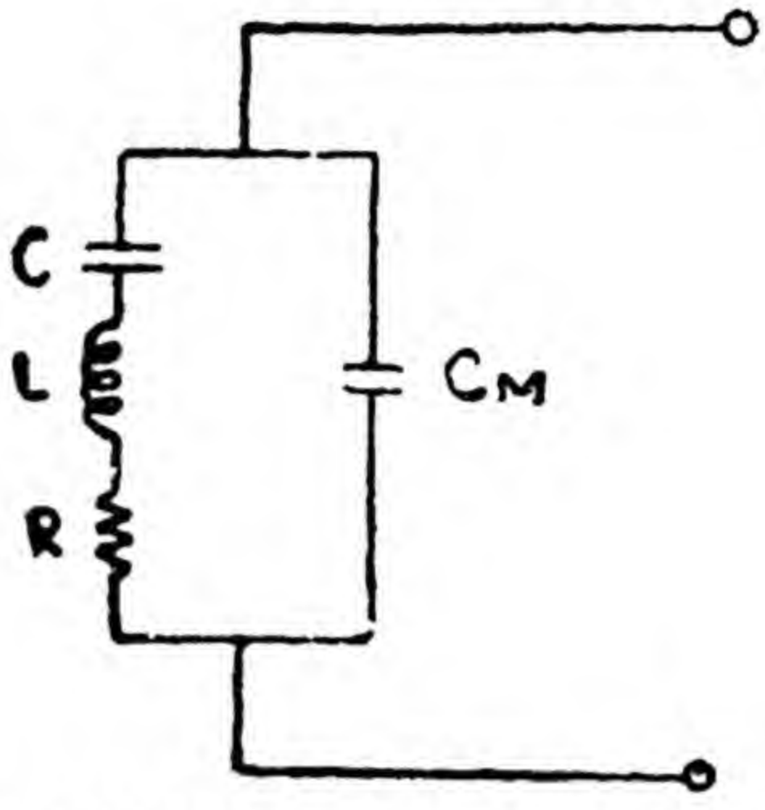


Fig. 23. (*b*) Equivalent circuit

The inductance L represents the mass of the crystal, the capacitor C the resistance or elasticity, resistance R the frictional losses, and the capacitor C_M is the actual capacitance formed by the crystal electrode, with crystal itself as the dielectric.

The inductance of quartz crystal is very high depending upon the cut and dimensions of the unit. On the other hand, the resistance is relatively low. This results in extremely high Q factors. The capacitance value C and C_M are low ; therefore at low frequencies, the reactance of this network is high. At some definite frequency, the reactances of L and C will be numerically equal to each other. This is the series resonant frequency and corresponds to the natural frequency of crystal. At some frequency higher than the resonant frequency, the combined effective reactance L and C is inductive and will be numerically equal to the reactance of capacitor C_M. At this frequency the crystal circuit acts as a parallel resonant circuit and its impedance is maximum. The circulating current in the circuit $CLRC_M$ is maximum and therefore the crystal vibrations will be maximum.

In a crystal oscillator. quartz crystal is put instead of resonant circuit which determines the frequency of oscillation. Fig. 24 shows a circuit of the crystal oscillator.

A study of equivalent circuit shows that the circuit of the crystal oscillator is equivalent to tuned grid-tuned plate oscillator. The feed back takes place through the internal plate grid interelectrode capacitance C_{gp} of the tube. From the study of Miller effect, we know that the plate load affects the grid cathode impedance and if the plate load is inductive, the input impedance of the

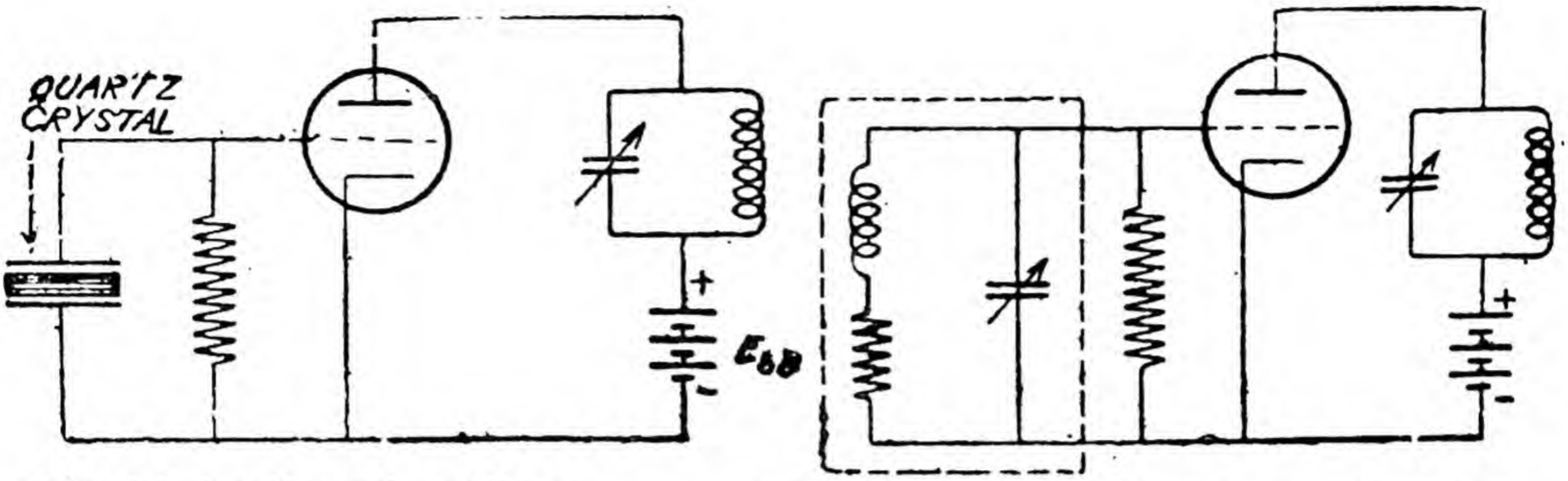


Fig. 24. (*a*) Crystal oscillator. **Fig. 24.** (b) Equivalent electrical circuit

valve may have a negative resistance The required inductive load is obtained with the help of *LC* circuit, tuned to a frequency slightly greater than the mechanical resonant frequency of the crystal. When the value of plate tuning capacitance is increased in such a way that the negative resistance component of the input impedance acquires a sufficient value, the energy fed back to grid circuit increases and circuit oscillates.

It is advantageous to use a pentode instead of a triode because in case of pentode a low grid voltage singal is required to drive it and due to this fact, the voltage across the crystal will be small so that very small heating of crystal takes place. As the value of C_g is small in case of pentode, it is necessary to place a small capacitance between the plate and grid. This capacity provides sufficient feed back so that the sustained oscillations may take place.

15·12. RELAXATION OSCILLATORS :

Multivibrators : Multivibrator is a device which produces extended voltage waveforms of almost square shape and also voltage pulses occuring periodically. According to one method of classification multivibrators may be classified as :

(*a*) *Free running Multivibrator* : This may generate pulses and extended waveforms independently without the necessity of any driving or external synchronizing voltage pulse.

(*b*) *Single shot Multivibrator* : This requires one driving pulse for generation of each cycle of wav form.

(*c*) *Bistable Multivibrator* : This required two driving pulses one for each half cycle of the waveform.

Free Running (Astable) multivibrator :

Essentially the circuit of multivibrator consists of a simple two stage *RC* coupled amplifier, without cathode bias, but with the output from V_1, fed-back as the input to V_1 since each tube (V_1 and V_2) produces a 180 deg. phase shift, any output from V_2

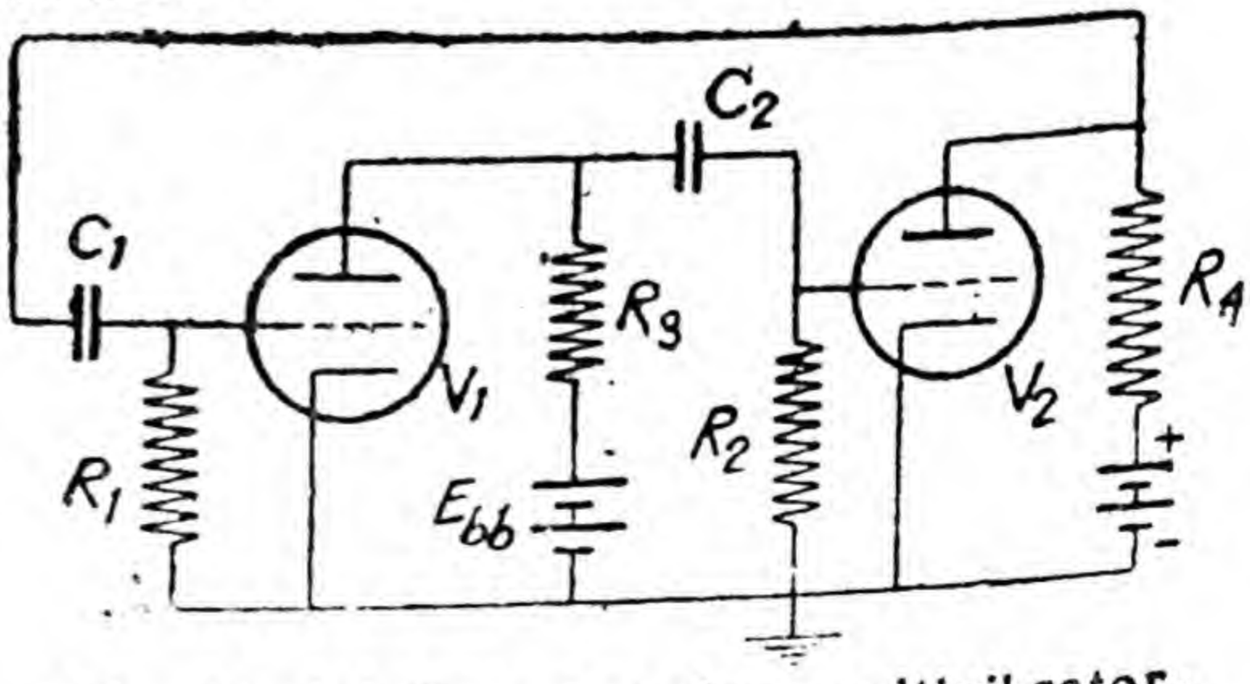


Fig. 25. (a) Free running multivibrator.

is in phase with the input to V_1.

In other words, we have positive or regenerative feed back. Since the coupling between the two tubes is *via* the plate of the other tube, this type of free running multivibrator is also known as a plate coupled multivibrator, whose circuit is given in the fig. 25(*a*), (*b*).

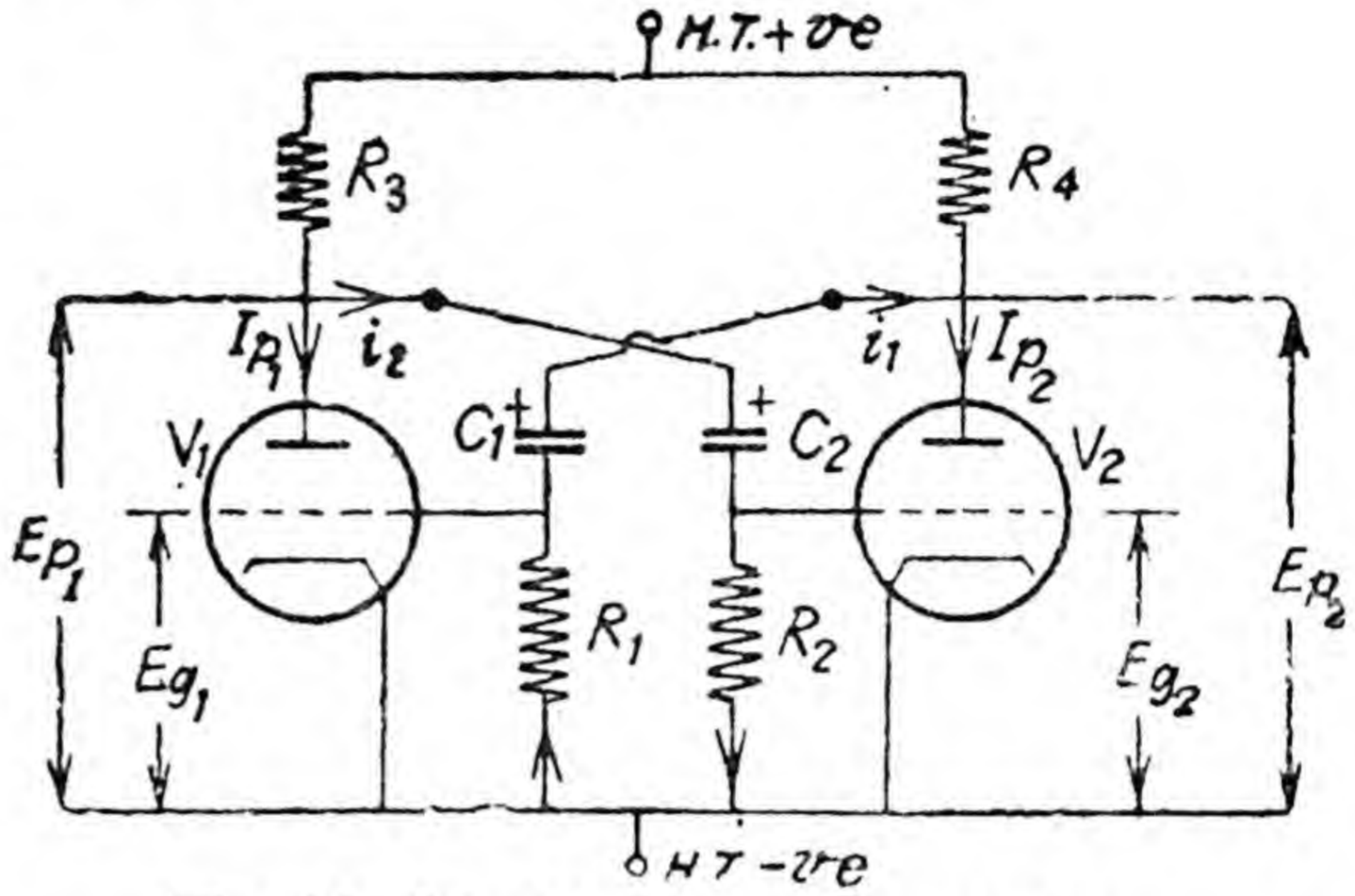


Fig 25. (b) Rearanged form of free running multivibrator.

Operation: Initially let us suppose that there is no charge on condensers and grid to the tubes are at zero potential. When plate supply voltage is switched on, the plate current flows in both the tubes and also the voltages across the condensers increase. If the two halves of the circuit are perfectly symmetrical, equal plate currents will flow However, this is not always possible and there is a slight difference which causes a cumulative unbalance.

Consider that the grid potential of valve V_1 becomes slightly negative, the plate current immediately falls and a positive signal is applied to the grid of valve V_2. As a result, the current in V_2 increases and the output signal is decreased. As the signal is coupled to the grid of V_1, the grid becomes more negative and plate current falls still further. This effect is cumulative—the potential of valve V_1 is driven so far negatively, almost instantaneously, that the current in V_1 is cut off. Simultaneously, V_2 is driven into maximum condition. With V_1 still cut-off the charge on capacitor C_1 leaks through R_1 and at some point the potential of the grid of valve V_1 becomes such that the tube starts conduction. Due to conduction of valve V_1, a negative potential is applied to the grid of valve V_2. A cumulative chain occurs and the valve V_2 is cut-off. Thus there are two unstable limiting conduction, one in which valve V_1 is cut-off and the second in which valve V_2 is cut off.

Detailed circuit operation : Let us now suppose that the grid potential of the first valve Eg_1, becomes slightly negative. The anode current I_{p_1} immediately falls, and the anode potential Ep_1 rises. A charging current i_2 then flows through the condenser C_2 and the resistor R_2. This current sets up a potential difference i_2R_2 across the resistor, which makes the grid of the second valve positive. The corresponding increase of anode current lowers the anode potential of the second valve, Ep_2 ; capacitor C_1, connected between plate of V_2 and ground in series with R_1, cannot hold its previous charge It discharges as shown in figure 26 (*a*). The timing being primarily determined by the time constant R_1C_1, Since

current flows down through R_1, the grid of V_1 is driven negative. The plate current of V_2 will decrease and the plate potential will increase.

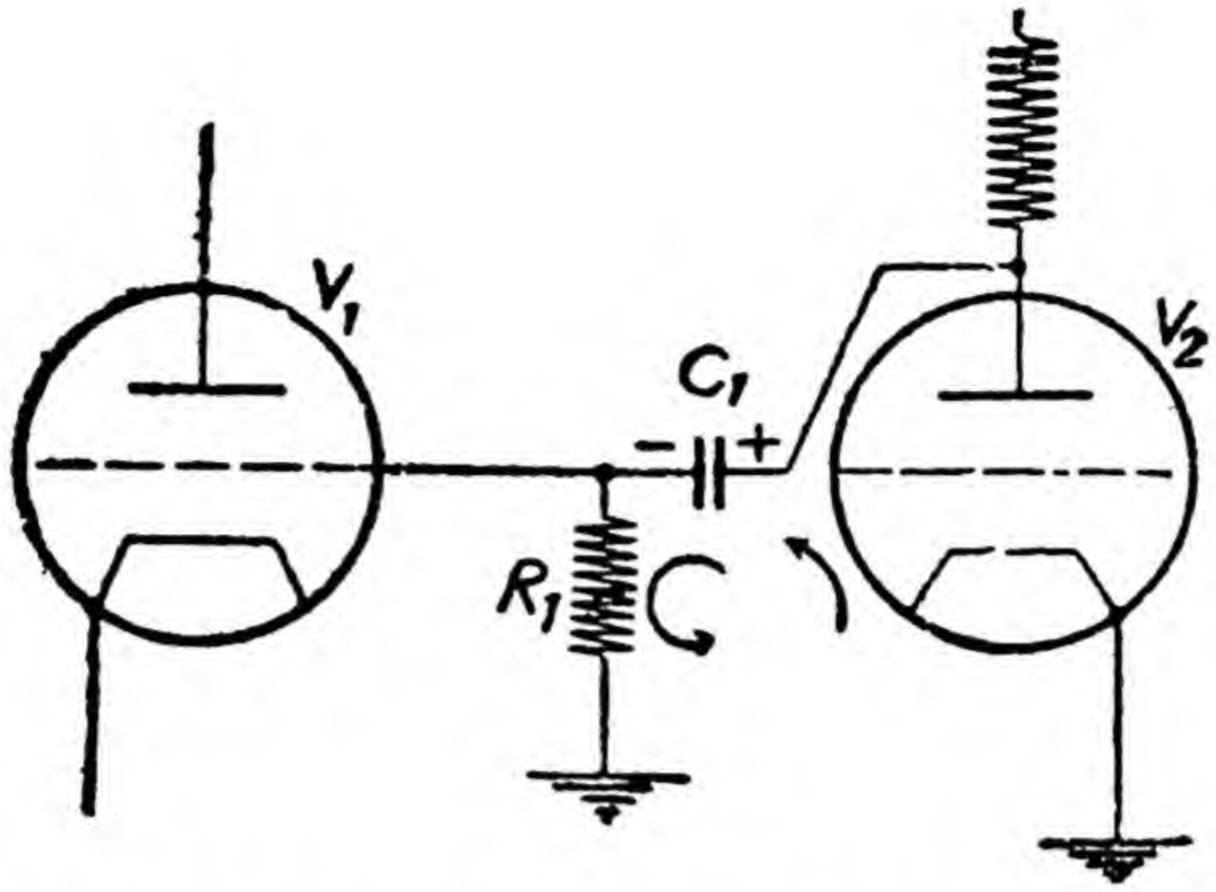


Fig. 26 (*a*) Discharge of C_1

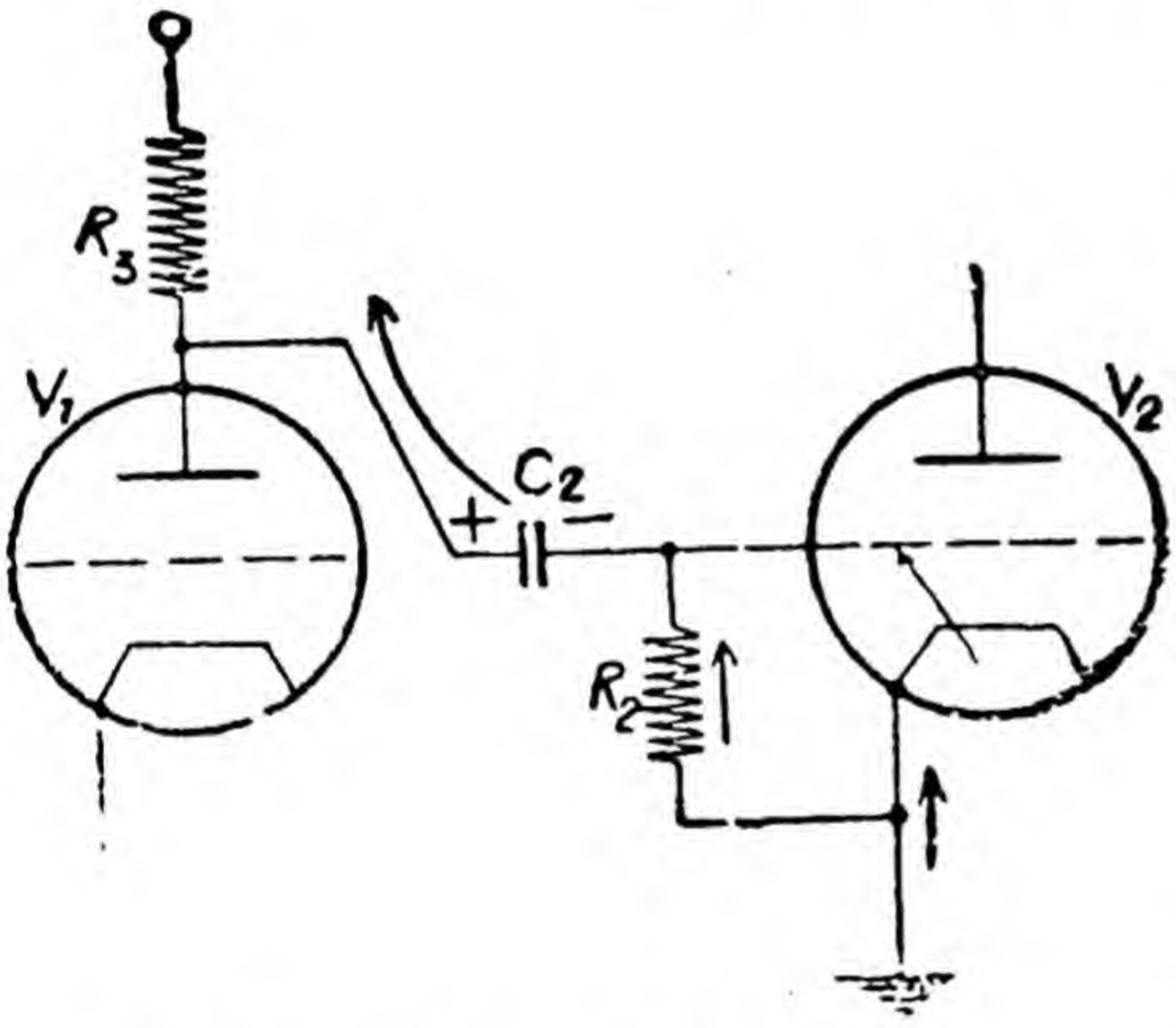


Fig 26 (*b*) Charging of C_2

As capacitor C_2 is connected between the plate of V_1 and ground (in series with resistor R_2, the charge on capacitor C_2 will increase. The charging path is shown in the fig. 26 (*b*).

Since the grid of V_2 is driven positive, its plate current increases. But this was the starting premise. In other words, the action is commulative (positive feedback) and almost immediately V_1 will be driven deep into cut-off, its plate potential rising to full E_{bb} value. Simultaneously, V_2 is driven into maximum conduction and its plate potential falls to its minimum value. At this point no further changes can take place and the plate potentials remain constant for a while at these values.

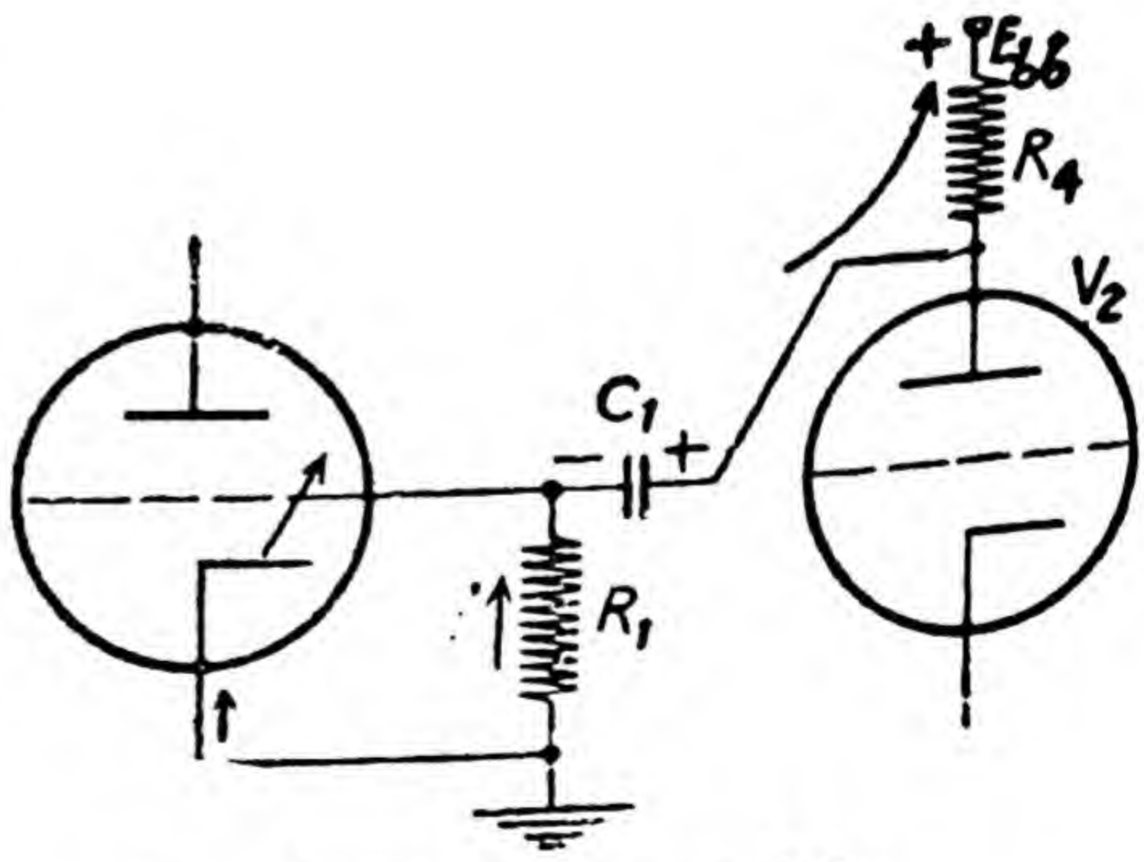


Fig. 26. (*c*) Charging of C_1.

With time, C_1 discharges sufficiently (and the discharge current falls to a value low enough) as to bring V_1 just out of cut-off. Plate current Ip_1 begins to flow, plate potential E_{p1} begins to

drop and capacitor C_2 starts to discharge, applying a negative voltage to the grid of V_2. Plate current I_{p_2} will begin to decrease, plate potential E_{p_2} rises, capacitor C_1 begins to charge as shown in the figure 26 (*c*). This drives the grid of V_1 positive.

This time the grid potential of the first valve is rising,

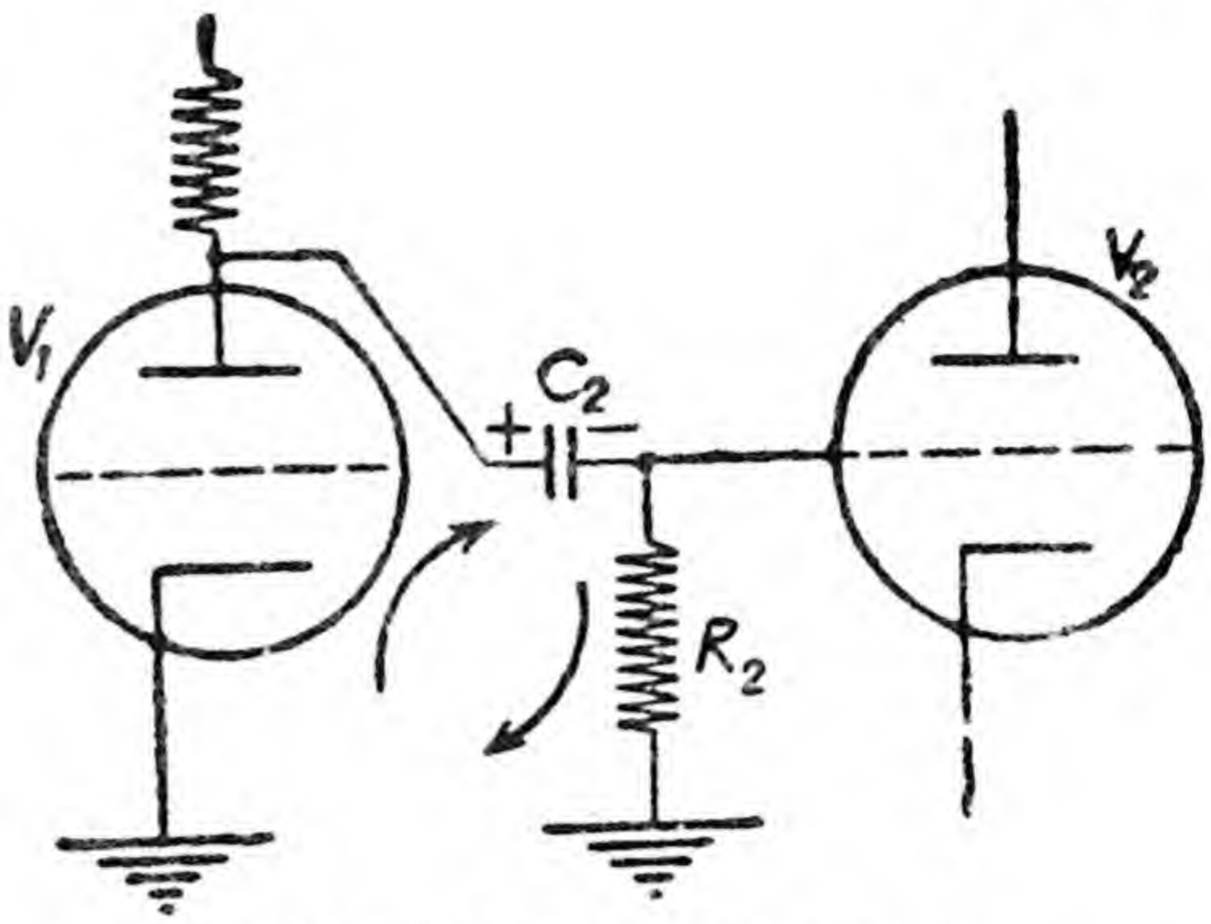


Fig. 26 (*d*) Discharging of C_2

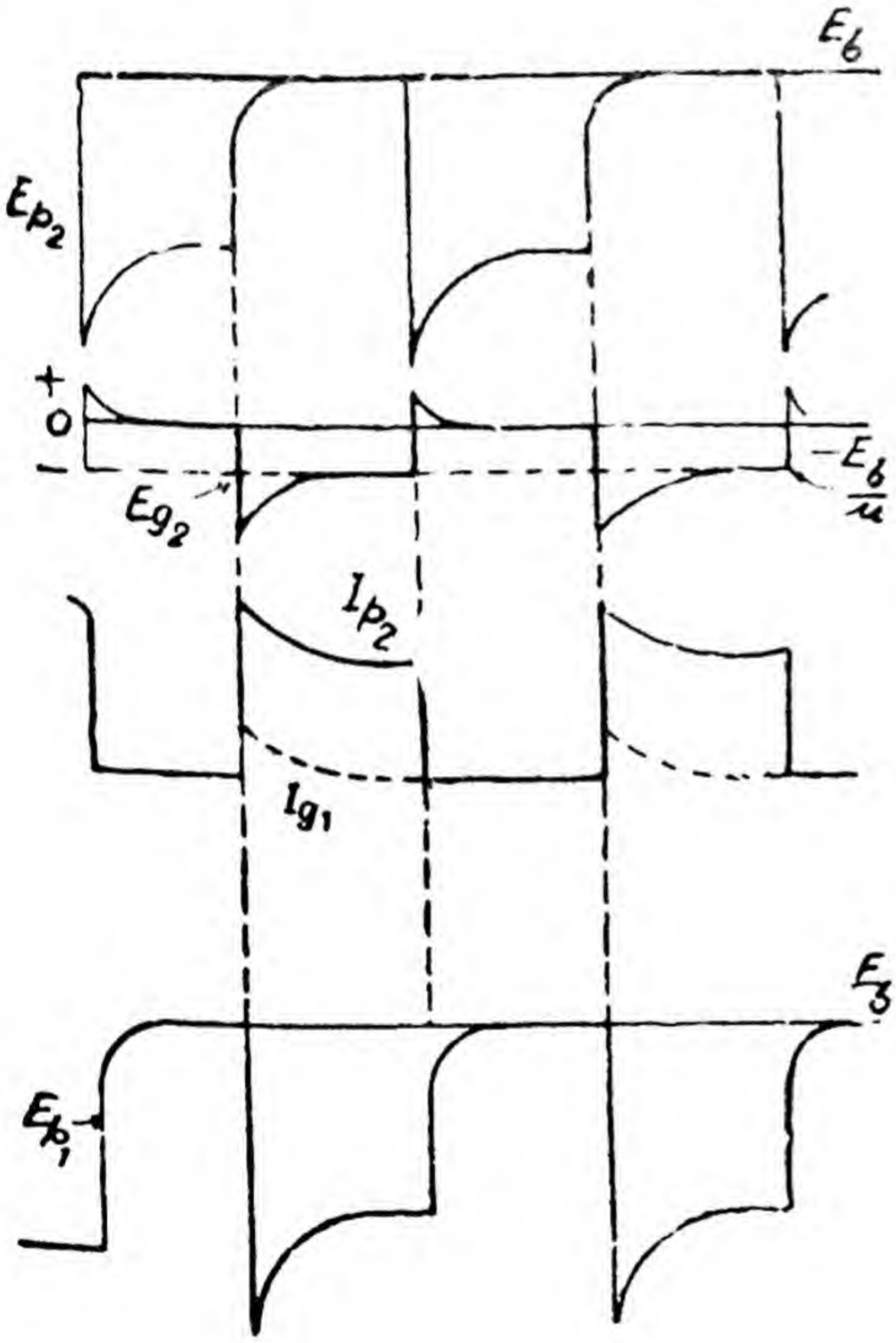


Fig. 26 (*e*) Waveforms.

and the avalanche takes place in opposite direction until the second valve is biased beyond cut off. Exponential decay of the grid potentials follows, and the cycle repeats itself indefinitely.

Potential waveforms : Potential waveforms of free running multivibrator are shown in fig. 26 (*e*).

The frequency of oscillations : If T_1 is the time taken by the first valve to recover after being cut off and similarly T_2 for the second valve (which depends upon the time constant associated with its grid condenser), then approximately,

$$T_1 \simeq R_1C_1$$

and $$T_2 \simeq R_2C_2,$$

The total time for full cycle

$$T = T_1 + T_2 \simeq R_1C_1 + R_2C_2.$$

The frequency of the output wave shape must be reciprocal of the period.

$$f \simeq \frac{1}{R_1C_1 + R_2C_2}.$$

When capacitances C_1 and C_2, and the resistances R_1 and R_2 are equal, the multivibrator is called symmetrical and gives the waveforms of the figure 26 (*e*). By making the time constant R_1C_1 different from R_2C_2, the relative lengths of the two parts of the cycle can be varied to give output waveforms ranging from square waves to sharp pulses.

To obtain a more general expression for the frequency of multivibrator, we proceed as follows :

Let E_{bb} = Potential of supply source,

E_{p_1} = Plate potential across valve V_1 when it is conducting,

E_{p_2} = Plate potential across valve V_2 when it is conducting,

E_{01} = Cut off value of grid potential of valve V_1,

E_{02} = Cut off value of grid potential of valve V_2,

T_1 = Time during which the valve V_1 remains non-conducting,

T_2 = Time during which the valve V_2 remains non-conducting.

When first valve is cut off, the time T_1 for which it is non-conducting depend upon the product R_1C_1 and similarly, the time T_2 for which valve V_2 is non-conducting depends upon R_2C_2. The total time T is the sum of these two time periods. The time T_1 for which valve V_1 is non-conducting is the time required for E_{g_1} to rise exponentially from $-(E_{bb}-E_{p_2})$ to cut off level $-E_{01}$. Therefore

$$-E_{01}=-(E_{bb}-E_{p_2})\, e^{-T_1/R_1C_1}$$

$$\therefore \quad T_1=R_1C_1 \log_e \frac{(E_{bb}-E_{p_2})}{E_{01}} \qquad \ldots(1)$$

Similarly, from the equation to T_2, we have

$$-E_{02}=(E_{bb}-E_{p_1})\, e^{-T_2/R_2C_2}$$

$$\therefore \quad T_2=R_2C_2 \log_e \frac{(E_{bb}-E_{p_1})}{E_{02}} \qquad \ldots(2)$$

Total time period T is given by

$$T=T_1+T_2=R_1C_1 \log \frac{(E_{bb}-E_{p_2})}{E_{01}}+R_2C_2 \log_e \frac{(E_{bb}-E_{p_1})}{E_{02}}.$$

If the two halves are identical, *i.e.*,

$$E_{p_1}=E_{p_2}=E_p,$$

$$E_{01}=E_{02}=E_0,$$

and $R_1=R_2=R$, $C_1=C_2=C$, we have

$$T=2RC \log_e \frac{(E_{bb}-E_p)}{E_0}. \qquad \ldots(3)$$

The frequency f is given by $= \frac{1}{T}$.

$$\therefore \quad f = \frac{1}{2RC \log_e \frac{(E_{bb} - E_p)}{E_0}} \qquad \ldots(4)$$

The expression is valid only for low frequency reception as it does not include stray and tube capacities.

The choice of plate load resistance R_L is important, as its proper choice gives optimum frequency.

Stability. To deduce the value, consider the fractional change in period of system. From equation (3) we have

$$dT = \frac{RC\, d(E_{bb} - E_p)}{(E_{bb} - E_p)}$$

and

$$\frac{dT}{T} = \frac{d\,(E_{bb} - E_p)}{(E_{bb} - E_p) \log_e \frac{(E_{bb} - E_p)}{E_b}} \qquad \ldots(5)$$

This expression will be minimum when the denominator is maximnm ; when this is done, we have

$$E_{bb} - E_p = 2{\cdot}71\ E_0.$$

Thus for optimum frequency stability

$$R_L \simeq \frac{2{\cdot}71\ E_0}{I_p}$$

where I_p is the plate current when valve is conducting.

15·13. ELECTRON MOTION IN TIME VARYING FIELD :

In tubes operating at microwave frequencies, the time of transit of an electron between two electrodes in the tube may be relatively large in comparison with the period of electric oscillations. It is interesting to study the motion of electrons through the field when a large transit time is involved.

For an example, let us investigate the motion of electrons in the time varying field of a parallel plate diode in which the current is temperature limited. In the diode shown in the figure 27, the instantaneous potential difference between cathode and plate is assumed to have a direct component V_0 and an alternating component $V_1 \sin \omega t$. Then

$$V = V_0 + V_1 \sin \omega t. \qquad \ldots(1)$$

Since the space charge density is negligible, the electric intensity in the diode space is given by

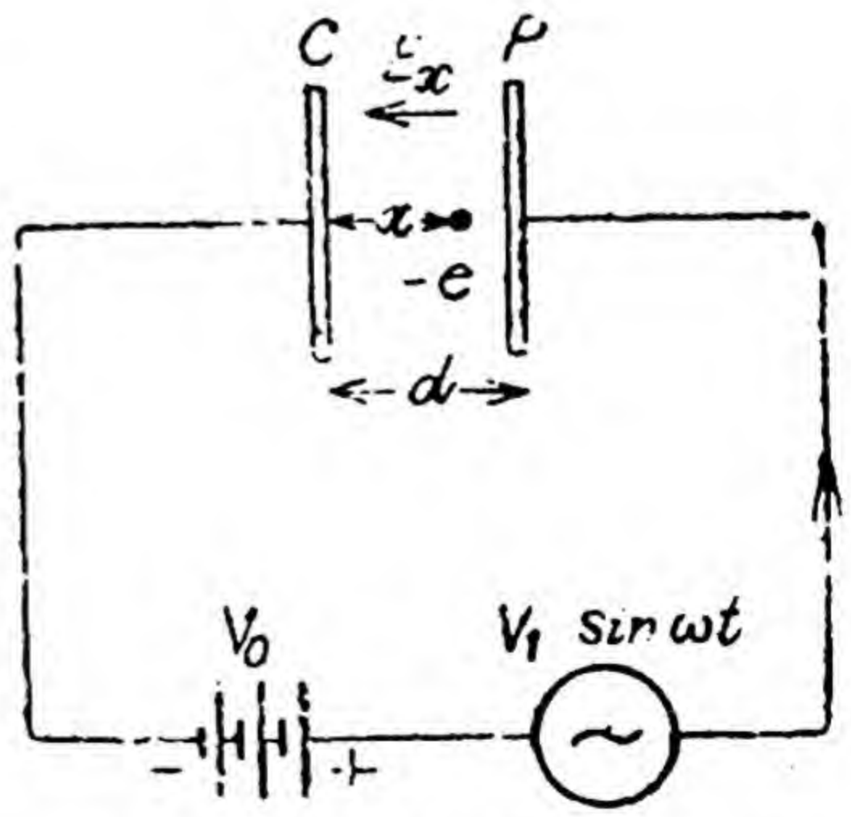


Fig 27. Parallel plate diode with d. c. and a. c. applied potentials.

$$E_x = -\frac{V}{d} = -\frac{V_0 + V_1 \sin \omega t}{d} \qquad \ldots(2)$$

Hence, force experienced by the electrons

$$-eE_x = \frac{e}{d}(V_0 + V_1 \sin \omega t). \qquad \ldots(3)$$

If m is the mass of the electron and $\frac{d^2x}{dt^2}$ is its acceleration, then the expression for force is equal to $m \frac{d^2x}{dt^2}$; so that

$$m \frac{d^2x}{dt^2} = \frac{e}{d}(V_0 + V_1 \sin \omega t). \qquad \ldots(4)$$

or
$$\frac{dx}{dt} = \frac{e}{md}\left[V_0 t - \frac{V_1}{\omega} \cos \omega t \right] + C_1. \qquad \ldots(5)$$

where C_1 is the integration constant.

Suppose the electron leaves the cathode with velocity v_0 at the instant t_0,

$$v_0 = \frac{e}{md}\left[V_0 t_0 - \frac{V_1}{\omega} \cos \omega t_0 \right] + C_1,$$

$$C_1 = v_0 - \left[V_0 t_0 - \frac{V_1}{\omega} \cos \omega t_0 \right] \cdot \frac{e}{md}.$$

Putting this value in equation (2), we get

$$\frac{dx}{dt} = \frac{e}{md}\left[V_0 t - \frac{V_1}{\omega} \cos \omega t \right] + v_0 - \frac{e}{md}\left[V_0 t_0 - \frac{V_1}{\omega} \cos \omega t_0 \right]$$

$$= \frac{e}{md}\left[V_0 t - \frac{V_1}{\omega} \cos \omega t - V_0 t_0 + \frac{V_1}{\omega_1} \cos \omega t_0 \right] + v_0$$

$$= \frac{e}{md}\left[V_0 (t - t_0) - \frac{V_1}{\omega} (\cos \omega t - \cos \omega t_0) \right] + v_0 \qquad \ldots(6)$$

Again integrating with respect to t, we get

$$x = \frac{e}{md}\left[V_0 \frac{(t - t_0)^2}{2} - \frac{V_1}{\omega}\left(\frac{1}{\omega} \sin \omega t - t \cos \omega t_0 \right) \right] + v_0 t + C_2 \qquad \ldots(7)$$

At the instant $x = 0$, $t = t_0$,

$$0 = \frac{e}{md}\left[0 - \frac{V_0}{\omega}\left(\frac{1}{\omega} \sin \omega t_0 - t_0 \cos \omega t_0 \right) \right] + v_0 t_0 + C_2,$$

$$\therefore \quad C_2 = \frac{e}{md}\left[\frac{V}{\omega}\left(\frac{1}{\omega} \sin \omega t_0 - t_0 \cos \omega t_0 \right) \right] - v_0 t_0,$$

Putting in equation (7),

$$x = \frac{e}{md}\left[V_0 \frac{(t - t_0)^2}{2} - \frac{V_1}{\omega}\left(\frac{1}{\omega} \sin \omega t - t \cos \omega t_0 \right) \right] + v_0 t$$

$$+ \frac{e}{md}\left[\frac{V_1}{\omega}\left(\frac{1}{\omega} \sin \omega t_0 - t_0 \cos \omega t_0 \right) \right] + v_0 t_0$$

$$=\frac{e}{md}\left[V_0\frac{(t-t_0)^2}{2}-\frac{V_1}{\omega}\left(\frac{1}{\omega}\sin\omega t-t\cos\omega t_0-\frac{1}{\omega}\sin\omega t_0\right.\right.$$

$$\left.\left.+t_0\cos\omega t_0\right)\right]+v_0(t-t_0).$$

Putting $t-t_0=T$, the transit time of the electron to move from the cathode to a distance x and let $\omega t_0=\phi$, phase angle of alternating potentia! : then

$$\frac{x}{K}=A+\frac{V_1}{V_0}B+C,$$

where

$$K=\frac{e}{md}\cdot\frac{V_0}{\omega^2}$$

$$A=\frac{(\omega T)^2}{2},$$

$$C=\frac{v_0.T}{K}$$

$$B=\cos\phi\,(\omega t-\sin\omega t)+\sin\phi\,(1-\cos\omega t).$$

15·14. ULTRA-HIGH FREQUENCY EFFECTS IN CONVENTIONAL TUBES :

It is an experimentally observed fact that conventional tubes do not operate favourably and efficiently at high frequencies. In case of amplifier gain decreases; input impedance and the maximum impedance that can be realised in the plate circuit also decrease. When valves operate as oscillators, even output decreases ; they become less efficient and output obtained is limited by both plate dissipation and grid dissipation at ultra-high frequencies. The above effects in conventional tubes enter due to the following detrimental limitations :

2. Circuit reactance limitations.
2. Circuit and tube-loss limitations.
3. Electron transit time limitation.

15·14-1. CIRCUIT-REACTANCE LIMITATIONS ·

The electrode leads have got small finite inductance and two electrodes in combination form an effective capacitor. At moderate frequency, the reactance offered by lead inductances and interelectrode capacities is small, but at the ultra-high frequency region the values of lead inductances and inter-electrode capacitances becomes large and therefore the tube offers an appreciable reactance at ultra-high frequency region. The situation is shown in fig. 28. Because of the appreciable magnitude of the lead inductances, the voltages applied across the external terminals do not appear across the electrode. In addition,

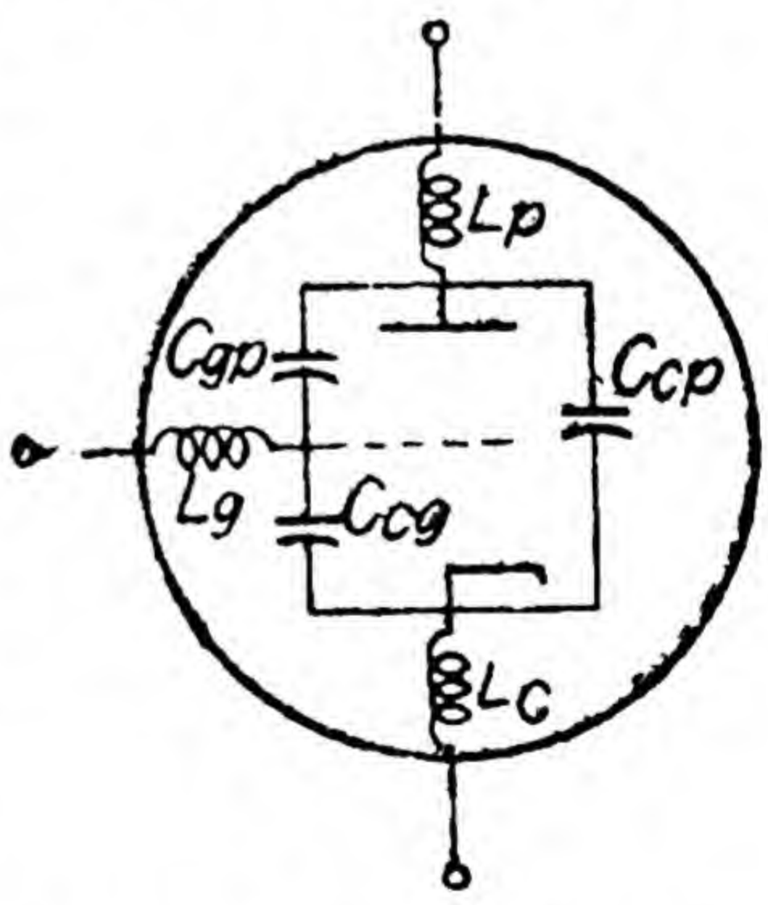


Fig. 28. Reactive elements associated with triod$_e$ tube.

cathode lead inductance introduces a feed back through the grid circuit. The effect at this feed-back is that it modifies the input admittance as given by $Y_{in}=j\omega C_{cg}\ (1-\omega L_c\ g_m)$ ...(1)
where g_m is the mutual conductance and other quantities bear the physical significance as shown in the fig. 28. The first term of input admittance is normal capacitive susceptance of the tube and second part represents the conductive component of the input admittance and has got the value $g_{in}=\omega^2 L_c C_{cg} g_m$. ...(2)

It means that in effect a resistive component is shunted across the input terminals. The magnitude of this resistance is proportional to the square of the frequency. The resistance consumes power, which increases as the square of the frequency for a given driving voltage and hence reduce the amplifying capabilities of the circuit. At suffieiently high frequency this makes the gain of amplifier less than unity and oscillator output falls to zero.

The inter electrode capacity is appreciable at high frequency and it determines the plate load resistance. The plate-load resistance fixes the gain and power output of the tube. The equivalent shunt resistance of a parallel resonant circuit is

$$R_{sh}=\frac{1}{\omega_0 \cdot RC^2} \quad ...(3)$$

or

$$R_{sh}=\frac{Q}{\omega_0 C} \quad ..(4)$$

where Q is equivalent series resistance, C is the total capacity determining the resonance and ω_0 is the resonant frequency. For operation at given resonant frequency, the shunt resistance is increased by decreasing the capacity and increasing the inductance to maintain the same resonant frequency. There is a limit below which capacity cannot be decreased. At the most we can make external capacity equal to zero and can minimise the inter-electrode capacity. So the conclusion is that larger is the inter electrode capacity, smaller will be the shunt resistance that can be realised. Hence power output decreases as the load resistance increases or as the square of the frequency when frequency increases.

The gain band-width product of an amplifier is given by

$$ABF=\frac{1}{\pi}\frac{g_m}{C}. \quad ...(5)$$

Hence the gain-band-width changes inversely as the inter electrode capacity.

The tube resistance also influence operation at nigh frequencies in that they form significant portion of the tuned circuit, *i e.*, at ultra-high frequencies part of the resonant circuit lies inside the tube. Thus lead inductances and interelectrode capacitances limit the actual operation of the tube. Even if the tube resistance does not give rise to resonance, it may mismatch the generator and the load.

The above adverse effects are minimised by making the tube small which reduces inductances and interelectrode capacities and making tube structure such that electrode leads are incorporated into external concentric line resonators. Individually, lead inductances are reduced by making lead small in the the length and of large diameter, and capacities are reduced by projecting the leads widely separated outside the tube.

15·14-2. CIRCUIT AND TUBE LOSS LIMITATIONS :

The power loss associated with the tube and circuit increases with the frequency. The associated resistance (due to skin effiect) and losses increase as the square root of frequencies. Dielectric and hysteresis losses increase as the first power of the frequency. The radiation loss increases as the square of the frequency. These deterimental effects are lessened as follows :

(*i*) Resistance losses are reduced by increasing the current carrying surface area.

(*ii*) Dielectric losses are lessened by proper positioning of glass with respect to the points of low electric field.

(*i i*) Radiation losses are reduced by enclosing the tube and circuits or by using concentric line construction.

15 14-3. ELECTRON-TRANSIT TIME LIMITATION :

In all tube applications, a finite time is required for electrons to traverse the space between two electrodes because of finite electron mass. The time taken by an electron in going from one electrode to other is known as transit time and it is of the order of 10^{-8} to 10^{-10} seconds. Its exact value depends on electrode separation and impressed voltage. For the frequencies below 10 mc/s, it is neglected because it is less than the period of signal cycle. At high frequencies, the transit time becomes comparable or even greater than the period of signal cycle. At this stage a phase lag is introduced between current carried by the electrons and potential that determines the current. This phase lag affects the operation of the tube adversely.

Transit time calculation : To calculate the transis time between two electrodes, we must know the velocity of electrons between these electrodes or the potential distribution between them. To do this, following assumptions are made :

(*i*) The electrons are emitted from the cathode almost with zero initial velocity and hence potential minimum near cathode is neglected.

(*ii*) The grid acts like an equipotential surface of the potential V_o where

$$V_o=\left(V_b+\frac{V_p}{\mu}\right), \qquad \text{...(1)}$$

with V_g, the grid potential, V_p anode potential and μ the amplification factor of the valve.

(*iii*) The grid behaves like the anode of a diode with potential V_g.

The last assumption yieds for following potential distribution

$$V = V_g \left(\frac{x}{x_{cg}} \right)^{4/3} \qquad \ldots(2)$$

where V is potential at a distance x from the cathode and x_{cg} is the distance between cathode and grid.

Now if the valve is working as an amplifier, then grid potential is given by

$$V_g = E_c + V_1 \sin \omega t, \qquad \ldots(3)$$

where E_c is bias voltage.

The attraction of the signal are amplified by the valve and appear in the anode potential. Thus both the terms in equation (1) vary with time. It means that potential distribution and therefore velocity between cathode and grid vary with time. In radio-frequency amplifier

$$V_1 \leqslant | E_c |. \qquad \ldots(3)$$

Therefore, the effective potential in the grid plane may be taken as constant as far as its effects on grid potential are concerned. The potential and velocity are then functions of distance and not of the time.

Now
$$V_g = \left(E_c + \frac{V_q}{\pi} \right), \qquad \ldots(5)$$

where V_q is the quiescent value of anode potential.

The velocity of the electron at any point of potential V is given by

$$v = \sqrt{\left(\frac{2e}{m} V \right)}$$

$$= \left\{ \frac{2e}{m} V_g \left(\frac{x}{x_{cg}} \right)^{4/3} \right\}^{1/2} \qquad \ldots(6)$$

Hence time taken by the electron to go from cathode to grid is given by

$$T_{cg} = \int_0^{x_{cg}} \frac{dx}{v}$$

$$= \int_0^{x_{cg}} \frac{dx}{\left\{ \frac{2e}{m} V_g \left(\frac{x}{x_{cg}} \right)^{4/3} \right\}^{1/2}}$$

$$= \frac{x_{cg}^{4/3}}{\sqrt{\left(\frac{e}{m} V_g \right)}} \int_0^{x_{cg}} \frac{dx}{x^{2/3}},$$

$$T_{cg}=\frac{3x_{cg}}{\sqrt{\left(\frac{2e}{m}V_g\right)}} \qquad ...(7)$$

Now to calculate the transit time between grid and anode, we assume that potentials of electrodes do not vary with time and the potential distirbution between the grid and anode is linear. It is justified when valve is not working near cut-off. Under such situations the electron is accelerated by constant electric field. The average velocity of electron is given by

$$v=\frac{v_g-v_p}{2}, \qquad ...(8)$$

where v_g is the velocity with which electron crosses the grid plane and v_p is velocity with which it strikes the anode plane. Hence the tiansit time between grid and anode is given by

$$T_{gp}=\frac{x_{cp}-x_{cg}}{\frac{1}{2}(v_g+v_p)}. \qquad ..(9)$$

The velocities of electron at the grid and anode are

$$\left.\begin{aligned} v_g&=\sqrt{\left(\frac{2e}{m}\right)V_g},\\ v_p&=\sqrt{\left(\frac{e}{m}\right)v_p}. \end{aligned}\right\} \qquad ..(10)$$

The following values are typical for a voltage amplifying valve

$x_{cg}=0{\cdot}05$ cm.,
$x_{cp}=0{\cdot}15$ cms.,
$E_g=-1$ volt.
$V_p=200$ volts,
$\mu=50$,

$$\frac{e}{m}=1{\cdot}76\times10^8 \text{ coulombs./gm.}$$

For these valves, we have

$v_g=1{\cdot}03\times10^8$ cms./sec.,
$v_p=8{\cdot}38\times10^8$ cms./sec

$$V_g=-1+\frac{200}{50}=3 \text{ volts.}$$

Hence

$$T_{cg}=\frac{3\times5\times10^{-2}}{\sqrt{(2\times1{\cdot}76\times10^8\times3)}}$$
$$=1{\cdot}5\times10^{-9} \text{ seconds} \qquad ...(11)$$

and

$$T_{gp}=\frac{10^{-1}}{\frac{1}{2}(1{\cdot}03+8{\cdot}38)\ 10^8}$$
$$=2{\cdot}1\times10^{-10} \text{ second.} \qquad ...(12)$$

Thus we see that electrons go much quickly from grid to anode than cathode to the grid.

Transit time effect :

A high input admittance results at ultra-high frequencies due to transit time effect. If the tube is working as an amplifier or as an oscillator, at moderate frequencies, electrons flow from cathode to anode during the positive half cycle of the grid potential. The electrons are accelerated by the field of a.c. grid potential in cathode-grid region and retarded in grid anode region. Therefore net exchange of energy. when transit time effects are negligible, between electrons and field is zero. Hence input conductance due to this effect is also zero.

At ultra-high frequencies, the transit angle is relatively large. What happens that a.c grid potential reverses its phase before the electrons reach the plate. As a consequence of this, electrons are accelerated in both cathode-grid region and grid-anode region. Under these conditions there is a net transfer of energy from the grid potential source to the electrons. This increases the input conductance. The magnitude of input admittance and input conductance is given by

$$Y_{in}=k_1 g_m f T, \qquad \ldots(13)$$

$$G_{in}=k_2 g_m f^2 T^2, \qquad \ldots(14)$$

where k_1 and k_2 are proportionality constants and T is the quantity proportional to the transit time. The magnitudes of Y_{in} and G_{in} increase with the frequency.

We now consider the energy transfer with respect to a.c. plate potential. Normally, the electron should flow through the grid plate region during the regarding phase of a. c. anode potential so that energy is transferred by the electrons to the field, *i e*., to the load impedance. However, at high frequency, a.c. anode potential reverses its phase before the electrons reach at the anode. Thus energy is extracted from a.c. plate potential during a portion of cycle. This decreases the power cutput and efficiency. Thus the electron transit time decreases power output efficiency of the tube and increases input conductance of valve as well as the capacity between cathode and grid. The last two in turn again hamper the tube operation.

An expression for input conductance : Consider two plane electrodes, one donor and the other acceptor of electrons. Let the potential difference applied between the two be E Then as an electron flies from one electrode to the other, it induces charge on both of them. The magnitude of the induced charge on the electrode to be approached increases, while on the other decreases. This means that due to the electronic charge flow, instantaneously changing induced charges of opposite nature appear on two electrodes. These induced charges result in generating a pulse of electric current, the duration of the pulse is equal to the time taken by electrons to reach from one electrode to the other, *i.e*., transit time.

Consider that acceptor electrode is at potential E with respect to donor at zero. Let the charge be $-e$ passing at a point where the potential is V, the work done on it by battery is eV. The work done eV is equivalent to flow of charge by battery given by $e\left(\frac{V}{E}\right)$, with the result that the acceptor electrode acquires a positive charge $e\left(\frac{V}{E}\right)$. The donor plane is left with an image charge $+e$ due to emission $-e$ charge from it ; hence reduction of charge on it is $e\left(\frac{V}{E}\right)$. These changes in charges on two electrodes are due to changes in induced charge. The change in induced charge is due to changes in electrostatic lines of force due to flow of e. The charge starts at the instant of emission and finishes at the instant of reaching the acceptor electrode, *i.e.*, a current pulse flows for this duration due to induction effects.

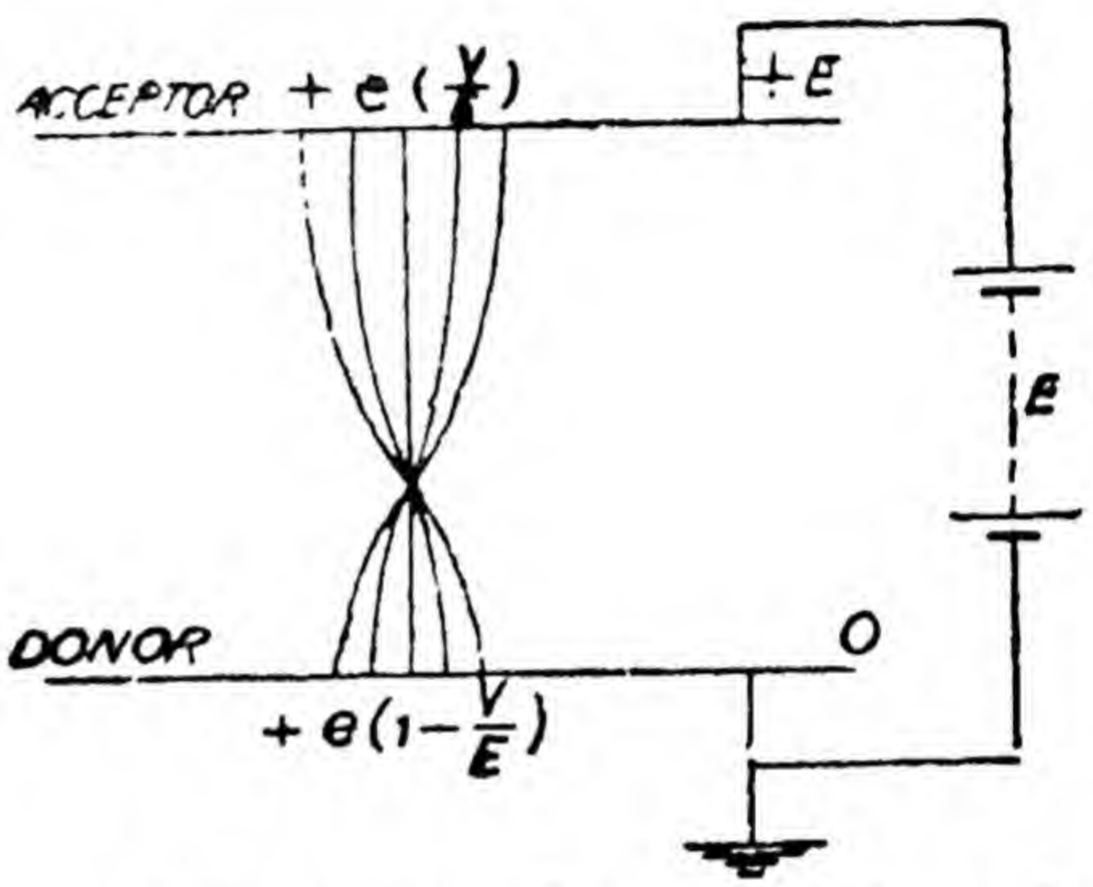


Fig. 29 (a) Charges of lines of force due to passage of charge between two electrodes and appearance of changing induced charges.

If now acceptor piane is replaced by grid, a mesh work of wire, the pulse direction will change as soon as the electronic charge has passed it, *i.e* , a current will flow in one direction when the electrons are reaching it and in the other direction when receding from it. Usual transit time is of the order of 10^{-9} sec. If the time period of the signal applied is greater than 10^{-9} sec , the r.f. pulses produced by induction cancel each other. The necessary condition is that the potential of the grid should remain constant during this time *i.e.*, for the time of passage of electron from the plate. The changes in such pulse current are shown graphically where T_1 and T_2 are the transit time from the cathode to the grid and grid to anode respectively. Total transit time is (T_1+T_2). For complete cancellation the upper and lower halves should be identical.

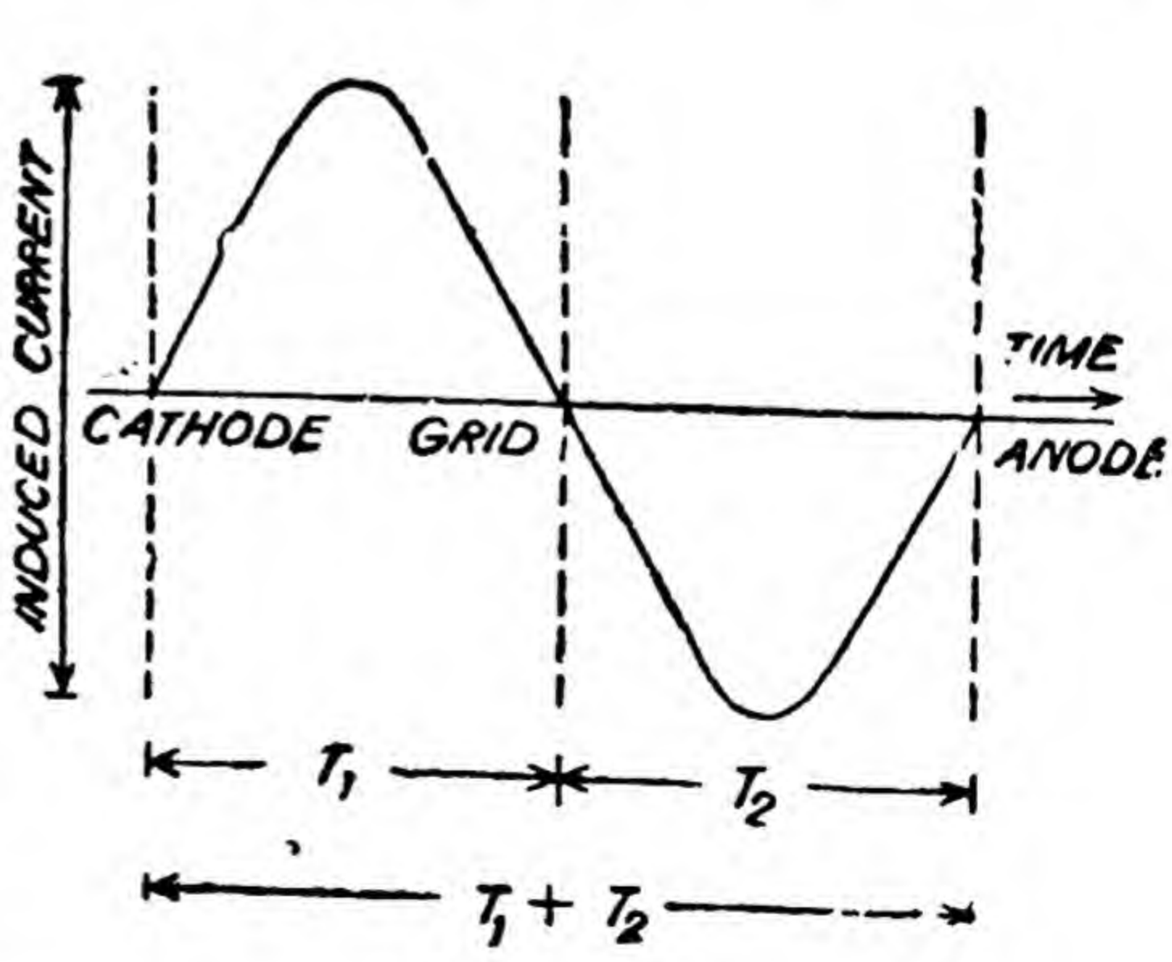


Fig. 29. (b)

For frequencies upto 10 mc./s. the cancellation is almost complete. But above these frequencies, the flow of

current through the grid is sufficient and decreases the input impedance seriously. Input conductance is approximately given by

$$\frac{1}{R_g} = g_m\omega^2 (T_1T_2 + \tfrac{1}{2}T_2^2),$$

where g_m is the mutual conductance of the valve. North has shown that

$$\frac{1}{R_g} = \frac{g_m\omega^2 T_1^2}{10} \text{ approximately.}$$

Let the signal voltage be $V = V_0 \sin \omega t$. The current induced by electrons approaching the grid

$$I_1 = \frac{\mu V_0}{r_p} \sin \omega (t - T_1) = g_m V_0 \sin \omega (t - T_1),$$

where μ and r_p are the amplification factor and internal resistance of the valve respectively.

$(t - T_1)$ shows the time for which current pulse flows ; I_2, the current induced in the grid by receding electrons is

$$I_2 = -g_m V_0 \sin \omega (t - T_1 - T_2),$$

negative sign is for reversed current.

Net grid current is therefore

$$I_g = I_1 + I_2 = g_m [V_0 \sin \omega (t - T_1) - V_0 \sin \omega (t - T_1 - T_2)]$$

or $$I_g = 2g_m V_0 \cos \omega [t - (T_1 + \tfrac{1}{2}T_2] \sin \tfrac{1}{2}\omega T_2.$$

The product ωT_1 or ωT_2 is very small even at high frequencies, so that $$I_g = 2g_m V_0 \omega (T_1 + \tfrac{1}{2}T_2).\tfrac{1}{2}\omega T_2 \text{ approximately.}$$

or $$\frac{I_g}{V_0} = g_m\omega^2 (T_1T_2 + \tfrac{1}{2}T_2^2) \text{ approximately.}$$

Only resistive component has been taken into account.

Thus $$\frac{I_g}{V_0} = \text{input conductance} = g_m\omega^2 (T_1T_2 + \tfrac{1}{2}T_2^2).$$

For $g_m = 5mA/V$, $(T_1 + T_2) = 10^{-9}$ sec $(T_1 \simeq T_2)$,

$$\frac{I_g}{V_0} = 5 \times 10^{-12} f^2 \text{ mhos.}$$

For a frequency of 10^8 cycles/sec.. the value of input conductance is sufficiently high and undesirable.

15·15. POSITIVE GRID OSCILLATOR :

In 1919 Barkhausen and Kurz observed that when the grid of a triode is at a high positive potential and the anode is at a negative potential, high frequency oscillations appeared in the external circuit. The frequency depends on the spacing and potentials of the electrodes of the tube and on the constants of the external circuit used. The positive grid oscillator consists of a parallel plane cylindrical element triode tube with the grid at a positive d.c. potential with respect to its cathode and plate as shown in the figure 30 and 31.

In Barkhausen Kurz oscillator a load is a connected between anode and cathode. At high frequencies, the load is in the form of short circuited transmission line. Normally, it is represented

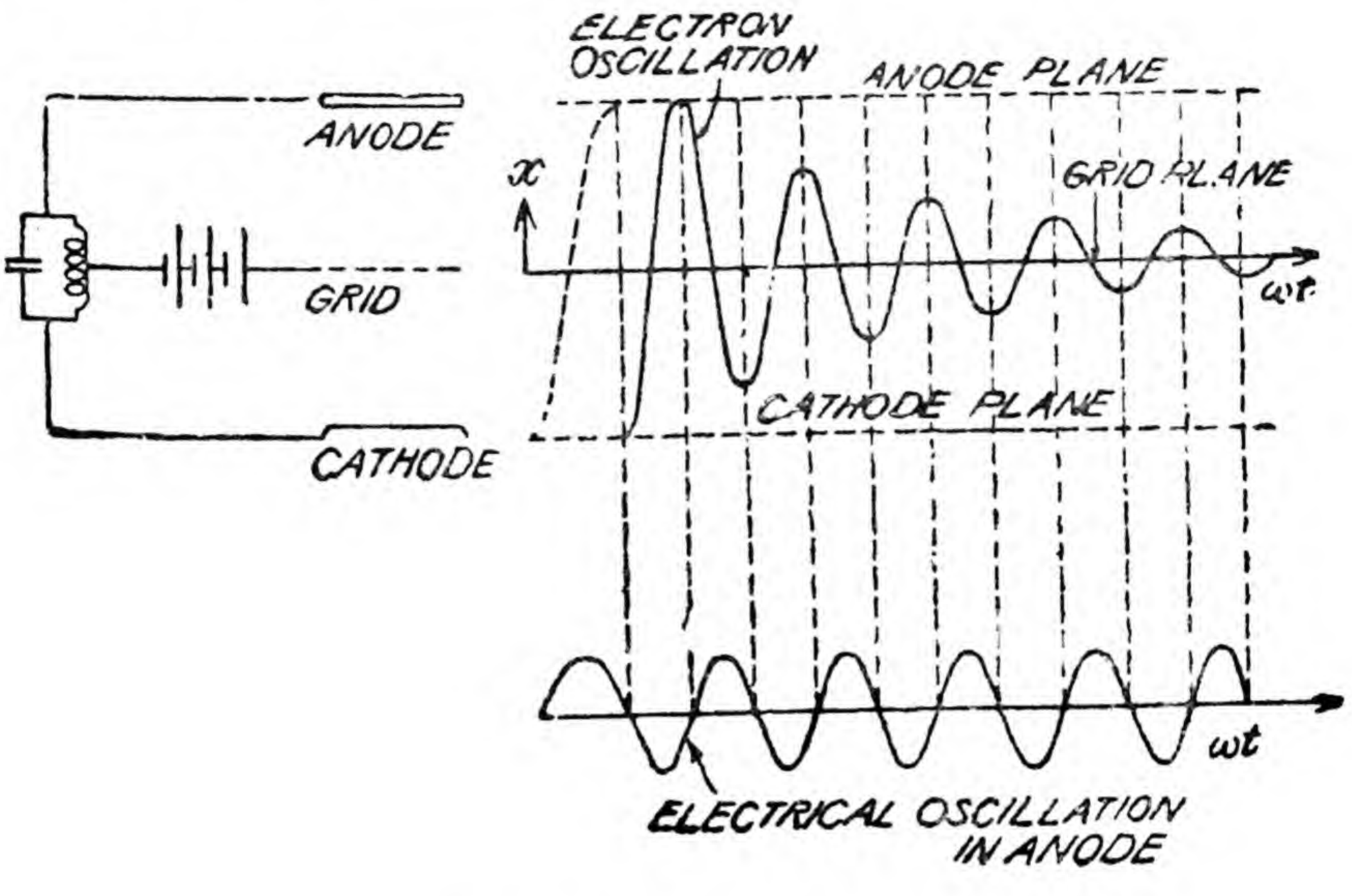


Fig. 30. Positive grid oscillator.

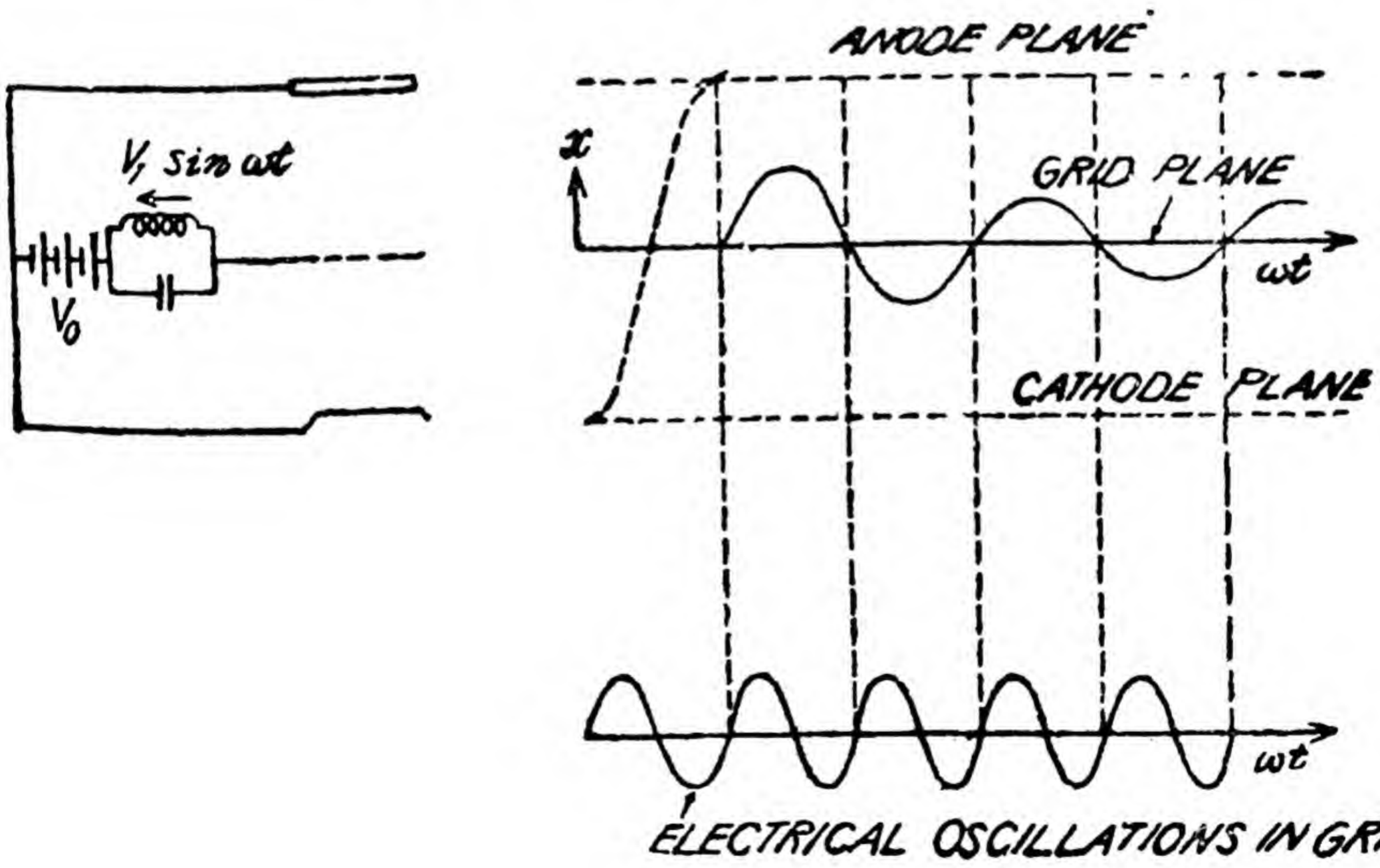


Fig. 31.

by an LC circuit. A high tension bypass condenser is also used to connect alternating quantities to earth from the grid (not shown in the fig.). In the fig. 31, an alternating arrangement is shown in which load is connected between grid and other electrodes.

Principle :

In the absence of load impedence, there is only accelerating field between cathode and grid, and anode and grid. Therefore, whenever an electron approaches the grid, it is accelerated by the d.c. applied potential. The electron which has been already accelerated in cathode grid region enters into grid-plate region where it is decelerated. It comes momentarily at rest somewhere near the plate and then starts travelling towards the grid plane and is accelerated. The electron again faces deceleration in grid cathode region, comes at rest near the cathode and reverses its direction of travel. Thus the electrons continues to oscillate back and forth round the grid plane. In the presence of load impedence, the oscillating space induces an alternating component of current in the external circuit. The resulting voltage drop across the load impedance establishes an alternating field in the inter-electrode space of the tube. The electrons could oscillate about the grid plane in such a fashion that it is retarded by this field and hence transfer a part of energy to the external circuit of the tube during each cycle of electron oscillation. This transfer of energy maintains sustained oscillations at the output.

Mechanism of oscillations :

In Barkhausen Kurz oscillator, the measured period of oscillation is of the order of the transit time and, therefore, it is assumed that time of vibration of electron about the grid plane is equal to the period of electrical oscillation (*fig. 30*). To understand the transfer of energy, we divide the alternating potential in two halves, namely, positive half when anode is positive with respect to cathode, and negative half when anode is negative with respect to cathode. Thus in the positive half cycle, the potential difference between cathode and anode increases the accelerating field strength and decreases the retarding field strength and reverse is the case in negative half cycle. Thus an electron which is emitted at the beginning of positive half cycle, gains more energy in cathode grid region than it loses in grid anode region. The electron is left with same residual energy and therefore this electron strikes the anode and dissipates its energy there in the form of heat.

The electron which is emitted from the cathode at the start of negative half cycle covers the cathode-grid region in weakend field and therefore gains less kinetic energy than it would require to cover the grid-plate region because there it faces much strengthen field. The consequence is that electron comes at rest somewhere near the cathode for a moment and then reverses its direction of travel. By the time the polarity of electric field changes, *i.e.*, anode becomes positive with respect to cathode and so the electron is again accelerated by a weakend field and retarded in

grid cathode region by a strengthen field. Thus the electron comes at rest near the cathode and again starts repeating the process of back and fourth journey. It should be noted that after completing a cycle of oscillation electron comes at rest nearer and nearer to the grid plane, because each time it is accelerated by a lesser amount of potential than its predecessors. Hence the motion of the electron is somewhat analogous to a damped pendulum as shown in fig. (30). Finally the vibrating electron strikes the grid wires and dissipates its residual kinetic energy in the form of heat. The residual kinetic energy corresponds to the potential though it was less accelerated and it is very much less than the initial kinetic energy gained by the electron from the d.c. source.

It is seen that electron emitted from the cathode at the start of negative half cycle faces stronger retarding field than the accelerating field in each cycle and hence it delivers energy to the external circuit, while the electron which is emitted at the start of positive half cycle gains energy from the field. But the latter electron strikes the plate in its first excursion. hence there is net energy delivered to the load by the former electron and oscillations are sustained. The difference of energy drawn from the d.c. source and power dissipated in the form of heat both at the grid and plate gives us the power generated in the form of electrical oscillations.

When the load is connected in the grid circuit of fig. 31 :

In this case frequency of the current induced is doubled, *i.e.* transit time of the electron between the grid and cathode or anode is equal to half of the period of electrical oscillations. In this case, we call the half cycle as positive when the grid is positive with respect to grid potential V_0 and negative when the grid is negative with respect to grid potential V_0. In order to consider the maintenance of oscillations in the load, we consider the behaviour of electrons in the valve when the potential across the load is alternating through half cycles.

Consider two case of the electron which leaves the cathode at the start of a positive half cycle. When the electron moves from cathode to grid it experiences the accelerating field which is stronger than the average, and hence the electron passes through the grid with more kinetic energy than corresponding to the mean grid potential. Thus the electron gains more kinetic energy between cathode and grid, In the following negative half cycle, the electron travels towards the anode with a weakened retarding field. Due to this weakened retarding field, it loses less kinetic energy and therefore strikes the plate, and dissipates its residual kinetic energy as heat.

Now consider the case of the electron which leaves the cathode at the start of a negative half cycle. Between cathode and grid it is accelerated by the weakened electric field and thus gains less kinetic energy. Now on the following positive half cycle between grid and plate it is retarded by a strengthened field, it thus loses its whole energy acquired between cathode and grid, and therefore fails to reach the plate. At the instant, the electron comes to rest, the varying grid potential starts on the negative half cycle and the field which draws the electron back to the grid is weaker than the average. It thus gains less kinetic energy. Now on the following positive half cycle, it is retarded by a strengthened field between grid and cathode and therefore loses its all acquired kinetic energy. Thus the electron is alternately accelerated by a weakened field, and retarded by a strengthened field, and it oscillates about a grid plane with a decreasing amplitude. As the electron does so, it induces currents in the grid circuit which delivers energy to load. Finally the electron strikes the grid.

The electrons which leave the cathode at the positive half cycle pass only once through the tube and draws energy from the load, while the electrons leaving the cathode at a negative half cycle strikes the grid after making several vibrations, inducing current in the grid circuit which delivers energy to the load. On an average, the electrons give energy to the load, and so they maintain continuous oscillations.

Analysis. In order to derive the expression for frequency, we assume two things :

(*i*) space charge density is small, so that it does not affect the electric field distribution in the interelectrode space,

(*ii*) distances are measured from the grid.

The potential difference between grid and plate is

$$V = V_0 + V_1 \sin \omega t. \qquad \ldots(1)$$

The motion of an electron in time varying field is given by (according to article 15·13)

$$\frac{x}{K} = A + \frac{V_1}{V_0} B + C, \qquad \ldots(2)$$

where $K = \frac{-e}{md} \cdot \frac{V_0}{\omega^2}$, $A = \frac{(\omega T)}{2}$, $C = \frac{\omega T}{K}$,

$$B = \cos \phi \, (\omega t - \sin \omega t) + \sin \phi \, (1 - \cos \omega t),$$

K is negative in our present case, because the terminals of battery are changed.

We now assume that the electron completes one half cycle of oscillation and returns to the grid while the alternating potential completes one full cycle, *i.e.* $T_1 = \frac{2}{f}$ where T_1 is the time period of electron oscillation.

Hence the electron leaves the grid plane $x=0$ when $\omega T=0$ with an initial velocity v_0 and returns to the grid plane when $\omega T=2\pi$.

Putting these values in equation, we get

$$\frac{x}{K}=\frac{(\omega T)^2}{2}+\frac{V_1}{V_0}\{(\omega T-\sin\omega T)\cos\phi+(1-\cos\omega T)\sin\phi\}+\frac{v_0 T}{K},$$

at the instant at which the electron returns to the grid plane by setting $x=0$ and $\omega T=2\pi$, we obtain

$$0=\frac{4\pi^2}{2}+\frac{V_1}{V_0}\{2\pi\cos\phi+0\}+\frac{v_0 T}{K}\frac{\omega}{\omega}$$

$$=2\pi^2+\frac{V_1}{V_0}2\pi\cos\phi+\frac{v_0 2\pi}{K\omega},$$

$$0=\pi+\frac{V_1}{V_0}\cos\phi+\frac{v_0}{K\omega},$$

$$\frac{v_0}{\omega}=-K\left(\pi+\frac{V_1}{V_0}\cos\phi\right),$$

$$\frac{v_0}{\omega}=\frac{e}{md}\frac{V_0}{\omega^2}\left(\pi+\frac{V_1}{V_0}\cos\phi\right),$$

$$\omega=\frac{e}{m}\frac{V_0}{dv_0}\left(\pi+\frac{V_1}{V_0}\cos\phi\right). \qquad ...(3)$$

Period of electron oscillation is

$$T_1=\frac{2}{f}=\frac{2}{\omega/2\pi}=\frac{4\pi}{\omega},$$

$$T_1=\frac{4\pi m v_0 d}{eV_0\left(\pi+\frac{V_1}{V_0}\cos\phi\right)}. \qquad ...(4)$$

The above equation shows that period of electron oscillation depends upon the initial phase angle ϕ (*at the grid plane*). Since the electrons cross the grid plane at various values of ϕ, the period of electron oscillation is different for various electrons and it depends upon their initial phase. The period of electron oscillation also depends upon its velocity at the grid plane. The electron which vibrates about the grid plane so as to impart energy to the load, gets its velocity reduced in successive cycle of vibration and hence its period of oscillation. Thus every electron gives rise to a situation of electron oscillation of varying frequencies. The frequency of electrical oscillation is determined by the composite effects of all such electrons.

The velocity v_0 in equation (4) is determined largely by the d.c. potential ; hence we may assume that

$$\tfrac{1}{2}mv_0^2=eV_0,$$

$$v_0=\sqrt{\left(\frac{2eV_0}{m}\right)}. \qquad ..(5)$$

Then
$$T_1=\frac{4\pi m d\sqrt{\left(\frac{2eV_0}{m}\right)}}{eV_0\left(\pi+\frac{V_1}{V_0}\cos\phi\right)}$$
$$=\frac{4\pi d\sqrt{(2m)}}{\sqrt{(eV_0)}\left(\pi+\frac{V_1}{V_0}\cos\phi\right)}. \quad \ldots(6)$$

Now the value of a.c. field is very small in comparison with d.c. field, so that the term $\frac{V_1}{V_0}\cos\phi$ is negligible. Thus

$$T_1=\frac{4\pi d\sqrt{(2m)}}{\sqrt{(eV_0)}\,\pi}=\frac{d}{\sqrt{(V_0)}}\frac{4\sqrt{(2m)}}{\sqrt{e}}.$$

Putting the values of m and e,

$$T_1=1\cdot35\times10^{-5}\frac{d}{\sqrt{(V_0)}}\text{ seconds.} \quad \ldots(7)$$

Frequency of electrical oscillation is,

$$f=\frac{2}{T_1}=\frac{2\sqrt{(eV_0)}}{4d\sqrt{(2m)}}=1\cdot48\times10^5\frac{\sqrt{(V_0)}}{d}\text{ cycles/sec.} \quad \ldots(8)$$

Wavelength of the electrical oscillation will be

$$\lambda=\frac{c}{f}=\frac{cd}{1\cdot48\times10^6\sqrt{(V_0)}}=2020\frac{d}{\sqrt{(V_0)}}\text{ meters,}$$

where c is the velocity of light.

For $V_0=200$ volts and $d=2$ mm. we find that

$f\approx1050$ mega cycles. $\ldots(9)$

When load is connected between plate and cathode, then

$$f=0\cdot74\times10^5\frac{\sqrt{V_0}}{d}\text{ cycles/sec.} \quad \ldots(10)$$

$=525$ mega cycles for above mentioned values.

The above derivation is based on the assumption that frequency of electrical oscillations depends only on the geometry and electrode potential of the valve. This is known as Barkhausen Kurz criterion of oscillation. Later on, Gill and Morell observed that oscillations are more pronounced when external circuit is resonant than when it is not resonant (other working conditions are kept constant). The slight mistuning in external circuit changes Gill Morell condition to Barkhausen-Kurz conditions. The situation becomes clear when it is understood that the amplitude of potential developed across the load is small when load is nonresonant and it is appreciable when load is resonant. Therefore, the potential developed in the latter case forces the electron vibration in synchronism with natural oscillations of the load and hence pronounced effect under Gill-Morell conditions.

Practical aspects in positive grid oscillator :

(*i*) In analysis it is assumed that there is space-charge near the grid and current is temperature limited. But there is space-charge due to slowly returning electron towards the grid plane. It is the result of experience that the efficiency of the oscillator is maximum when the current is at the verge of two limits.

(*ii*) Since it is not desirable to have space-charge limited current, oxide coated cathode are not used but tungsten filament surrounded by grid and anode is used.

(*iii*) In a cylind rical valve the potential distribution is not uniform even when space-charge is absent. It is due to the large ratio of radii of anode and cathode. This makes transit time of an electron less in grid cathode region than in the grid-anode region.

(*iv*) It has been discussed particularly when load is connected between grid and other electrodes, that transit times in two regions should be equal. But it is not the case due to asymmetrical distribution of potential above the grid plane. A negative anode equalises the two transit times and hence increases the efficiency of the oscillator.

(*v*) Positive grid oscillators are inefficient because a large number of electrons strike the grid plane in their first excursion and dissipate their energy in the form of heat. This warrants heavy grids capable of dissipating quantity of heat.

(*vi*) Since the period of electrical oscillations is determined by potential of high tension supply, a carefully regulated power supply is used.

(*vii*) The power output is also due to the fact that most electrons do not oscillate in favourable phase, and due to the tendency of electrons to shift from a favourable phase to unfavourable phase as they continue to oscillate.

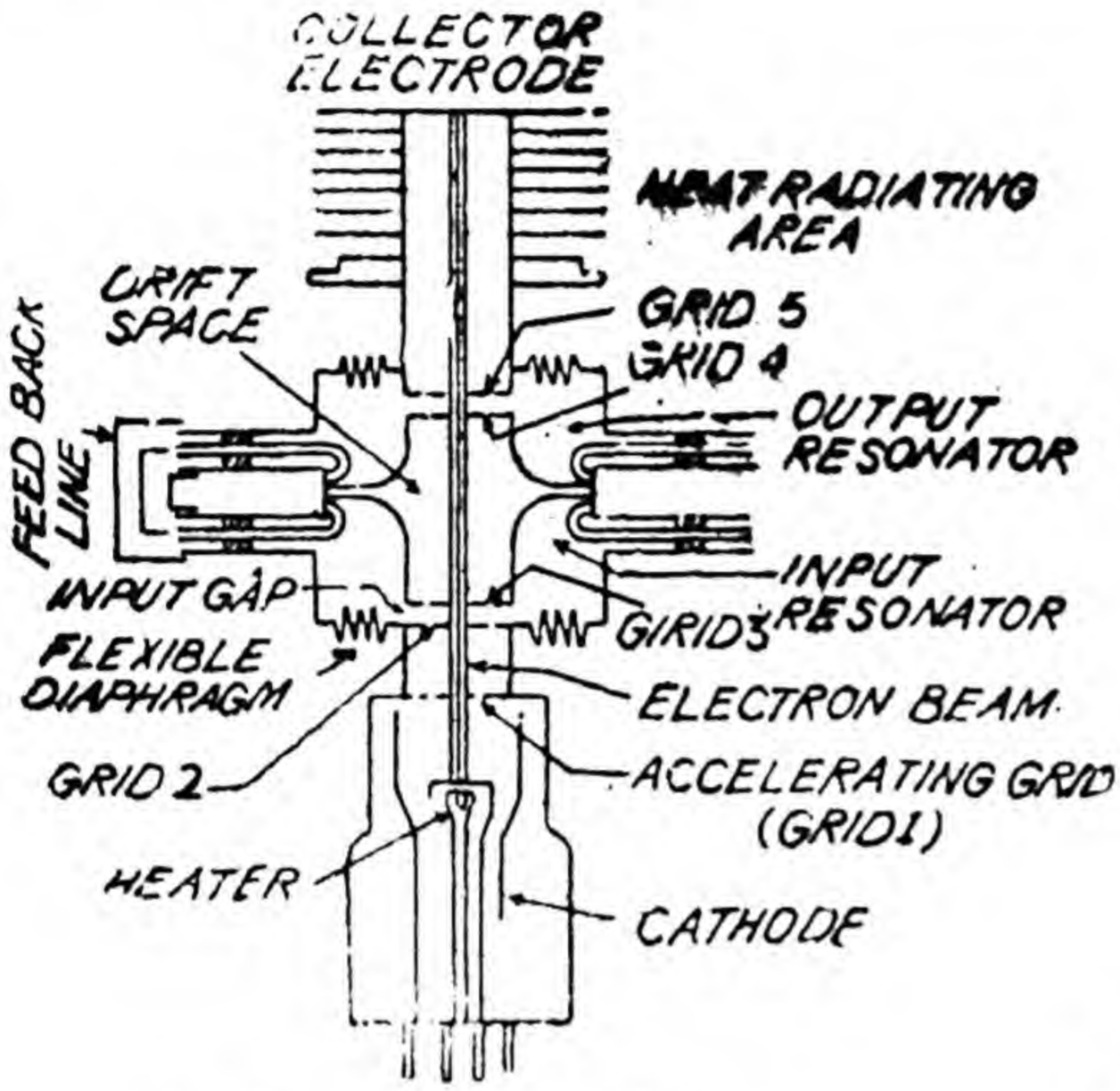
15·16. KLYSTRON OSCILLATOR :

Several types of tubes have been developed, in which the electrons are initially accelerated to a high velocity by a d.c field before they enter the retarding a.c. field. The electrons then have a relatively high velocity as they travel through the a.c. field ; hence the electron transit angle is greatly reduced. The double resonator klystron and the reflex klystron are examples of such tubes.

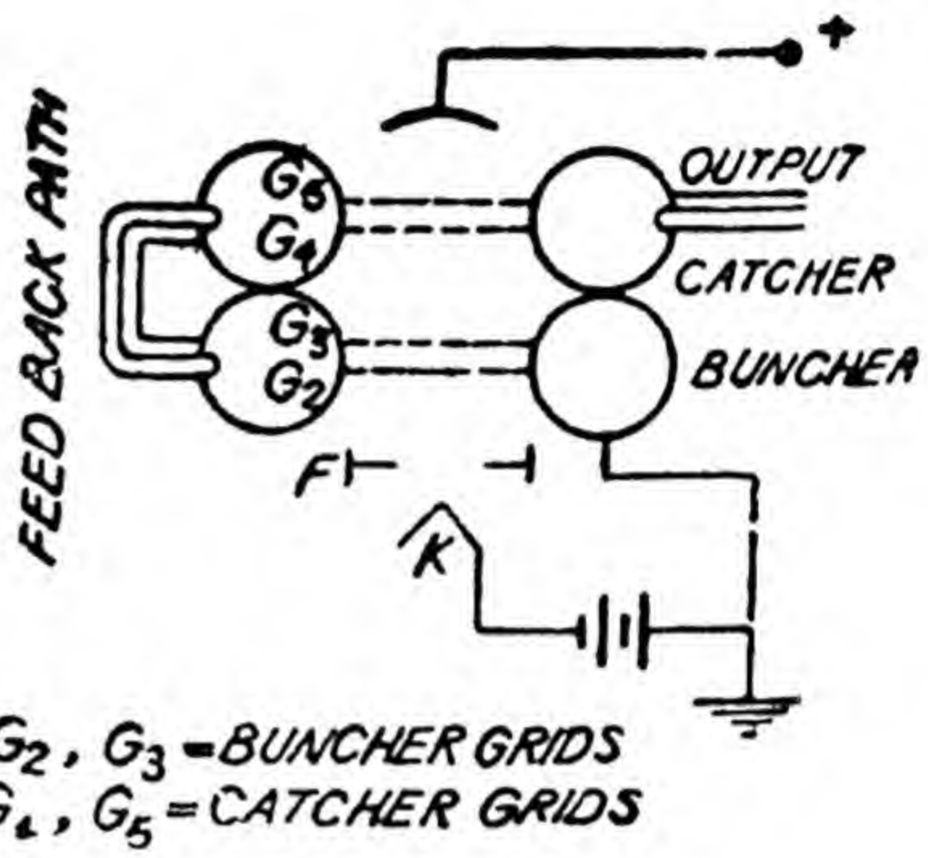
Double Resonator Klystron :

Figure 32 shows the working part of a klystron. It consists of (*i*) oxide coated cathode, (*ii*) a control grid, (*iii*) two metallic resonators, and (*iv*) collector anode. The resonator nearest to the cathode is known as buncher or input resonator and the second resonator is the catcher or output resonator. The input signal is applied to one of the resonators and the amplified signal is obtained

from the other. If a portion of the output signal is returned to the input, the tube work as an oscillator.



(a) Crosssection of a two cavity klystron.



(b) Two cavity klystron without vacuum envelope.

Fig. 32

In klystron, the cathode is a part of an electron gun structure for which the accelerating grid, the two resonators, and the collector electrode form the anode. Due to an accelerating voltage applied to the grid, electrons travel along the axis of the tube, some electrons are intercepted by each of the five grids and remaining reach the collector electrode. The input gap is analogous to the cathode grid space of a triode valve. With the help of flexible diaphragms and a tuning mechanism (not shown in figure), the resonant frequency is controlled by the adjustment of the gap spacing. The collector electrode is given some positive potential to collect even slowest electrons that emerge from the cathode grids.

The input signals applied to the input gap cause an electric field in this gap which brings a variation or modulation of the electron beam. In case of a triode valve, the electronic beam is *density modulated* because when the grid is positive, a dense beam of electron flows from cathode to plate while few electrons (or none) traverse when the grid is negative. The beam is *velocity modulated* in case of klystron. The principle of velocity modulation is shown in figure 33 (*a*).

Let a signal $V_1 \sin \omega t$ be applied between the two grids—grid 2 and grid 3 denoted as G_2 and G_3 and grid system is at a high positive potential. For the first half cycle for which grid G_3 is positive with respect to G_1, the electrons emerge from the grid system with an increased velocity *i.e.* accelerated, while for the

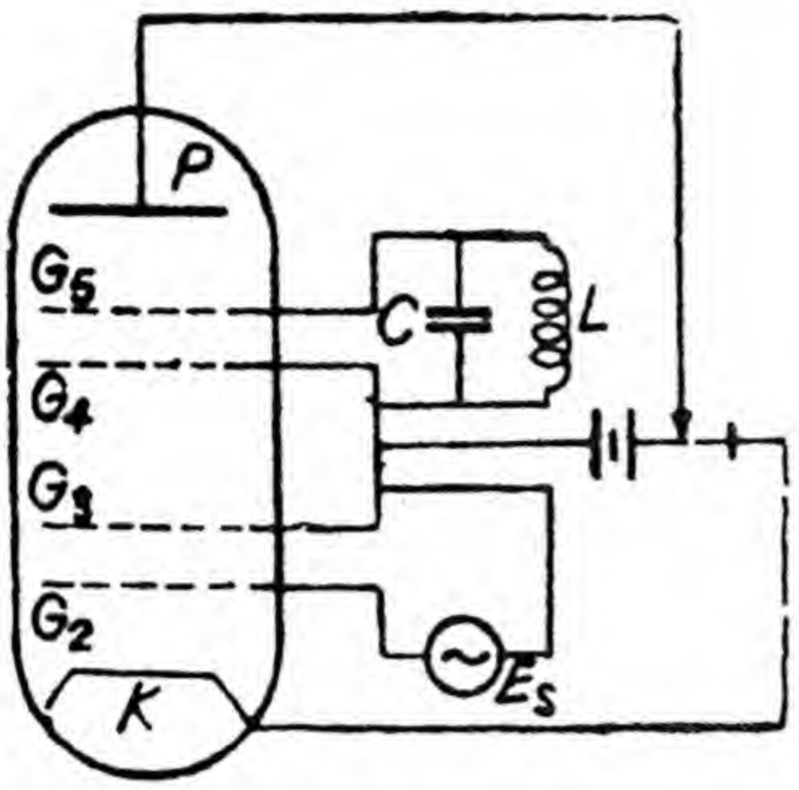


Fig. 33. (a) Principle of velocity modulation.

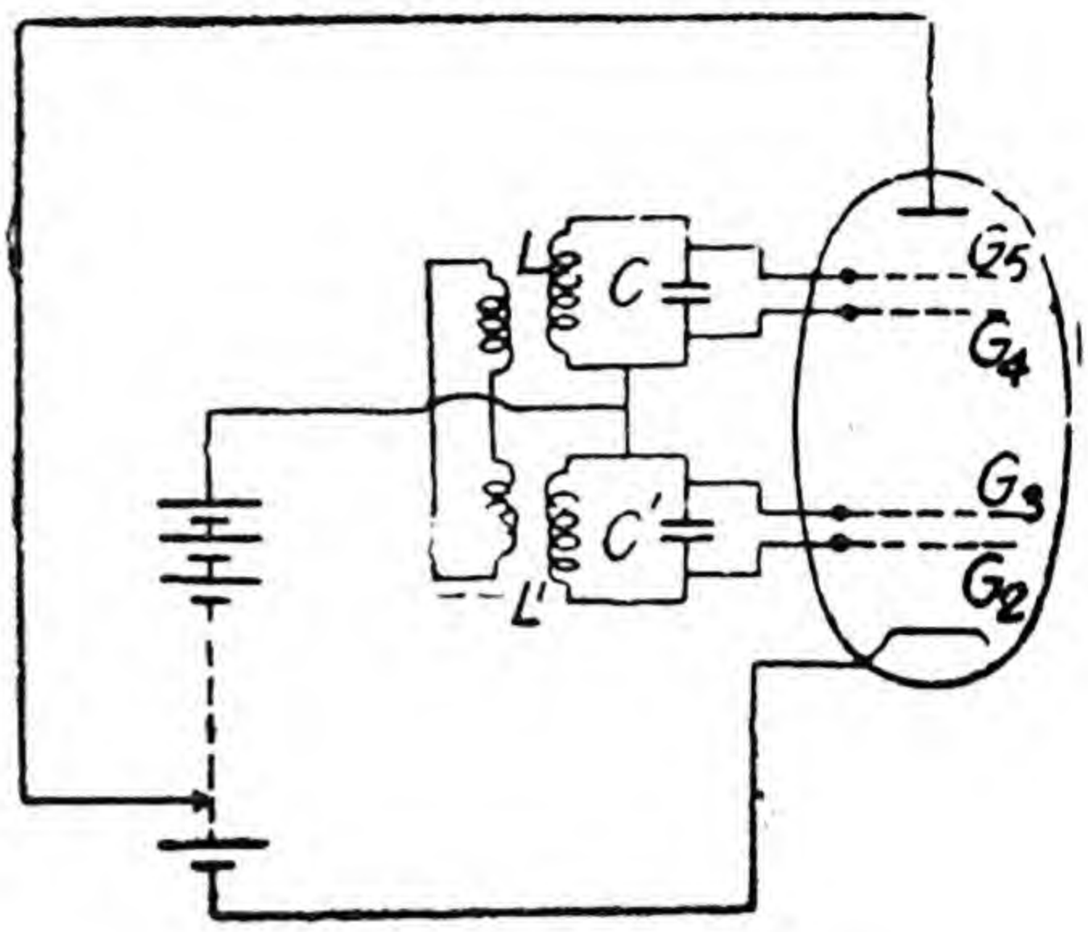


Fig. 33 (b)

next half cycle of the signal, the electrons are decelerated. The electric field in the input gap affects only the velocity of the electrons and thus they are velocity modulated.

The advantage of velocity modulation is that the electrons are accelerated before they reach in the input gap and due to this fact the transit time is very short even though the gap has the appreciable width. The disadvantage of velocity modu lation is that it cannot deliver pulses of current to the tuned circuit at its resonant frequency, *i.e.* it is not usable for production of output power. The difficulty is removed with the help of field free drift space, where the velocity modulated beam is converted into density modulated beam.

In the field free space between buncher and catcher grids, the high velocity electrons overtake the low velocity electrons which left the buncher grids at an earlier phase. The results in a bunching of the electrons as they drift towards the catcher resonator. The bunching process is shown by the Applegate diagram of fig. 34. The figure shows the electron displacement as a function, of time for electrons leaving the buncher at different phases of buncher voltage.

Each line in the diagram has a slope proportional to the velocity of the electron whose progress it represents. At the bottom, the parallel lines represent electrons arriving at the buncher with a velocity v_0. The electrons emerge from the buncher with velocities greater or less than v_0, according to the direction of the electrostatic field between the grids, and the lines representing them become more or less steep. Where any two lines intersect, two electrons come together; the electrons are therefore bunched in the neighbourhood of the line representing centre of catcher resonator. In order to induce the maximum current in the output

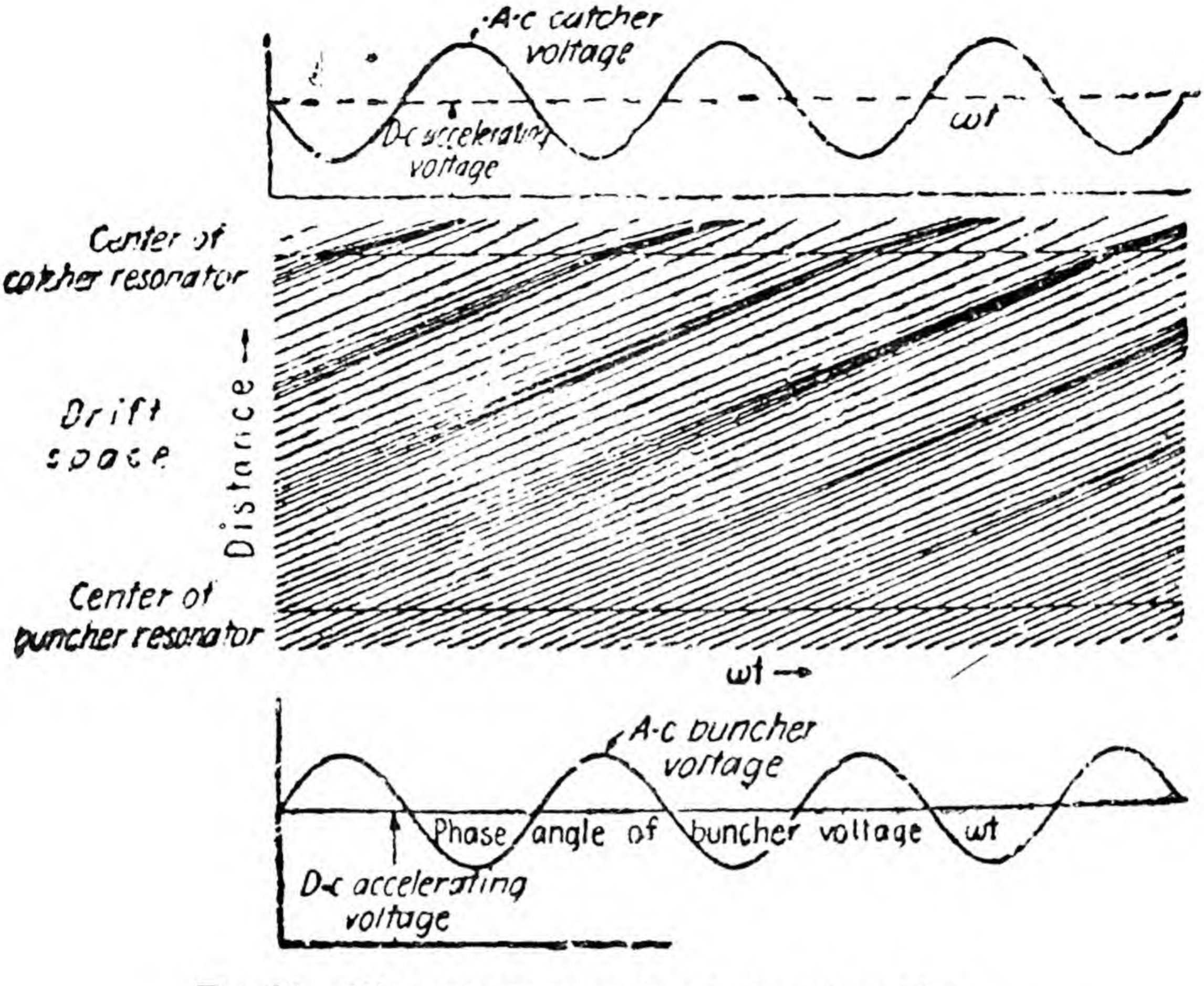


Fig. 34. Diagram representing electron bunching.

cavity of klystron, *the output cavity should be placed in a region where maximum bunching takes place.*

When the bunch of the electrons passes through the output cavity, it induces a.c. which flows through the external circuit of the grid from G_4 to G_5. This induced current initiates the oscillations in the cavity at its natural frequency. For the first half cycle, the field between grids G_4 and G_5 is retarding as G_4 is positive with respect to G_5 and therefore the field takes kinetic energy from the bunch and converts it into the electric energy. For the next half cycle, the field is accelerating as grid G_5 is positive with respect to G_4 and therefore the field transfers energy from resonator to the electrons. If the natural frequency of the resonator is equal to the frequency of bunching process then the density of electrons between grids G_4 and G_5 will be minimum for maximum accelerating field and maximum for decelerating field. The net result is that the net power is delivered to the output cavity. Tne electrons emerge from the catcher grids with reduced velocity and finally terminate at the collector. Thus a small signal applied across the buncher resonator is amplified and system works as a klystron amplifier. If a part of this energy is fed back from catcher to buncher through a coaxial line and no input signal is applied

[shown in fig. 33 (*b*)], the system works as an oscillator and sustained oscilla tions are produced.

The reason for maintaining small transit time for both input and output gaps of klystron is as follows : The appreciable input gap transit time reduces the amount of velocity modulation produced by a given r.f. voltage at the gap while the appreciable output transit time reduces the power delivered to the resonator by a given beam current.

Phase Relationships :

Maximum bunching takes place when the electrons of unchanged velocity catches the electron of least velocity and the electron of unchanged velocity is in turn caught up by the electron of fastest velocity. Thus the centre of the bunch centres round the electron of unchanged velocity. This electrons leaves the buncher system at the time when bunching voltage is zero, *i.e.* changing from *deceleration to acceleration.* If we ignore, for the moment, the phase shift introduced by the drift space, then bunching leads by an angle $\pi/2$, the bunching process. The phase lag introduced in the drift space is $\omega S/v_0$ and bunching lags the bunching voltage by the amount

$$\left(\frac{\omega S}{v_0} - \frac{\pi}{2}\right).$$

When bunch of electrons goes to the catcher grids G_4 which is positive with respect to G_5, its positive sense is taken the same as that of catcher voltage. It means negative maximum of catcher voltage coincides with maximum of bunching. Thus catcher voltage lags bunching by π. The whole of the situation is shown in the fig. (35).

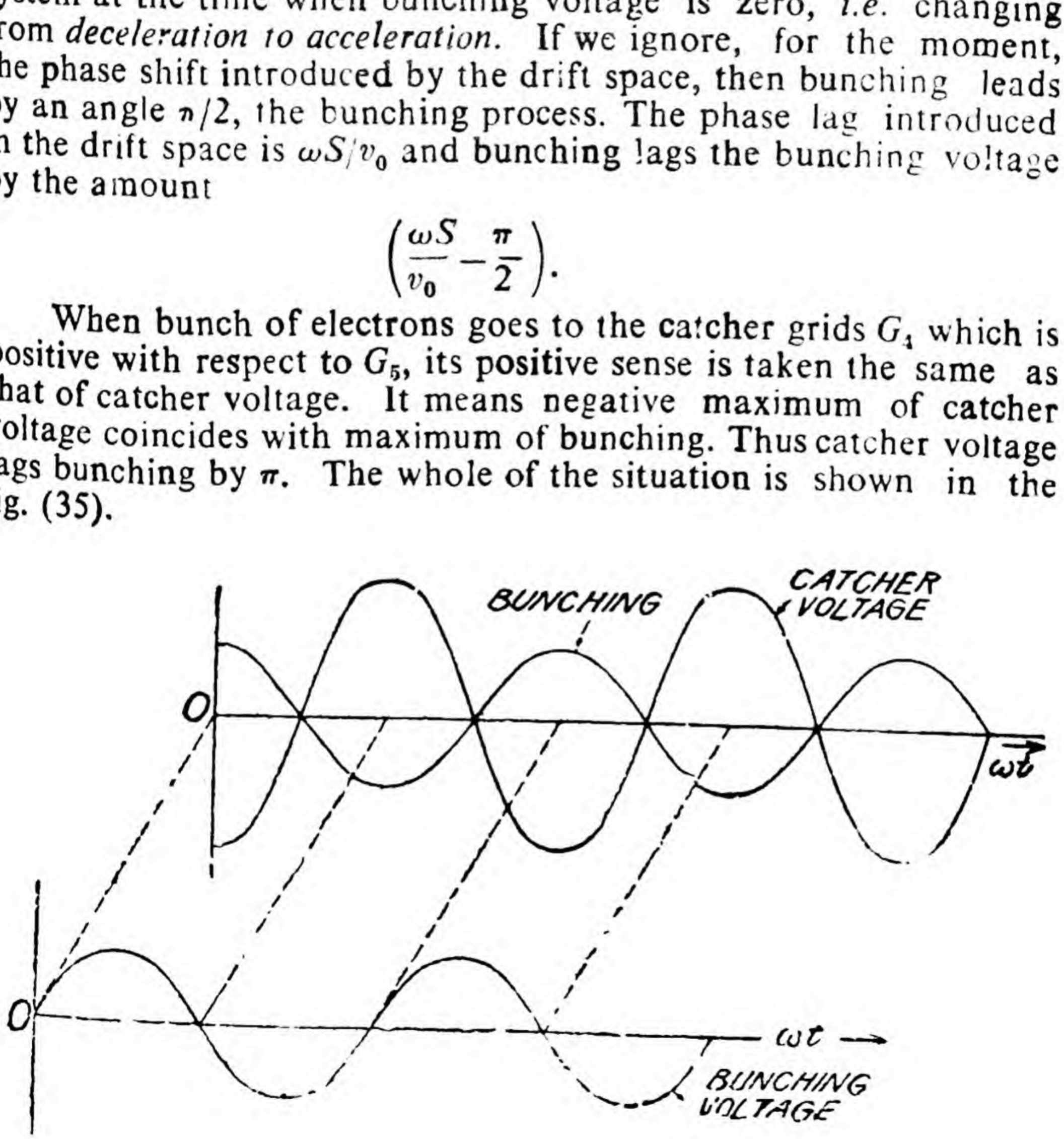


Fig 35. Phase relationships between bunching voltage, bunching and catcher voltage.

Now if klystron is working as an oscillator, a phase ϕ is introduced by the feed back cable. Hence the total phase lag at the input of the klystron becomes

$$\frac{\omega S}{v_0}-\frac{\pi}{2}+\pi+\phi.$$

For regeneration, the total phase lag introduced should be an integral multiple of 2π, *i e*. feed back should be in phase with the input. Now the feed-back cable is so designed as to make ϕ either equal to zero or an integral multiple of 2π. Hence total phase lag is written as

$$\frac{\omega S}{v_0}+\frac{\pi}{2}=2\pi n$$

or
$$\alpha=2\pi n-\frac{\pi}{2},$$

with
$$\alpha=\frac{\omega S}{v_0}$$

Analysis : The analysis of the klystron is based on the following simplifying assumptions :

(*a*) Space charge effects are ignored.

(*b*) The alternating voltage between buncher grids is assumed to be small in comparison with d.c. accelerating voltage.

(*c*) The electron transit angles through the buncher and catcher grids are assumed to be negligible.

(*d*) All the electrons which leave the cathode are assumed to pass through the catcher grids.

(*e*) The electron beam is assumed to have uniform density in the cross-section of the beam.

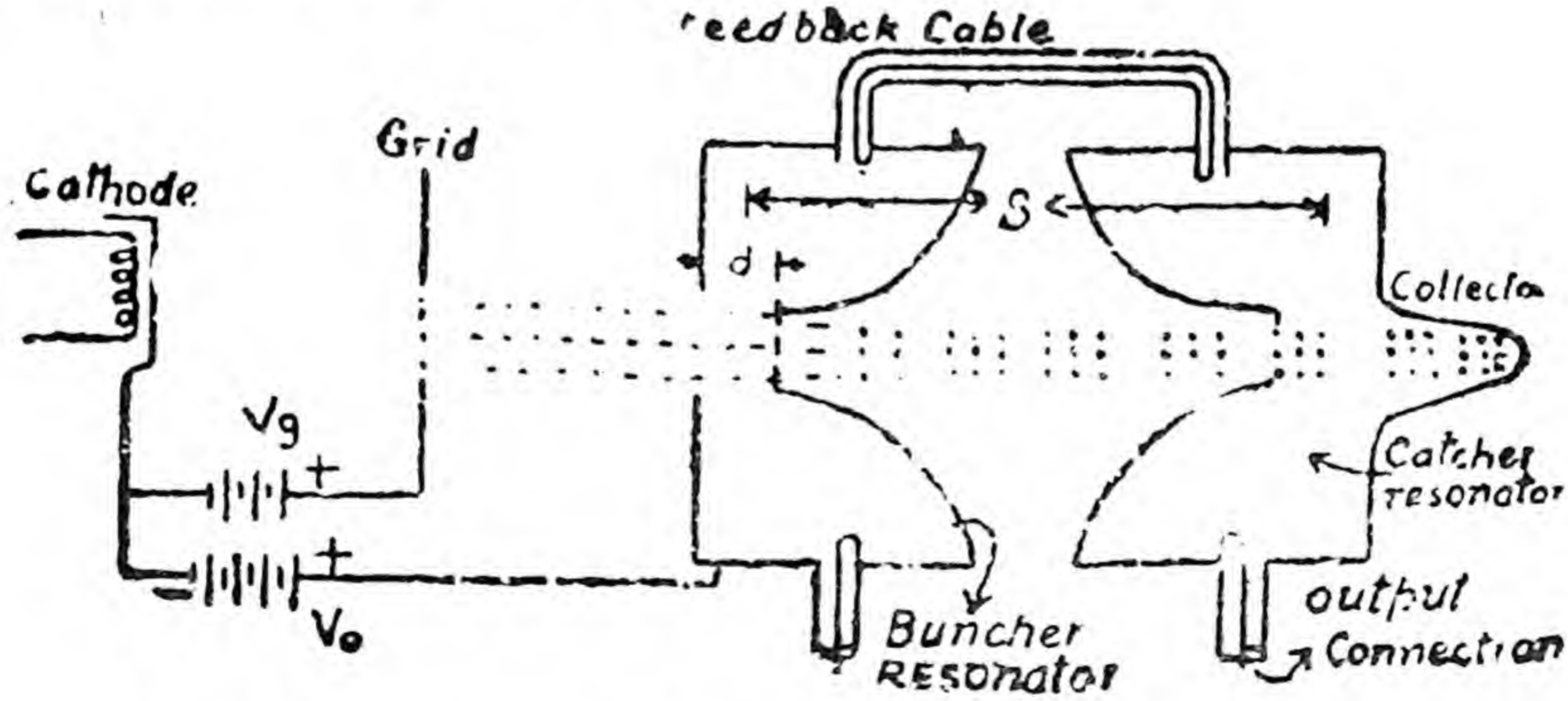


Fig 36. Working parts of Klystron.

Assume that the electrons leave the cathode with zero velocity. They are accelerated to an initial velocity v_0 by the d.c. field before they enter the buncher grids. The kinetic energy and the velocity of the elcctrons entering the buncher is given by

$$\tfrac{1}{2}mv_0^2 = eV_0,$$

$$v_0 = \sqrt{\left(\frac{2eV_0}{m}\right)}. \qquad \ldots(1)$$

Suppose the two grids of the buncher resonator are at the potential V_0 and $V_0 + V_1 \sin \omega t$ respectively with respect to cathode and consider an electron that passes through the buncher grid at the instant t_1.

After emerging from the buncher grids, the velocity of the electron is given by

$$\tfrac{1}{2}mv_1^2 = e\,(V_0 + V_1 \sin \omega t_1),$$

$$v_1 = \sqrt{\left[\frac{2e}{m} V_0 \left\{1 + \frac{V_1}{V_0} \sin \omega t_1\right\}\right]},$$

$$v_1 = v_0 \sqrt{\left(1 + \frac{V_1}{V_0} \sin \omega t_1\right)}. \qquad \ldots(2)$$

The expression shows that the velocities of electrons after buncher grids are sinusoidal, the frequency and phase being the same as that of bunching voltage.

The time T required for the electron to travel the distance S in the field free space between buncher and catcher grid is

$$T = \frac{S}{v_1} = \frac{S}{v_0 \sqrt{\left(1 + \frac{V_1}{V_0} \sin \omega t_1\right)}}. \qquad \ldots(3)$$

Using the number alternating voltage as the time reference, the electron arrives at the catcher grids at time $t_2 = t_1 + T$. The corresponding phase angle is given by

$$\omega t_2 = \omega\,(t_1 + T).$$

$$\omega t_2 = \omega t_1 + \omega \cdot \frac{S}{v_0 \left\{1 + \frac{V_1}{V_0} \sin \omega t_2\right\}^{1/2}}$$

$$= \omega t_1 + \frac{\omega S}{v_0}\left(1 + \frac{V_1}{V_0} \sin \omega t_1\right)^{-1/2}$$

$$= \omega t_1 + \frac{\omega S}{v_0}\left(1 - \frac{1}{2}\frac{V_1}{V_0} \sin \omega t_1\right)$$

$$\omega t_2 = \omega t_1 + \alpha\left(1 - \frac{1}{2}\frac{V_1}{V_0} \sin \omega t_1\right), \qquad \ldots(4)$$

where $\alpha = \dfrac{\omega S}{v_0}$.

This equation shows that an electron started from the buncher at a phase angle ωt_1 will come at the catcher at a phase angle ωt_2.

Power output and efficiency : To determine an expression for power output and efficiency, we assume that the buncher and catcher voltages are in time phase and, therefore, the catcher vol-

tage will be represented by $V_2 \sin \omega t$. During an electron's transit through the catcher grids, the electron gives up an amount of energy $W = -eV_2 \sin \omega t_2$. The negative sign signifies energy transfer from the electron to the field :

$$W = -eV_2 \sin \omega t_2,$$

$$W = -eV_2 \sin \left[\omega t_1 + \alpha \left(1 - \frac{V_1}{2V_0} \sin \omega t_1 \right) \right]. \quad \ldots(1)$$

For all electron entering the buncher grids during a complete cycle, *i.e.*, between $\omega t_1 = 0$ and $\omega t_1 = 2\pi$, we obtain the average energy per electron transferred to the catcher resonator as

$$W_{average} = \frac{1}{2\pi} \int_{\omega t_1 = 0}^{\omega t_1 = 2\pi} W \, d(\omega t_1)$$

$$= -\frac{eV_2}{2\pi} \int_0^{2\pi} \sin \left[\omega t_1 + \alpha \left(1 - \frac{V_1}{2V_0} \sin \omega t_1 \right) \right]. \, d(\omega t_1) \quad \ldots(2)$$

Solving the integral, we have $W_{average} = -eV_2 J_1(x) \sin \alpha$, ...(3)

where $x = \frac{\alpha V_2}{2V_0}$ and $J_1(x)$ is Bessel's function of first kind and first order.

Now let us consider the power relationship. Let N electrons be emitted from cathode per second; so the direct current emitted by the cathode is $I_0 = Ne$. ...(4)

$$\text{Power output} = -NeV_2 J_1(x) \sin \alpha$$

$$= -I_0 V_2 J_1(x) \sin \alpha. \quad \ldots(5)$$

Power supplied by by d c. potential source is

$$P_0 = I_0 V_0.$$

$$\text{Hence efficiency } \eta = \frac{\text{Power output}}{\text{Power supplied}} = -\frac{V_2}{V_0} J_1(x) \sin \alpha. \quad \ldots(6)$$

The graph plotted between different values of x as a function of $J_1(x)$ is of the form shown in fig. 37.

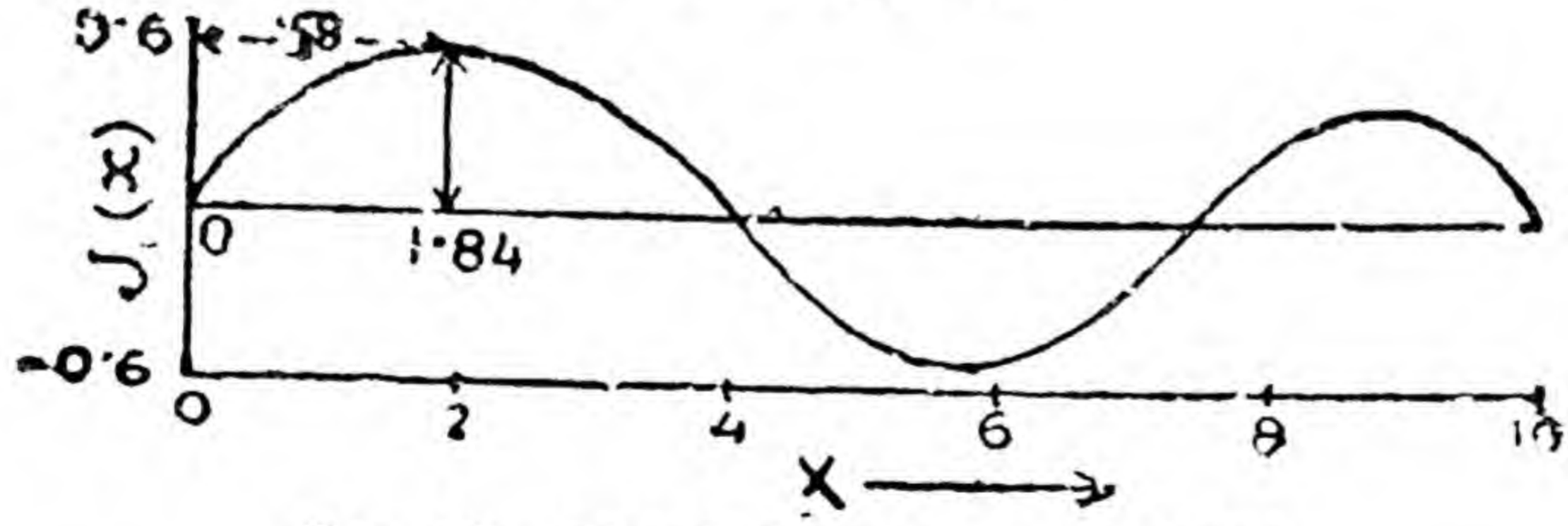


Fig. 37. Graph between x and $J_1(x)$.

Requirements of maximum output and maximum efficiency :

From equation (6), we write for maximum efficiency that

(*i*) the value of $\frac{V_2}{V_0}$ should be maximum,

(*ii*) $J_1(x)$ should be maximum. For $x = 1{\cdot}84$, $J_1(x) = 0{\cdot}58$, and is its maximum value.

(*iii*) $\sin\alpha=-1$, *i.e*, $\alpha=2n\pi-\pi/2$, where n in any positive integer, and

(*iv*) for power output I_0 and V_2 shculd be maximum.

For sustained oscillations $\sin\alpha=-1$ and hence

Power output $=I_0V_2\,J_1(x)$

Efficiency $=\frac{V_2}{V_0}J_1(x)$.

If $\frac{V_2}{V_0}=1$, the maximum theoretical efficiency would be 58%.

It should be remembered that power required for bunching of electrons is negligible. This is justified when transit angle of the electron is zero.

Further the criterion $\alpha=2\pi n-\pi/2$ is always satisfied It yields

$$\frac{\omega S}{v_0}=2n\pi-\pi/2$$

or
$$\frac{\omega^2S^2}{\frac{2eV_0}{m}}=(2\pi n-\pi/2)^2$$

or
$$V_0=\frac{m}{2e}\frac{\omega^2S^2}{(2n\pi-\pi/2)^2}$$

$$=0\cdot284\times10^{-11}\left(\frac{\omega S}{2\pi n-(\pi/2)}\right)^2 \qquad ...(7)$$

The equation shows that there are multiplicity of values of d.c. accelerating potential which will yield maximum power output, one for each integral values of n. The $n=1$ corresponds to highest value of d.c. accelerating potential and this is the reason that n is never kept equal to one. In other words it means that time of flight of electron through the drift space is generally more than one period of oscillations.

Now
$$x=\frac{\alpha V_1}{2V_0}. \qquad ...(8)$$

$J_1(x)$ is maximum when $x=1\cdot84$, hence

$$1\cdot84=\frac{V_1}{2V_0}(2n\pi-\pi/2)$$

or
$$\frac{V_1}{V_0}=\frac{3\cdot68}{2n\pi-\pi/2}. \qquad ...(9)$$

For each value of n, there is an optimum value of V_1/V_0 which gives maximum power output. It is also clear from equation (7) and equation (9) that as n increases both V_0 and the ratio V_1/V_0 decrease. The results are justified on the grounds that when n is high V_0 is small and v_0 is also small and consequently electron takes large time in travelling the drift space and therefore required less bunching voltage.

In klystron operation it is maintained that electron bunching takes place at catcher resonator. But it is also possible that electron several times bunch and debunch before reaching the catcher. This is shown in fig. (34) by several maxima and minima. $x=1{\cdot}84$ corresponds to single bunching and debunching.

Double resonator klystron in practice :

The output of the klystron furnishes sustained oscillations when feed back from output to input circuit is correct in phase and sufficient in degree. It is also required that two circuits are fully tuned. The degree of feed back is fixed by the size, position and orientation of loops at the feed back cable. The correct phasing is done by adjusting the field.

Klystron furnishes power output of the order of a fraction of watt to several hundred watts at frequencies of the order of thousand megacycles. The practical efficiency of double resonator klystron is much below its theoretical efficiency (58%). It is due to space charge effects causing debunching of electrons, collisions of electrons with grids, secondary emission at the grids, power consumed in bunching process, transit angle delay of electron in their passage through buncher and catcher systems and losses in resonators.

The frequency stability of electrical oscillations is a function of temperature of resonators and stability of supply voltage. As the resonator walls are metallic, their thermal coefficient of expansion is quite large. This makes the frequency of oscillations unstable. The thermal expansion in tuning mechanism is counteracted by expansion of resonator walls by proper choice of material. Carefully regulated power supply are used to stabilise the frequency of oscillations.

15·17. THE REFLEX KLYSTRON :

Fig. 38 shows the reflex klystron in which a single resonator acts both as buncher and catcher. After leaving the cathode and passing through the resonator grids, the electrons approach a plate called the reflector, which has a negative potential with respect to the cathode. The reflector electrode is ordinarily concave towards the resonator, so that there is a focussing effect which directs the electrons back towards the centre of the resonator grid. The gap between resonator grid and the reflector electrode is the drift space of the klystron.

If the resonator makes a transient oscillation, it modulates the velocity of the electrons as they cross its grid for the first time. The fast electrons penetrate the drift space more deeply than do the slow ones, and spend a longer time in it. Therefore they may return

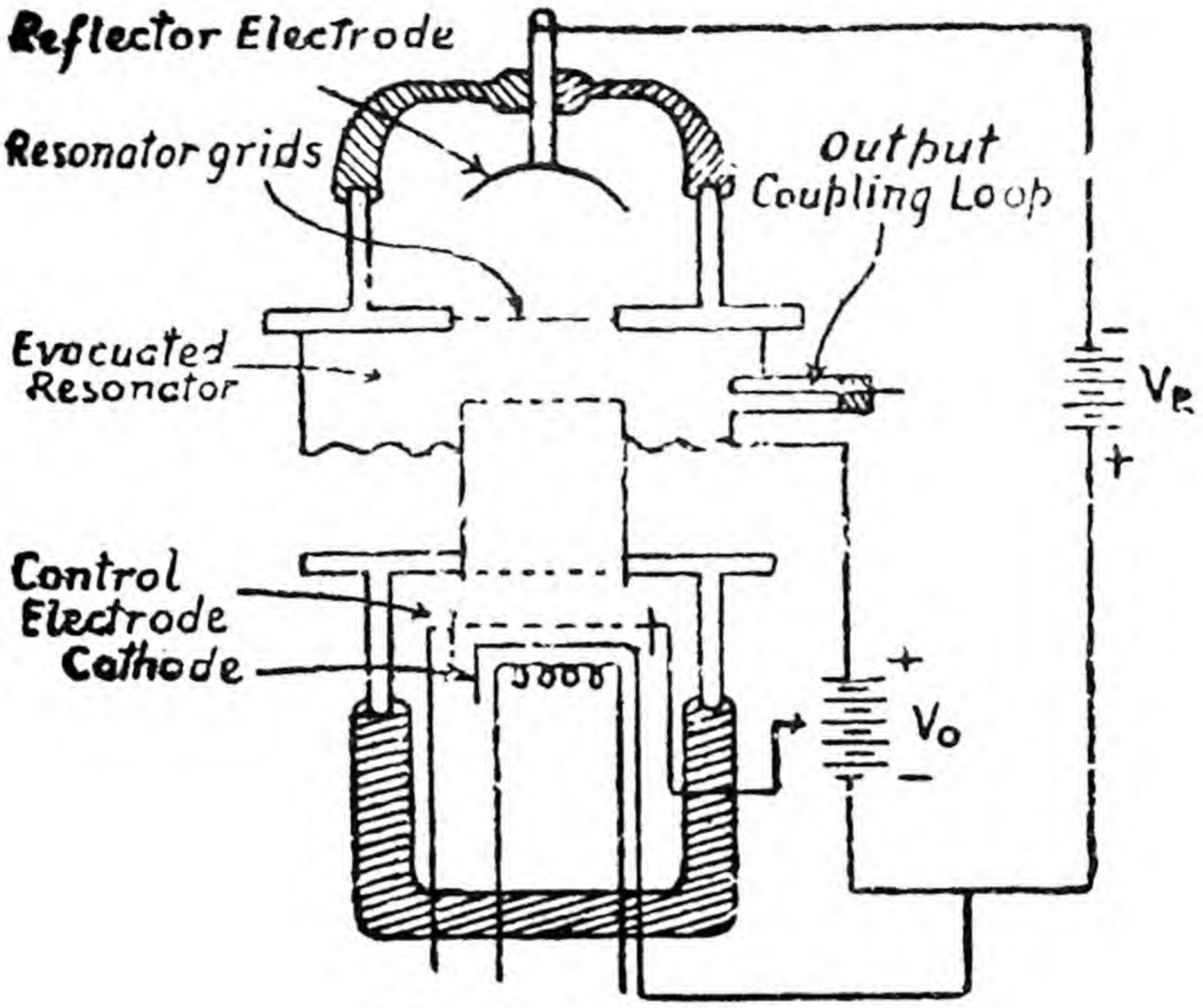


Fig. 38. Reflex klystron.

to the resonator girds at the same time as slow electrons which left the grids at a later instant than the faster ones. When this happens, bunches appear at the grids.

The type of bunching in reflex klystron is shown as reflector bunching. The Applegate diagram, illustrating the bunching action in reflex klystron, is shown in the figure 39. The space-time

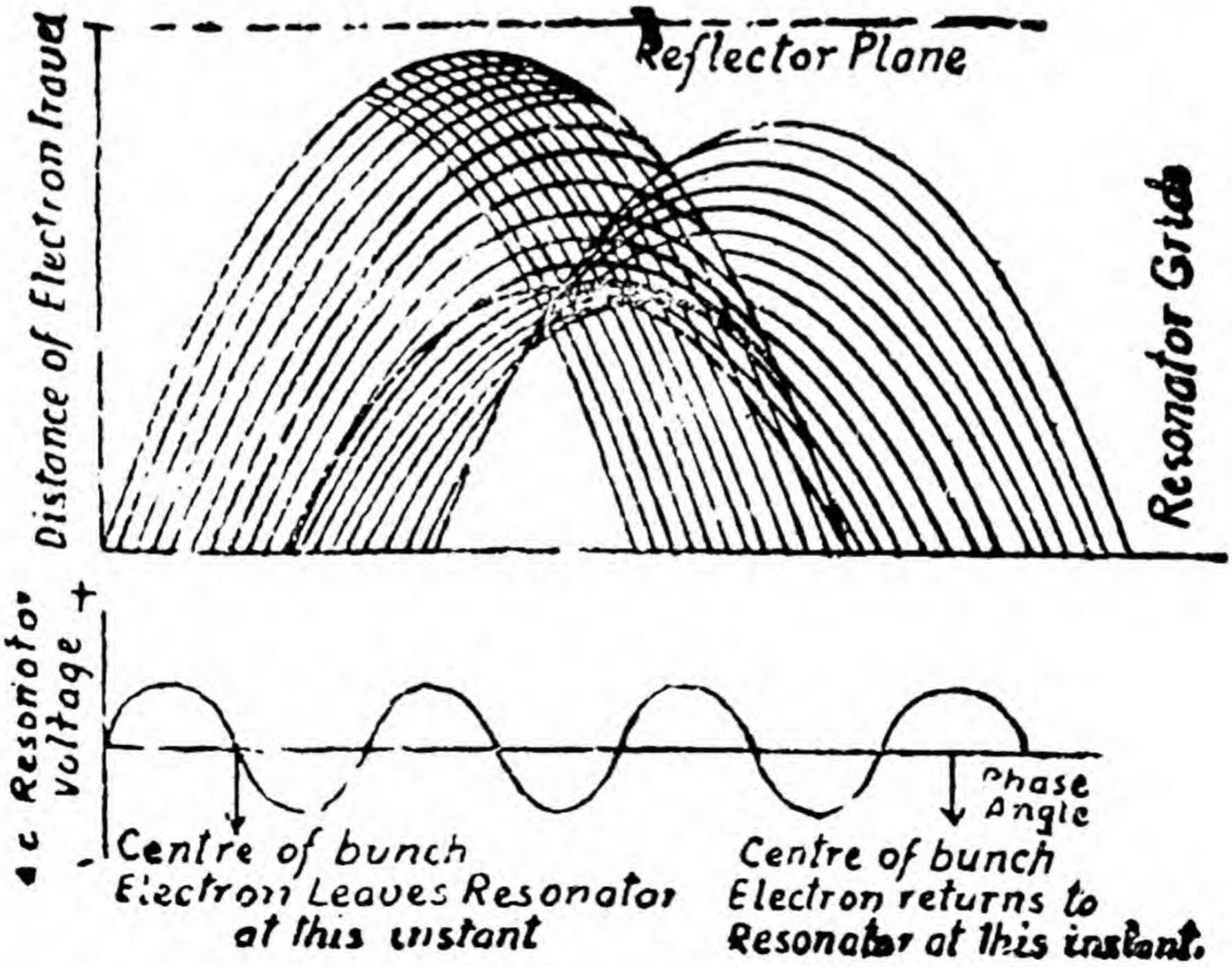


Fig. 39. Distance-time diagram of reflex klystron oscillator.

curve for each electron is a parabola. On their return journey, the electron bunches pass between the alternating field between the resonator grids during its retarding phase and thus energy is transferred from the electrons to the resonator alternating field. Oscillations may usually be obtained by merely adjusting the accelerating and reflector voltage. The frequency may be varied readily by changing the spacing between the grids of the resonator by means of suitable mechanism.

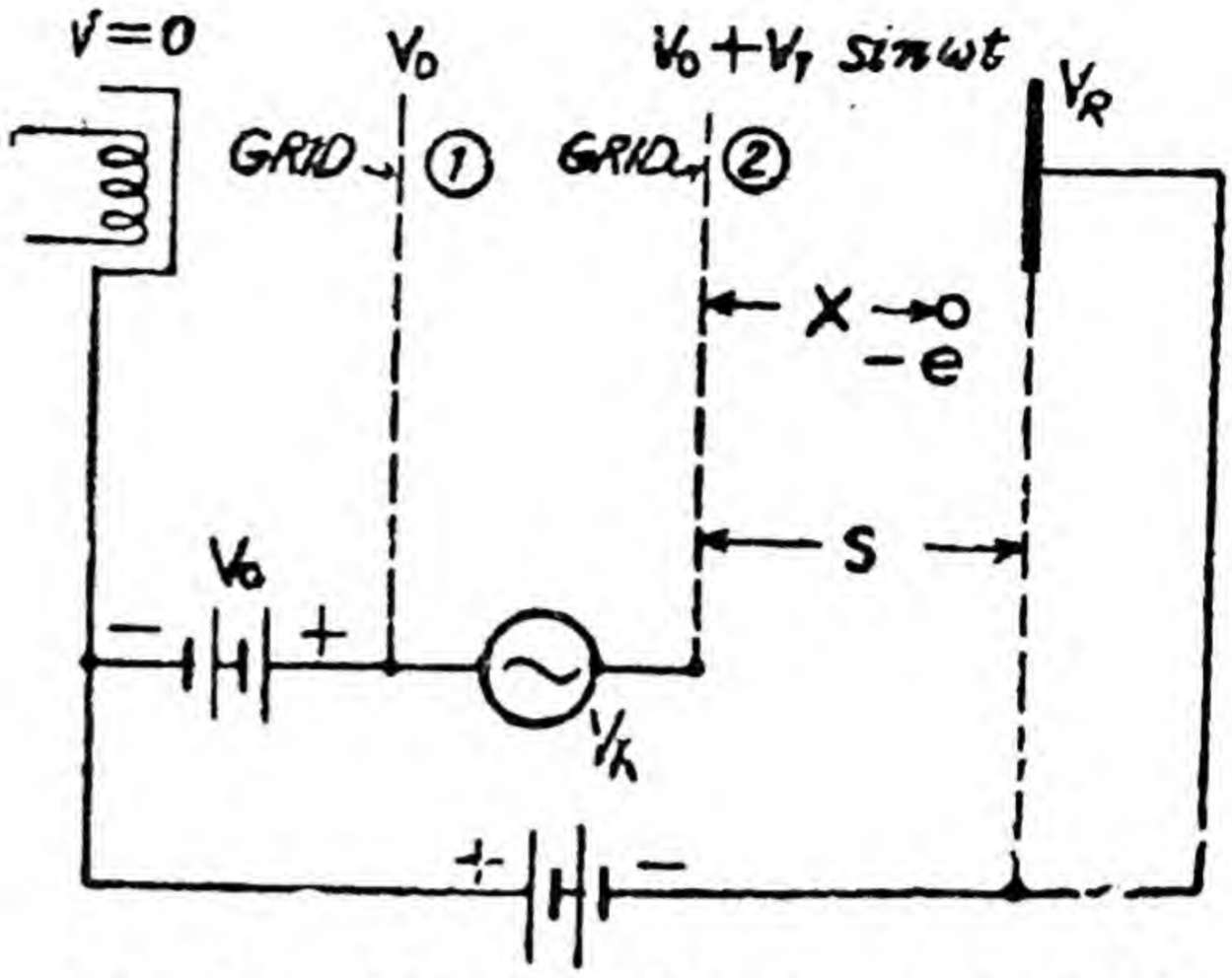


Fig. 40. Diagram for the analysis of the reflex klystron.

Comparison between the mechanism in double resonator klystron and reflex klystron :

1. In double resonator klystron all electrons travel same distance within the drift space and therefore for bunching process centre of the bunch lies at the zero of the bunching voltage when it changes from retardation to acceleration. In reflex klystron, the electrons with high velocity penetrate deep into the reflecting space and that is why the centre of bunch lies at the zero of the bunching potential when it changes from acceleration to deceleration.

2. In reflex klystron bunched electrons travel through the resonator grids in opposite direction to the direction of travel of bunched electrons in double resonator klystron. Therefore there is phase difference of π between the resonator voltage of the reflex klystron and catcher voltage of double resonator klystron at the time when energy is extracted from the electron.

3. In double resonator klystron bunching takes place in field free space and it is known as drift space bunching. In reflex klystron bunching takes place in the space in which field exists and it is known as reflector bunching.

4. The reflex klystron uses single resonator and therefore tuning mechanism and other adjustments are very simple.

Analysis of the reflex klystron : The analysis of reflex klystron is based on the same assumptions as adopted for the double resonator klystron.

Fig. 40 shows the potentials of various electrodes to the reflex klystron while oscillating. The cathode is at zero potential and a.c. potential difference $V_1 \sin \omega t$ is assumed to exist between the

grids. The potentials of grid (1) and grid (2) with respect to cathode assumed to be V_0 and $(V_0+V_1 \sin \omega t)$ respectively.

Referring to the figure, we find that the electron enter grid (1) with a velocity

$$\tfrac{1}{2}mv_0^2 = eV_0$$

$$v_0 = \sqrt{\left(\frac{2eV_0}{m}\right)}. \qquad \ldots(1)$$

and on the second grid,

$$\tfrac{1}{2}mv_1^2 = e\,(V_0+V_1 \sin \omega t),$$

$$v_1 = v_0 \sqrt{\left(1+\frac{V_1}{V_0} \sin\ \omega t\right)}. \qquad \ldots(2)$$

Potential difference between grid (2) and reflector electrode

$$= V_R - (V_0+V_1 \sin \omega t).$$

Hence we assume that $V_1 \sin \omega t$ is small in comparison to V_0 and hence neglected ; then

$$\text{potential difference} = V_R - V_0,$$

$$\text{electrical intensity } E = \frac{V_R - V_0}{S}. \qquad \ldots(3)$$

Force experienced by an electron at a distance X from the second grid

$$eE = m\ \frac{d^2X}{dt^2} \qquad \ldots(4)$$

From equations (3) and (4),

$$m\ \frac{d^2X}{dt^2} = eE = \frac{e\,(V_R-V_0)}{S},$$

$$\frac{d^2X}{dt^2} = \frac{e}{mS}\,(V_R-V_0). \qquad \ldots(5)$$

Integrating, we get

$$\frac{dX}{dt} = \frac{e}{mS}\,(V_R - V_0)\,t + C_1.$$

Applying the boundary conditions, *i.e.*, on grid (2), electron leaves with velocity v_1, when $t=t_1$,

$$C_1 = v_1 - \frac{e}{mS}\,(V_R-V_0)\,(t_1).$$

then

$$\frac{dX}{dt} = \frac{e}{mS}\,(V_R-V_0)\,(t-t_1) + v_1. \qquad \ldots(6)$$

Integrating again,

$$X = \frac{e}{mS}(V_R-V_0)\,\frac{(t-t_1)^2}{2} + v_1t + C_2,$$

when $X=0$, $t=t_1$.

$$0 = \frac{e}{mS}\,(V_R-V_0)\,\frac{(t-t_1)^2}{2} + v_1t_1 + C_2,$$

or

$$C_2 = -v_1t_1.$$

$$X=\frac{e}{mS}\,(V_R-V_0)\,\frac{(t-t_1)^2}{2}+v_1\,(t-t_1). \qquad ...(7)$$

The electron is decelerated as it approaches the negative reflector electrode, reverses its direction of travel, and starts back towards the resonator grids. Let t_2 represent the return arrival time of the electron, we may write equation (7) for the instant of arrival by substituting $t=t_2$ and $X=0$,

$$0=\frac{e}{mS}\,(V_R-V_0)\,\frac{(t_2-t_1)^2}{2}+v_1\,(t_2-t_1)$$

$$v_1=-\frac{e}{mS}\,(V_R-V_0)\,\frac{(t_2-t_1)}{2},$$

$$t_2-t_1=-\frac{2mSv_1}{e\,(V_R-V_0)}$$

$$t_2-t_1=-\frac{2mSv_0}{e\,(V_R-V_0)}\left(1+\frac{V_1}{V_0}\sin\omega t_1\right)^{1/2} \qquad ...(8)$$

The phase angle of the departure of the electron, with respect to buncher voltage, is ωt_1 and the phase angle of the arrival is ωt_2. Multiplying equation (8) by ω,

$$\omega t_2=\omega t_1-\frac{2mS\omega v_0}{e\,(V_R-V_0)}\left(1+\frac{V_1}{V_0}\sin\omega t_1\right)^{1/2}$$

$$\omega t_2=\omega t_1+\alpha'\left(1+\frac{V_1}{2V_0}\sin\omega t_1\right), \qquad ...(9)$$

where $\alpha'=-\dfrac{2mS\omega v_0}{e\,(V_R-V_0)}$

Power output and efficiency : The energy transferred by the electron is

$$W=-eV_1\sin\omega t_2$$

$$=--eV_1\sin\left[\omega t_1+\alpha'\left(1+\frac{V_1}{2V_0}\sin\omega t_1\right)\right]$$

$$=-eV_1J_1(x')\sin\alpha',$$

where $x'=-\dfrac{\alpha' V_i}{2V_0}$, $J_1(x')$ represents the Bessel function.

Suppose N electrons are emitted from the cathode per sec., so that the direct current from the cathode will be

$$I_0=Ne.$$

$$\text{Power output}=-NeV_1J_1(x')\sin\alpha'$$

$$=-I_0V_1J_1(x')\sin\alpha'. \qquad ..(10)$$

Power supplied by d.c. potential source $P_0=I_0V_0$.

$$\text{Efficiency } \eta=\frac{\text{power output}}{P_0}=--\frac{V_1}{V_0}J_1(x')\sin\alpha'. \qquad ...(11)$$

Referring now to the fig. '39), α' is the phase angle between zero of the alternating potential (when potential is changing from the acceleration to retardation) and positive peak of the resonator

voltage. Fig. (38) gives a value $2\pi n-\frac{1}{2}\pi$ for α', where n is any positive integer. It means that

$$\frac{-2mS\omega v_0}{e\,(V_R-V_0)}=2\pi n-\frac{\pi}{2}$$

or
$$\frac{V_0}{(V_R-V_0)}=\frac{(2\pi n-\frac{1}{2}\pi)^2\,e}{8mS^2\omega^2} \qquad ..(12)$$

Thus corresponding to different values of n, there are several values of $\frac{V_0}{(V_R-V_0)^2}$, which produce oscillations. The smallest value of n corresponds to highest value of reflecting voltage. For $\alpha'=2\pi n-\frac{1}{2}\pi$, the expressions for power output and efficiency may be written as

$$\text{power output}=\frac{2V_0I_0x'J_1(x')}{2\pi n-\frac{1}{2}\pi} \qquad ..(13)$$

$$\text{efficiency}=\frac{2x'J_1(x')}{2\pi n-\frac{1}{2}\pi}. \qquad ...(14)$$

The above equations show that as n decreases both power-output and efficiency increase. But the stability of the oscillator under variable load conditions become poor for small values of n. If a large value of n is used, the frequency of the oscillator is changed by large amount for small change in V_R. It is found that for stable performance the value of n between 1 to 5 is preferred, with $n=3$ most advantageous

The efficiency is maximum when the product $x'J_1(x')$ is maximum. The conditions for maximum efficiency and maximum power output cannot be simultaneusly satisfied because output voltage is also a bunching voltage and maximum bunching voltage produces broad bunches and there by reduces the efficiency. Therefore, a compromise is made between the power output and efficiency.

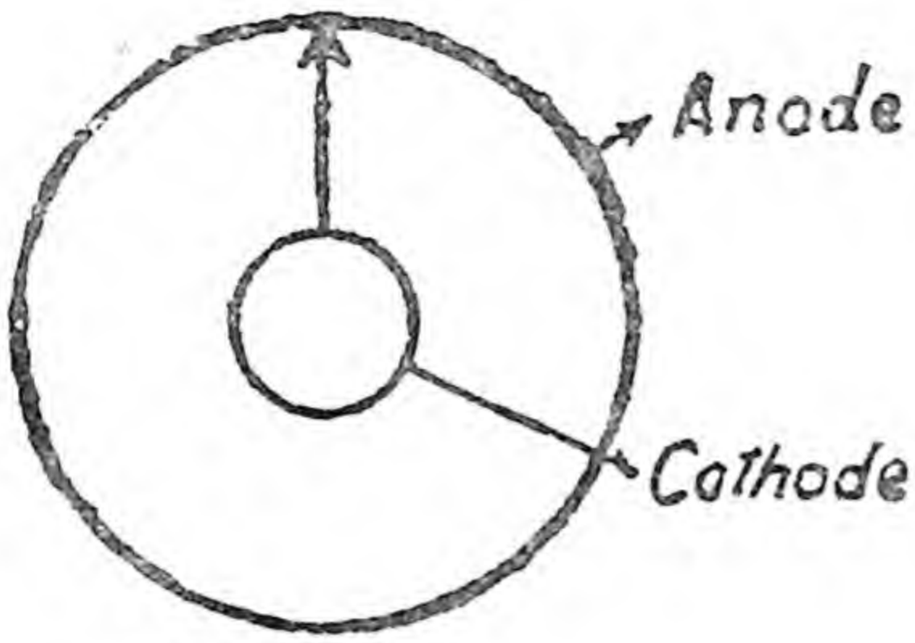
The maximum theoretical efficiency of reflex klystron lies between 20 to 30 percent Though it has low efficiency yet it has the advantages of having

(*i*) single resonator,
(*ii*) simplicity and adjustment, and
(*iii*) compactness.

15 18. MAGNETRON :

Valves in which the electrons move in a magnetic field, as well as in the electric field between the anode and cathode, are called magnetrons. Usually the electric and magnetic fields are perpendicular, and the arrangement is known as plane magnetron. Cylindrical structures are commonly employed. A cylindrical magnetron consists of a cylindrical anode along the axis of which is located the filament.

A high d.c. potential is applied between cathode and anode, setting up a radial electric field. An axial magnetic field is provided by either permanent magnet or an electromagnet.



(*a*) No magnetic field
(*b*) Weak field.
(*c*) Critical field.
(*d*) Strong field.

Fig. 41. Effect of magnetic-field on electron motion.

When no magnetic field is applied, the electrons will travel from cathode to anode in radial lines. With a weak magnetic field, the electrons are subjected to a deflecting force acting at right angles both to the magnetic field and to the direction of motion of the electrons at any instant and since a moving electron is just like a current carrying conductor, a deflecting force, which is at right angles to the direction of movement of the electron must produce circular motion. For this reason electrons travel in circular arc between the anode and cathode, the radius of curvature of which decreases with the increase in magnetic field strength. For a given value of d.c anode potential, there is a critical value of magnetic field strength which causes the electrons to just graze the anode. This value of magnetic field strength is known as the *cut-off value*. If the magnetic field strength exceeds the cut-off value, the curvature of the electron path is such that the electrons miss the anode and spiral back to the cathode.

If we plot a graph between anode current and strength of magnetic field, it is of the shape shown in figure 41 (*e*). The

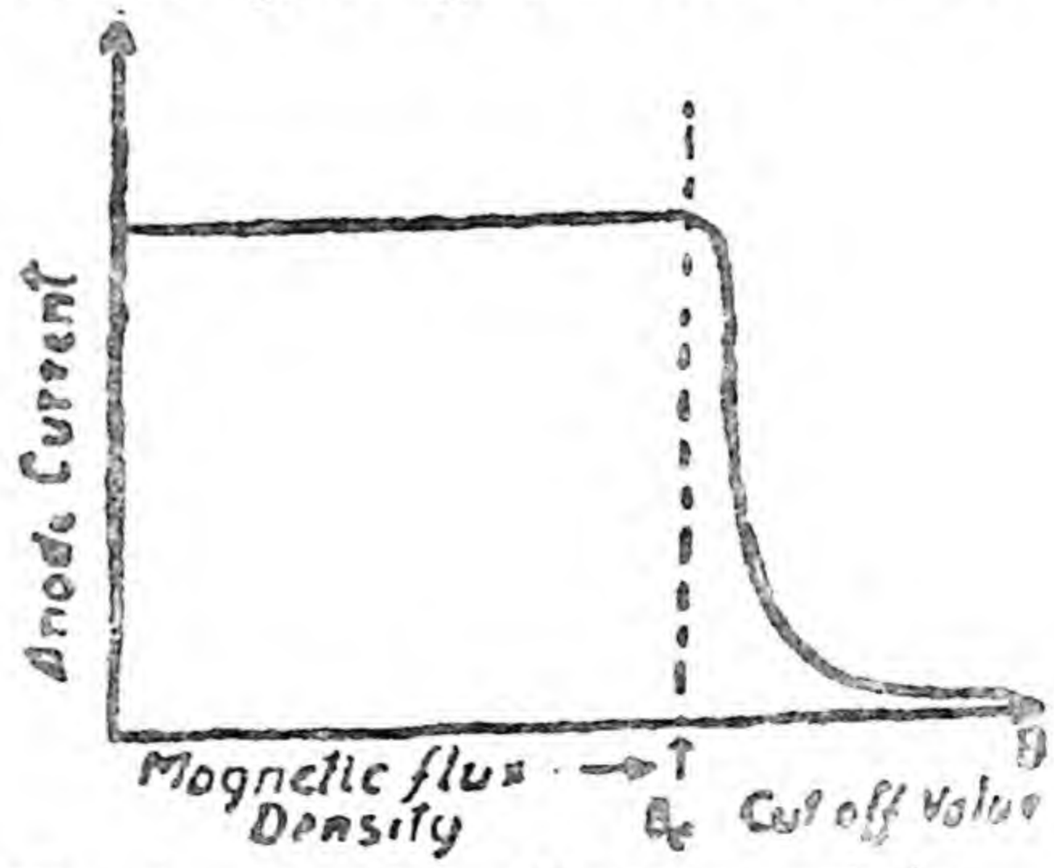


Fig. 4 (*e*). Cut-off characteristics.

anode current experiences as sharp cut-off as the ciritia! value of magnetic field is reached.

Mathematical theory of cylindrical magnetron (*critical value of magnetic field*) : Let the radius of cylindrical anode be r_a, and r_k be radius of cathode Further $r_k \ll r_a$

If F is the electrostatic force and f is magnetic force on the electron ejected from the surface of the cathode, then

$$F = -eE,$$

where E is the intensity of electric field and is given by

$$E = -\frac{dV}{dr}$$

$$\therefore \quad F = e\frac{dV}{dr}$$

and

$$f = Bev,$$

where v is the velocity of the electron and B is the magnetic flux.

If the velocity v is resolved into components v_r and v_θ, along and perpendicular to the radius vector r, then the components of f will be

$$f_r = Bev_\theta,$$
$$f_\theta = Bev_r.$$

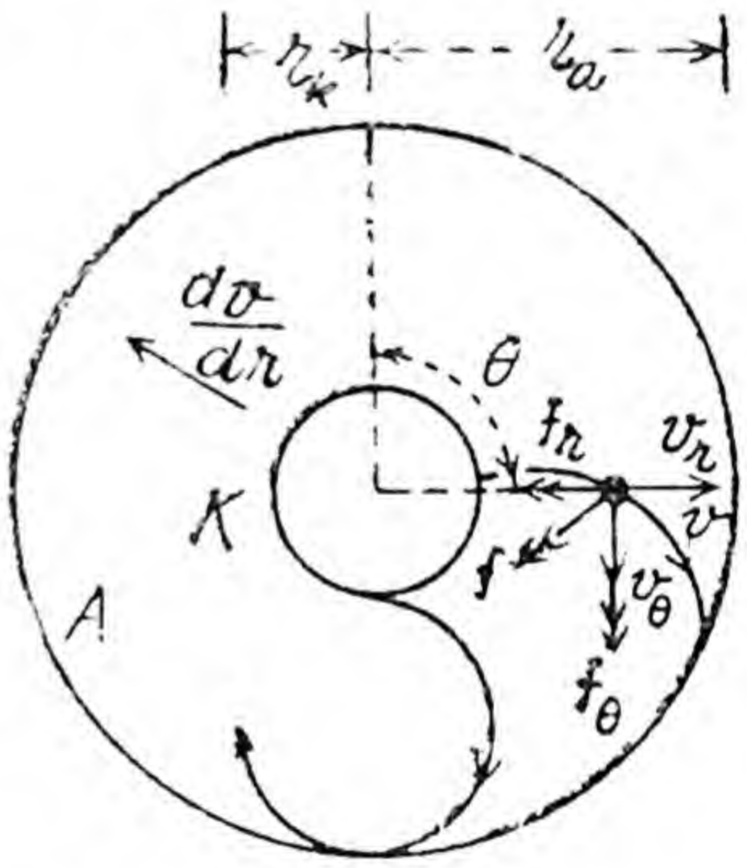


Fig. 42. Illustrating the theory of cylindrical margetron.

Let ω be the angular velocity and the radius vector r makes angle θ with an arbitrary line in the azimulthal plane, so that

$$\omega = \frac{d\theta}{dt} = \frac{v_\theta}{r} \text{ since } v_\theta = r\frac{d\theta}{dt}.$$

The equation of motion becomes

$$\frac{d}{dt}\left(m\frac{dr}{dt}\right) = F - f_r. \qquad \ldots(1)$$

Equation of motion in the azimuthal plane is obtained by equating the moment of the impressed force to the rate of change of angular momentum, *i.e.*

$$rf_\theta = \frac{d}{dt}(mv_\theta . r)$$

or

$$rBe\frac{dr}{dt} = \frac{d}{dt}(mr^2\omega).$$

Integrating, we get

$$\int Ber\,dr = \int d(mr^2\omega),$$

$$Be\frac{r^2}{2} = mr^2\omega + C.$$

To find the value of C, we note that $\omega=0$ at the moment the electron leaves cathode, when $r=r_k$.

$$C=\frac{Ber_k^2}{2},$$

$$\frac{Be}{2}(r^2-r_k^2)=mr^2\omega,$$

$$\omega=\frac{Be}{2m}\left(1-\frac{r_k^2}{r^2}\right). \qquad ...(2)$$

If we ignore the small velocities with which the electrons leave the cathode, then the velocity v of an electron, at a point where the potential is V, is given by

$$\tfrac{1}{2}mv^2=eV,$$

$$v=\sqrt{\left(\frac{2eV}{m}-\right)}.$$

When the current is just cut off, the electron is grazing the anode, and we have

$$V=V_a,$$

$$v_r=0,$$

$$v_\theta=v=\sqrt{\left(\frac{2eV_a}{m}\right)},$$

$$r=r_a$$

So $$\omega=\frac{v_\theta}{r_a}=\frac{1}{r_a}\sqrt{\left(\frac{2e}{m}V_a\right)} \qquad ...(3)$$

If B_c be cut off value (critical value) of magnetic field, then

$$B_c.\frac{e}{2m}\left(1-\frac{r_k^2}{r_a^2}\right)=\frac{1}{r_a}\left(\frac{2e}{m}V_a\right)^{1/2},$$

$$B_c=\sqrt{\left(\frac{8V_a}{e/m}\right)}.\frac{1}{r_a\left(1-\frac{r_k^2}{r_a^2}\right)}. \qquad ...(4)$$

Equation (4) shows that the value of B_c is independent of the potential distribution so that the presence of any other electrode or of space charge between the cathode and anode should not eflect the cut off conditions.

Magnetron as a high frequency oscillator : Usually two types of magnetrons, depending upon construction, are used :

(*i*) Split anode magnetron.

(*ii*) Cavity magnetron.

We shall discuss the operation of both these types of magnetrons in brief as follows :

(A) Split anode magnetron : In this type of magnetron the anode is split in two segments, each being at the same positive potential E_b with respect to cathode. Across these two anodes, an oscillatory circuit is connected as shown in fig. 43. By means of centre tap at the oscillatory coil, high tension E_{bb} is applied to

both the segments. As the the radius r_k of cathode is small compared to r_a, the anode radius, the electrostatic field rises steeply just beyond the cathode surface and then becomes almost constant over the remaining distance to the anode. The electrons flow from filament to the anode under the combined effect of electrostatic and magnetic fields. The magnetic flux density has about twice

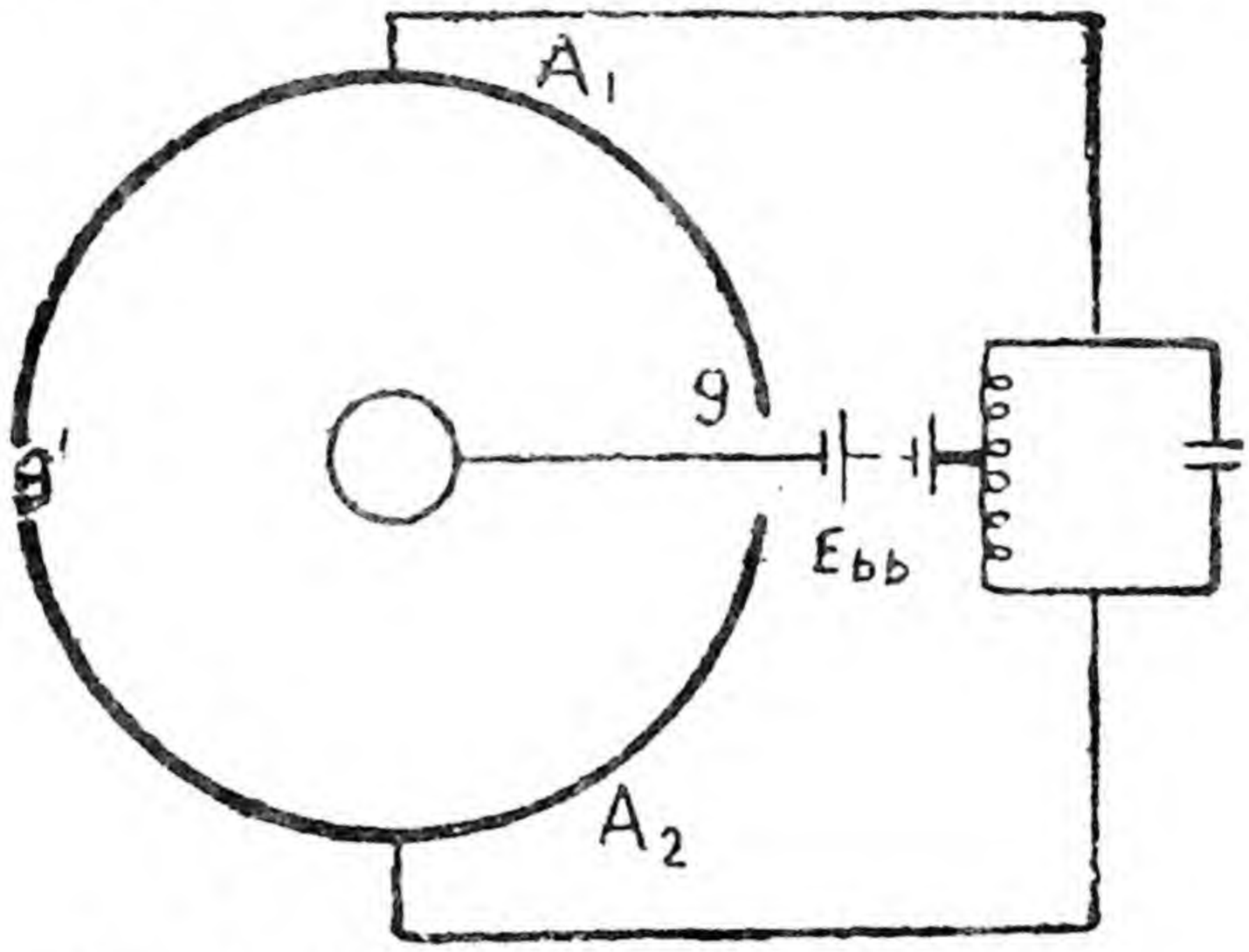


Fig. 43. Split anode magnetron circuit.

the critical value. The existence of a transient potential difference across the tank circuit will raise the potential of one anode segment and lower that of the other by equal amounts. The resulting distortion is shown by the continuous curved lines in space between anode segments in figure 44 (*a*) and 44 (*b*). It appears that the space between the electrodes is divided into two regions, of high and low potentials, and that the high-potential region invades the space within the low potential anode.

In figure 44 (*a*) the path of an electron, leaving the cathode on high potential side is shown. The electron travels in curved path which carries it out of that region. On entering low potential region, electron travels more slowly and its path has a smaller

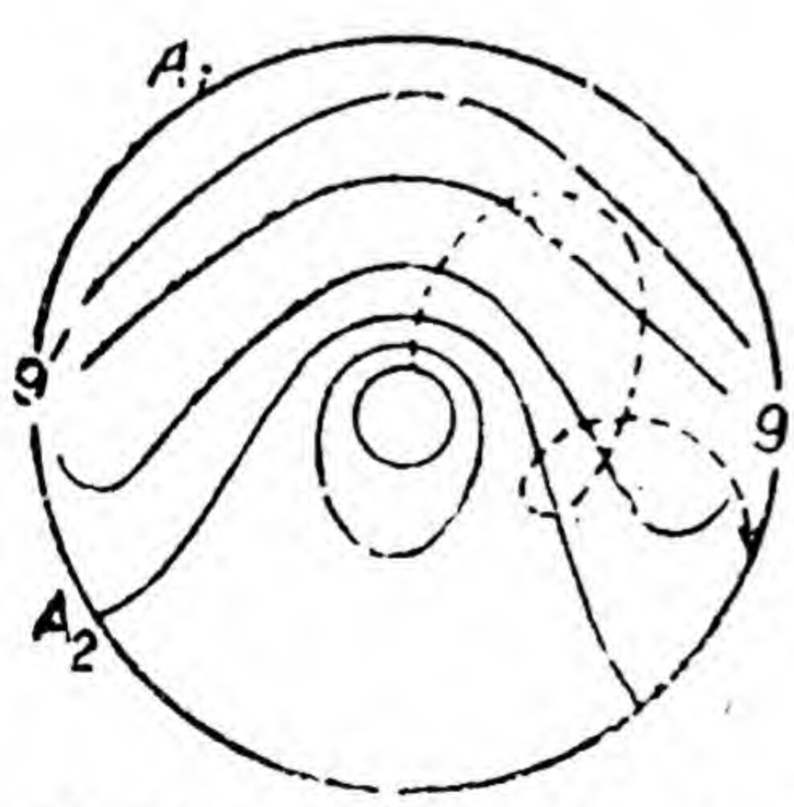


Fig. 44. (a) Electrons leaving cathode on high potential side.

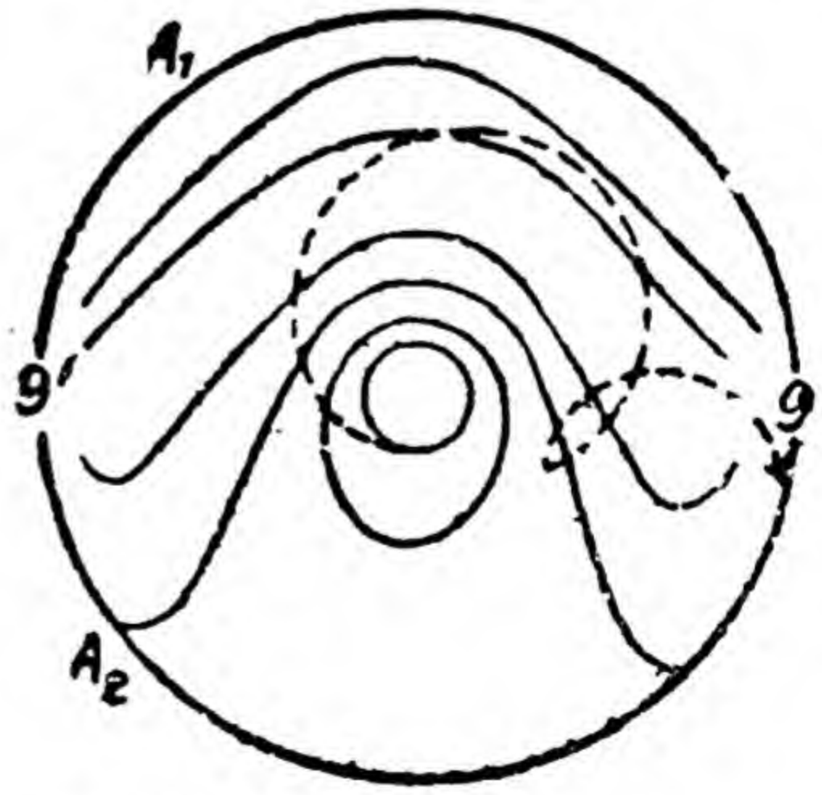


Fig 44. (b) Electron leaving cathode on low-potential side.

radius. Thus when it returns to high potential region and the radius of its path increases again, its penetration becomes smaller. The combined effect of magnetic and electrostatic fields on electron, in its each revolution, is to deflect it from high potential to low potential region and thus ultimately it strikes the low potential segment of the anode.

The behaviour of an electron leaving the cathode on the low potential side is shown in fig. 44 (b). The electron after swinging round into both potential regions ultimately strikes the low potential anode segment. Thus we conclude that when the flux density in a magnetron is greater than cut off value, and a potential difference is applied between the anode segments, an electron current flows to the segment at the lower potential.

With a change in the polarity of the transient potential, the upper segment will become of lower potential and now the electrons will strike the upper segment. Each anode thus delivers energy to the tank circuit during alternate half cycles in the same manner as happens in push-pull arrangement If we plot volt-ampere characteristic of such a valve, we find that in the vicinity of cut-off the characteristic gives a negative slope similar to dynatron. The property of its static characteristic can be utilized to produce oscillations in the tuned circuit.

There are three popular modes of operation of the split anode magnetron as described below :

(i) Negative resistance or Dynatron mode : Near the cut-off point, the volt-ampere characteristic gives a negative slope just as in the case of dynatron and therefore it is operated at this point. The tank circuit governs the frequency of such oscillation. The negative resistance of the valve in the dynatron region is a property of its static characteristic and is therefore independent of the frequency at which it is used.

At very high frequencies, the time taken by an electron to cross the space between the electrodes connected to the oscillatory circuit is comparable with the period of oscillation. Under this condition the dynatron action fails and the ratio of electron transit time τ to the period of oscillation T is the factor which determines the efficiency of the oscillator. For a ratio equal to or greater than than 0·20, oscillation ceases.

Magnetron is more efficient than a secondary emission dynatron because current flows in pulses to each segment having a low potential.

(ii) Transit Time Mode : A magnetron can maintain oscillations above its dynatron limits provided the external circuit is tuned to frequency which depends upon the size and working conditions of the valve. Oscillations take place when magnetic flux density is slightly greater than its critical value and the measured

wavelengths suggest that the period of oscillation is equal to the time taken by an electron to go from the cathode, towards the anode, and back to the cathode. For the calculation of this time, we assume that both segments of the anode are at the same potential E_b. The electrons are then accelerated over very short distance by intense electrostatic field just outside the filament; beyond this distance the speed of electrons may be taken as constant and having a magnitude

$$v = \sqrt{\left(\frac{2e}{m} E_b\right)}$$

The paths of electron are therefore circles of radius

$$r = \frac{m}{eB}$$

and their period of revolution is

$$\tau = \frac{2\pi r}{v} = \frac{2\pi m}{eB}.$$

Now we imagine the existence of a transient oscillation in the tuned circuit and the segments A_1 and A_2 as positive and negative according to instantaneous polarity of the tank circuit (ignoring the steady anode potential E_b component). Then an electron crossing gap g between the segments will be accelerated if segment A_2 is positive and retarded if A_1 is positive, If the electron is accelerated at the gap the radius of its path increases and it strikes the segment A_2, while, on the other hand, if the electron is retarded the radius of its path decreases and it fails to reach the either segment.

An electron which is accelerated by the field across the gap induces a current which takes energy from the tank circuit and when strikes with positive segment it dissipates its kinetic energy as heat.

An electron which is retarded at the gap loses kinetic energy and induces current which delivers energy to the tank circuit. As a result of losing kinetic energy, it fails to reach the cathode on its return journey and it turns towards the anode once more. If, when the electron approaches the gap 'g' again, the tank circuit has oscillated through one cycle, then the electron is again retarded at the gap, and the whole proocess is repeated. Thus in successive cycles, the electron loses its kinetic energy and approaches anode less closely and finally it comes to rest somewhere between anode and cathode. Since it would now start to move towards the more positive segment and damp the oscillations of tank circuit, therefore this electron must be removed from the space between the electrodes.

The external circuit gains more energy from electrons which are retarded at the gap, than it loses to electrons which are accelerated at the gap. The valve can therefore maintain oscillations in

the tuned circuit connected across its anode segments. The energy of oscillations comes from the high tension supply through the kinetic energy of retarding electrons.

(iii) Resonance Mode : The transition from dynatron to transit time oscillation occurs over a frequency range in which the period of oscillation is several times the electron's time of revolution. If the magnetic flux density is made greater than the value in the dynatron mode, powerful oscillations can be maintained within this frequency range. The frequency of these oscillations depends upon flux density B and mean anode potential E_b and their wavelength obey the empirical relation

$$\lambda \propto \frac{r_a^2 B}{E}.$$

The mechanism for maintaining the oscillations is that the half period of oscillation should be an integral multiple of the time of revolution of an electron in the valve. The action is similar to that in the dynatron mode.

Resonance oscillations have been assigned to depend upon the production of a form of negative resistance in which the electrons may have made several spiral orbits within the valve, being deflected in the same direction by successive anode segments before eventually striking on an anode segment at a low instantaneous potential. If any such process is occurring, it can be easily understood that the frequency of strongly sustained oscillation is ultimately connected with the period of oscillation of a single electron.

The required condition for the sustained oscillations in the resonance mode is that electron should complete whole number of orbits between successive gaps during an interval of time corresponding to only half the periodic time of the generated oscillation.

(B) Cavity Magnetrons : All magnetrons possess in common a cathode, an anode and an output coupling device.

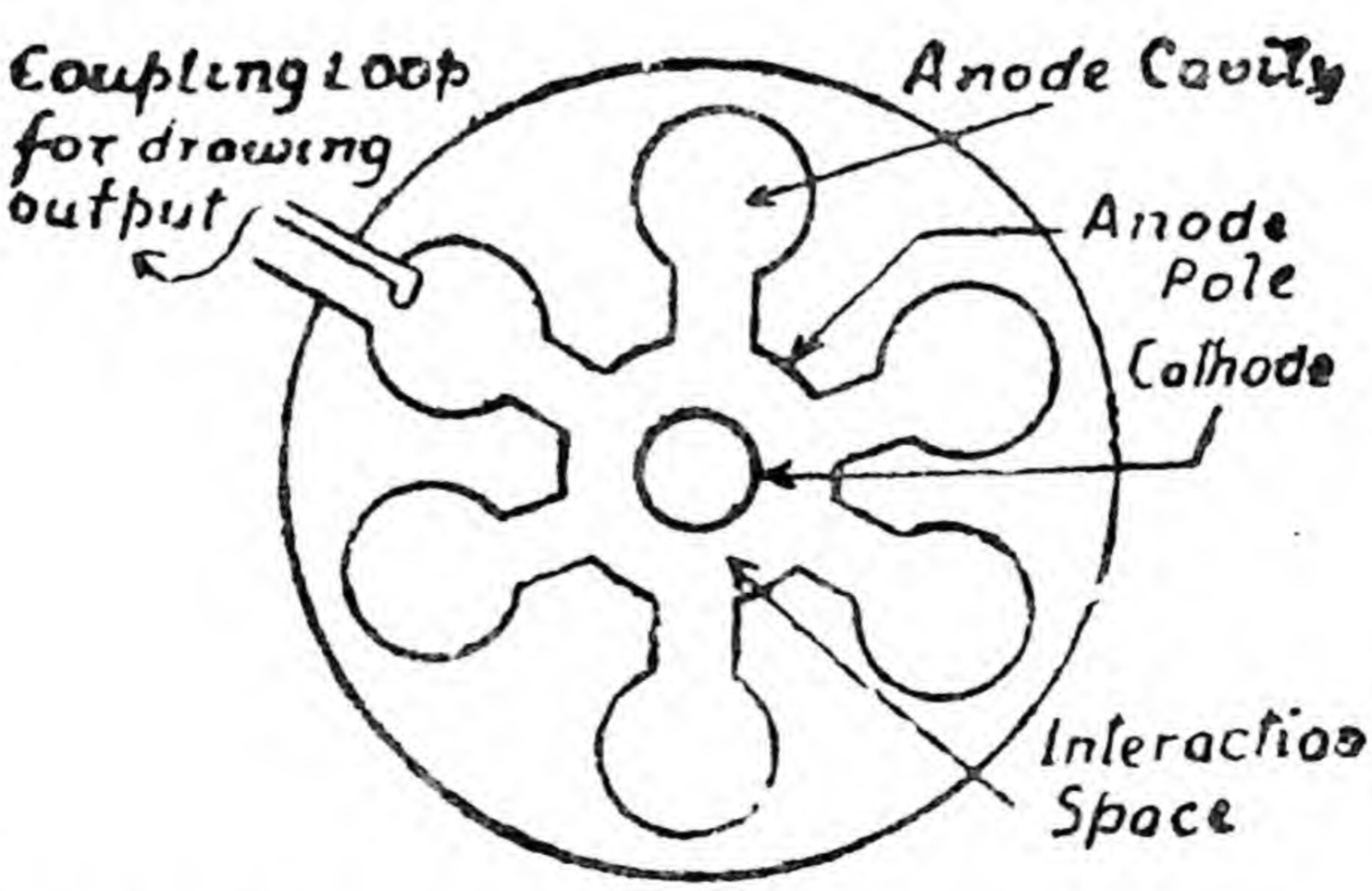


Fig. 45. Hole and slot type multicavity magnetron.

They may also have tuning system, mode suppressors and end plates. The multicavity magnetron exists in various forms, *e. g.*, hole and slot type, slot type, vane type, rising sun type. The structural form of cavity magnetron hole and slot type) is shown in figure 45. This consists of cylindrical cathode surrounded by an anode structure that possesses cavities opening into the cathode anode or interaction space by means of slots. Output power is withdrawn by means of a coupling loop. The multicavity magnetrons are put in metallic house. Glass is used at the points where high voltage leads come out of the enclosure.

Resonant Properties of Multicavity Magnetron : The anode cavities, together with spaces at the top and the bottom, of the anode block constitute the resonant system of the osciliator. There is a capacity across each gap in parallel with the inductance formed by the inner surface of the circular hole. The anode-cathode arrangement of figure 45 is expected to have the equivalent circuit shown in figure 46. The capacity C_2 represents the capacity between a pole face and the cathode. The capacity C_1 represents the capacity between two adjacent pole faces. The inductance L_1 is the inductance of the inner surface of circuit hole. The above equivalent circuit is poor as we have ignored the large mutual inductance between adjacent anode spaces. Let the circuit be developed by upwrapping the structure to give the arrangement as in figure 47. This is seen to be a low pass filter. Such type of filter will have a pass band in which the attenuation is substantially zero and in which there is a phase shift per section which increases uniformly from zero at zero frequency to π radians per section at cut-off. The standing waves will exist in the circular arrangement when the total phase

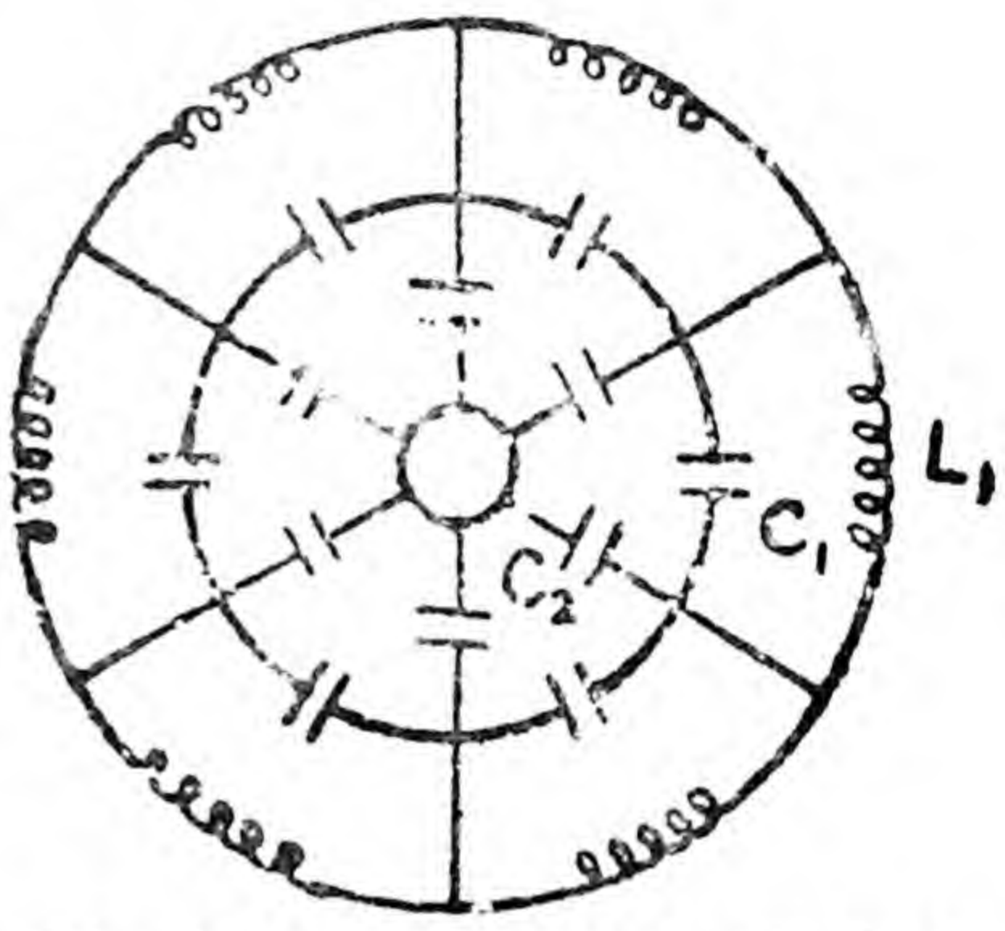


Fig. 46. Approximate equivalent circuit of fig. 45.

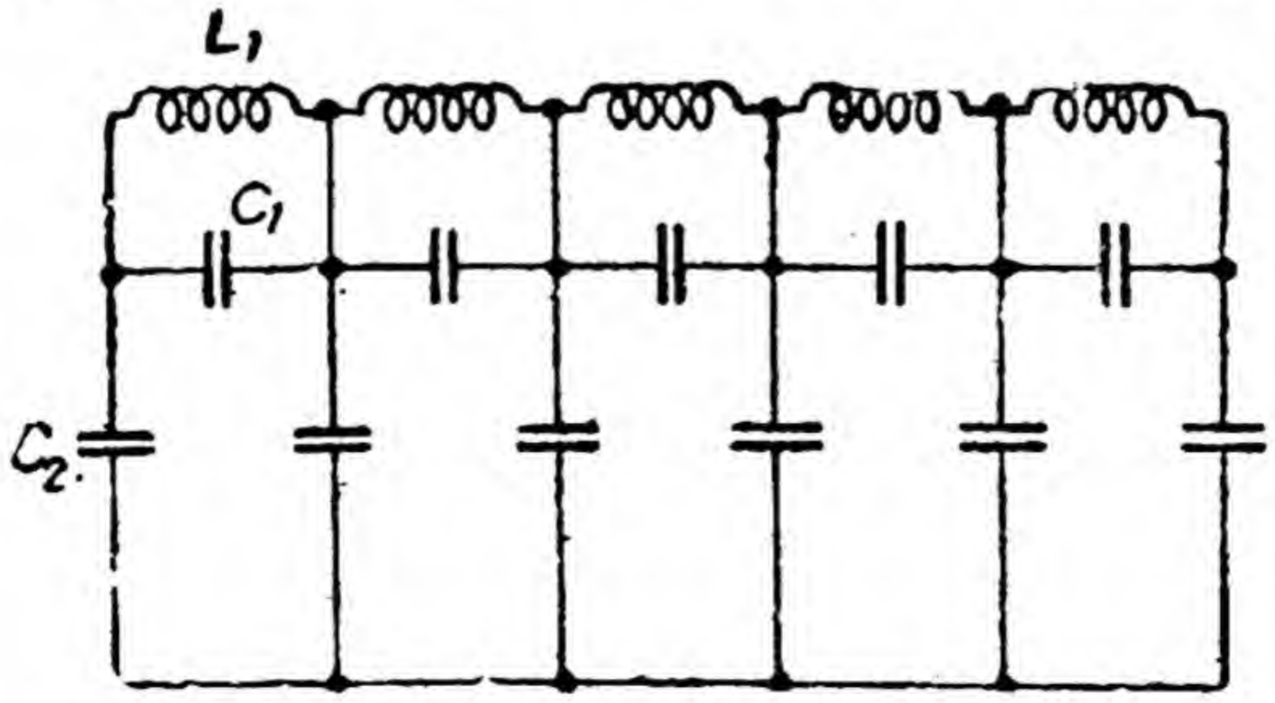


Fig. 47. **Developed form of the equivalent circuit of the magnetron.**

shift along N sections is any integral multiple of 2π. The two waves of equal amplitude and travelling in opposite directions will form actual resonant field. This happen when

$$\beta = \frac{2n\pi}{N},$$

where β represents phase shift per section in radians, n is mode number, *i e.* number of cycles in a travelling wave around the cathode, and N represents number of pole faces.

When $n = N/2$, the value of β is π which means that phase shift per section is π radians and the field reverses at adjacent gap. This mode of multicavity magnetron is generally used and is known as π mode.

This phase shift function of the circuit of figure 47 can be evaluated by applying Campbell's formula. In this passband this has the form

$$\cos\beta = 1 + \frac{Z_1}{2Z_2},$$

where Z_1 = total series impedance per section,

Z_2 = total shunt impedance per section.

In this case $$Z_1 = \frac{L_1/C_1}{j\omega L_1 + \frac{1}{j\omega C_1}},$$

which reduces to $$Z_1 = \frac{j\omega L_1}{1 - \omega^2/\omega_1{}^2},$$

where ω_1 is the angular resonant frequency of the parallel L_1C_1 combination equal to $\frac{1}{\sqrt{(L_1C_1)}}$ in magnitude.

The shunt impedance is given by

$$Z_2 = \frac{1}{j\omega C_2}$$

Putting the value of Z_1 and Z_2 in the expression for cos β, we have $$\cos\beta = 1 - \frac{\frac{1}{2}\omega^2 C_2{}^2/\omega_1{}^2 C_1}{1 - \omega^2/\omega_1{}^2},$$

which gives $$\left(\frac{\omega}{\omega_1}\right) = \frac{1}{\frac{C_2}{2C_1\left[1 - \cos\left(\frac{2\pi n}{N}\right)\right]} + 1},$$

where ω is now the angular resonant frequency corresponding to a given value of n and N.

The resonant frequencies for the assumed circuit will have the form as shown in fig. 48.

The figure shows that the frequency of the π mode is not very

different from the next resonant frequency. The situation is unsuitable when frequency separation is small.

For a precise analysis the mutual inductance between adjacent slots should be accounted.

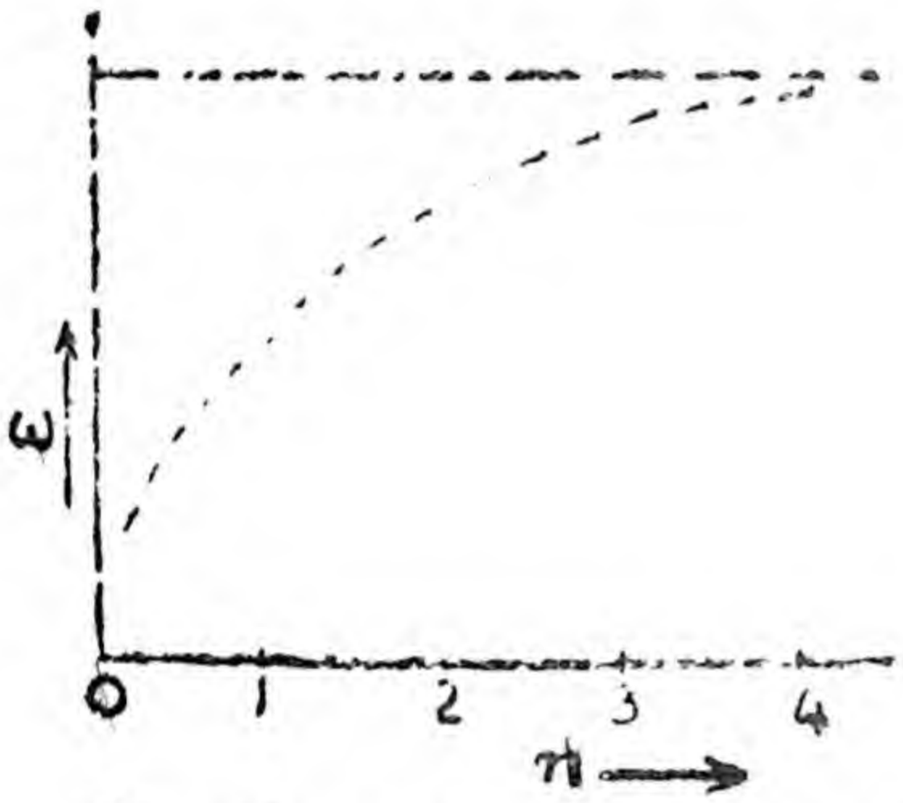


Fig. 48. Resonant frequencies of the circuit of fig. 47.

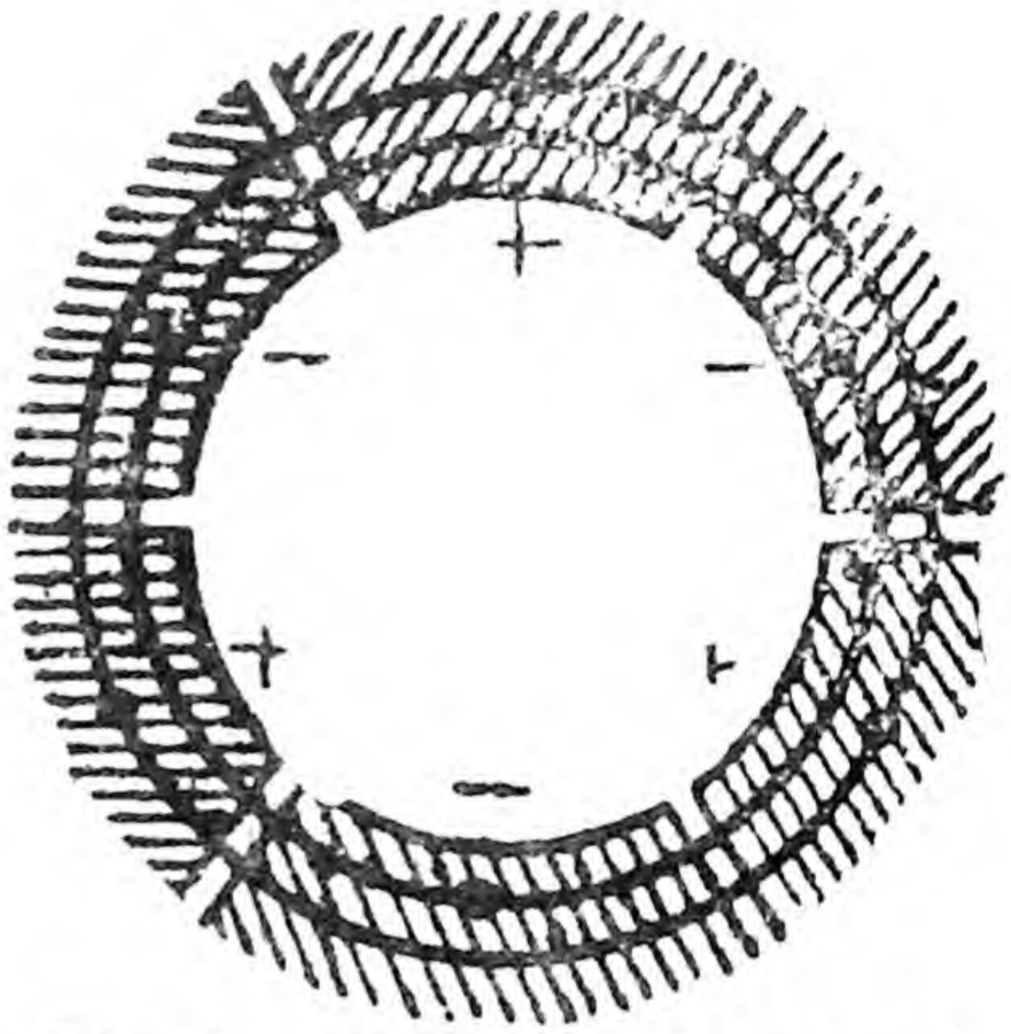


Fig. 49 Representation of ring strapping.

In order to overcome the situation of low frequency separation, a method of "strapping" is employed. In this arrangement two rings are run around the pole tips. Each ring is connected to alternate pole tips, one ring being connected to all the odd-numbered pole tips, and the other to all even numbered pole-tips. In the π mode of resonance, alternate poles are 180° out of time phase with each other. As a result, the straps will be 180° degrees out of phase with each other, and thus the capacity between the straps will be added in parallel with the capacity C_1 in the equivalent circuit. This lowers the resonant frequency of the π mode. Because of symmetry and phasing no current will flow in straps at the resonant frequency of the π mode. For other modes, the phase shift between adjacent poles is not 180 degrees, and so current will flow in the straps, the effect of which is to put more inductance in parallel with the inductances and thus rises the frequency of the adjacent resonances of slots. Both the capacity and inductances thus combine to increase the frequency separation between the π mode resonance frequency and the adjacent frequency.

Control of resonant frequency of the anode system : One way of controlling the resonant frequency of the anode system is to couple a tunable high Q resonant cavity to one of the anode cavities. Thus a sufficient reactance is coupled into the magnetron cavity system which modifies the resonant frequency by a moderate amount.

Another way consists in employing a 'C ring' as shown in figure 50. This adds capacitance between the straps of a ring

strapped magnetron and thereby lowers the resonant frequency of the π mode by an amount depending upon position of *C* ring.

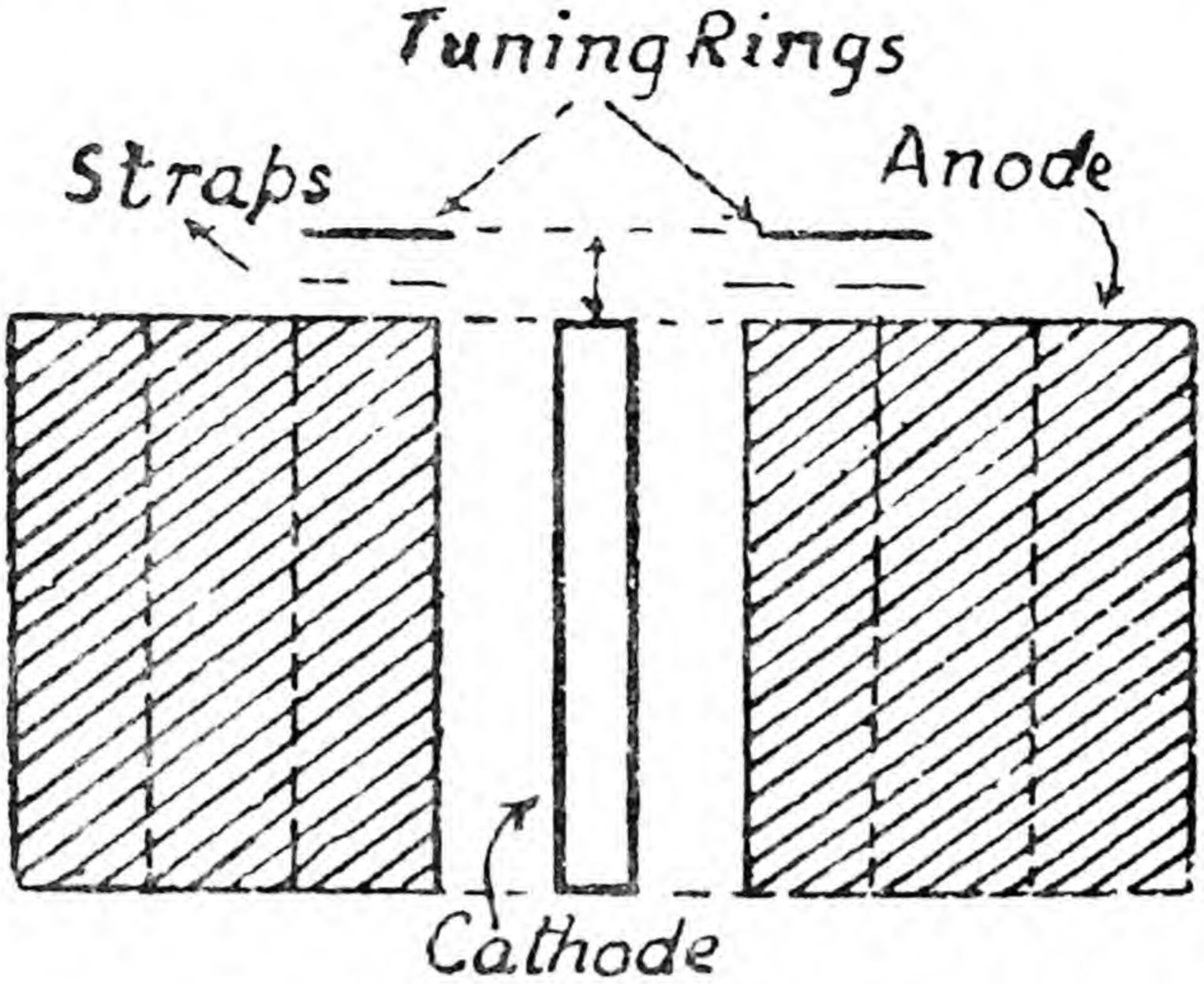


Fig. 50. C ring arrangement.

Mechanism by which the oscillations are generated : The nature of the field associated with the cavities is such that the alternating magnetic flux lines pass through the cavities parallel to the cathode axis, while alternating electric fields are confined largely to the slots and the interaction space region where the cavities open, and lie in planes perpendicular to the axis of the cathode.

To understand the mechanism by which electrons impart energy to the r. f. fields in the cavities, we assume in the first instant that there is no r. f. field present and there exists an axial magnetic field of intensity B (directed into the page) and an electric field of intensity E (directed radially outward). Under above mentioned conditions the electron moving in clockwise direction is acted upon by following forces :

(*i*) a centrifugal force equal to $\frac{mv^2}{r}$, where v is linear tangential velocity of the electron in its circular path, r is radius of the path in which it revolves and m is the mass of the electron.

(*ii*) a radial outward force equal to eE, where e is the electronic charge.

(*iii*) a third force equal to Bev directed radially inward and arises as the moving electron constitutes a current.

It is evident that third force opposes the first two forces. If magnetic field intensity B is chosen such that an electron revolves undisturbed round the cathode for a given intensity of electric field and

linear velocity, then we have

$$Ee\eta = eE + \frac{mv^2}{r}, \qquad \ldots(1)$$

which is a condition of balance of radial forces.

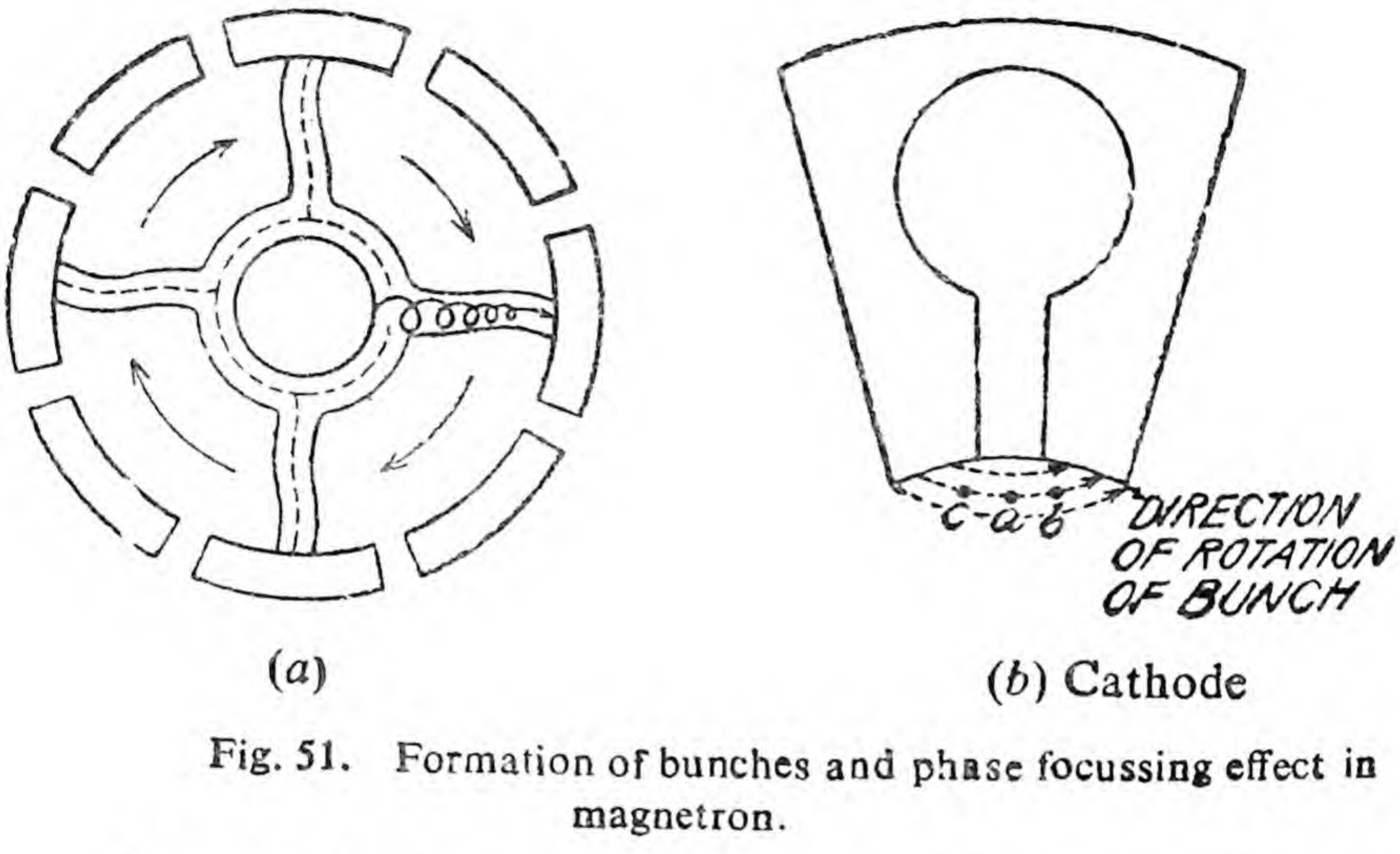
Now we assume the anode cavities are excited by some r.f. source and this produces an instantaneous r. f. field distribution in the interacti on space between cathode and anode. Let at any time electron arrives at point beneath the anode gap when r. f. field is at its peak value and in the direction of electron velocity. The consequency of this situation is that electron imparts energy to the radio frequency field, and if now velocity of electron and frequency of r.f field are co-ordinated, the electron will impact energy to the cavity while passing through each successive gap. Thus a circling electron continually gives energy to the field.

It appears at the first sight that synchronism between electron velocity and field cannot be retain ed because at each successive gap electron loses energy and lessens its velocity. But it is not true because as the electrons are slowed down, the radial inward force decreases and this permits the electric field to force the electrons towards the anode. In this process they regain their lost velocity and rotate in orbits of larger radius essentially with their initial velocity. Eventually the electron strikes the anode and is removed from the intera ction space. Thus an electron in the interaction space is regarded as a means of transfering energy from direct field to r.f. field. It should be noted that in normal magnetron operation, no external source of excitation is required because any r.f. disturbance suffices to give rise to oscillations in cav ties.

In the above idealised operation of magnetron, we assumed that electrons initially circle the cathode with correct angular velocity and phase to be retarded by r.f. fi eld. But in actual practice the electrons are emitted from cathode with zero initial velocity. Further there are many electrons to take energy from the field as these are to impart energy to the field. Thus, on the average, electrons will impart no energy to the field if their motion is not modified.

It has been discussed that in the absence of r.f. field and in the presence of magnetic field which is very well above cut-off value, the electron describes an epicycloidal path. But in the presence of r.f. field, electron does not describe an epicycloidal path because it is under an additional force of r.f. field. The total effect of r. f. field is to bunch the electrons into groups. The number of groups depends upon the mode of oscillations and number of cavities used in anode structure of magnetron. In an eight cavity magnetron oscillating in π mode, the electrons are bunched in four groups and rotate round the cathode with uniform velocity. Thus the

combined effect of magnetic field, direct electrostatic field and r.f. field is to produce a spoke-like arrangement revolving round the cathode. Tne situation is shown in fig. 51 (*a*).



(*a*) (*b*) Cathode

Fig. 51. Formation of bunches and phase focussing effect in magnetron.

To understand the formation of bunching we recall that there are some electrons with leave the cathode at the times and positions that they face an accelerating r. f. field and there is equal good number of electrons which leave the cathode at the times and positions that they face a retarding field. The accelerating electrons extract energy from the field and thereby gain speed. The gain in speed results in an increase in the magnetic field force which forces them back towards the cathode. Thus such electrons are removed from the interaction space after small radial excursion and no more energy is extracted from the field. The decelerating electrons impart energy to the field and loses velocity. The loss in velocity results in the decrease of magnetic field force and electrostatic field find opportunity to urge them towards anode. The process makes the electrons move in orbits of larger radius and compensate the loss of velocity. Thus the synchronism between velocity of electrons and r.f. field is maintained. As the electrons leave the interaction space at the anode after few cycles, other electrons enter into the space and thus bunching is preserved.

It is seen that electrons that find themselves in accelerating field are quickly removed from the interaction space and extract energy from the field for very short time, while those which face decelerating field remain for a long time in the space and transfer energy to field. This lesser amount of energy is extracted from the r.f. field, then the energy transferred to r.f. field from direct field by electrons and oscillations are sustained. The electrons

that are not included in space-charge cloud strike the cathode with considerable velocity on their return journey. The energy is dissipated in the form of heat. Therefore, care is always taken to prevent damage to the cathode through overheating.

As shown in fig. 51 (*a*) the spoke like arrangement of electrons revolves round the cathode in synchronism with r.f. field in a phase that enables the bunch to cross the gap when across it there exists maximum decelerating field. The actu al motion of electrons is like that shown by heavy line within the bunch.

The phase focussing effect of r. f. field preserves the electron bunches and help to maintain the synchronism. The phase focussing mechanism is illustrated in fig. 51 (*b*). When the electron is at the point '*a*' the combined effect of axial magnetic field, direct field and r. f. field is to make the electron rotate with some angular velocity. The electron at the point '*b*' is acted upon by force that opposes the direct field force and thereby reduces its velocity. Thus the electron at '*b*' falls in step with '*a*'. The electron at '*c*' is acted upon by a force which helps direct field force, increasing thereby its velocity so as to make this electron again in step with '*a*'. In this way bunch is preserved.

The performance of magnetron depends upon :

(*a*) strength of magnetic field.

(*b*) anode voltage, and

(*c*) load impedance.

The output power efficiency and frequency are affected by the variation in any of these quantities. The performance is usually studied with the help of Ricke diagram which depicts the performance as a function of load conditions for a given anode voltage and magnetic field strength. It has been observed that the frequency of oscillations is quite sensitive to changes in load impedance.

15·19. TRAVELLING WAVE TUBE :

The travelling wave tube is an amplifier in which a beam of electron is projected axially through a section of waveguide. Due to the interaction between the electrons and the axial component of electric field of wave propagated through the wave guide, amplification takes place. The travelling wave tube is particularly suitable for amplification of very high frequencies, such as 4000 mc/s. The physical construction of a typical travelling wave tube amplifier for the use a t frequencies of about 4000 mc/s is shown in figure 52.

In this tube, an electron gun, operating at a direct accelerating potential of 1500 volts, produces a pencil like beam of electrons. This beam passes through a long, loosely wound helix, and is collected by the collector at anode potential. To prevent the beam

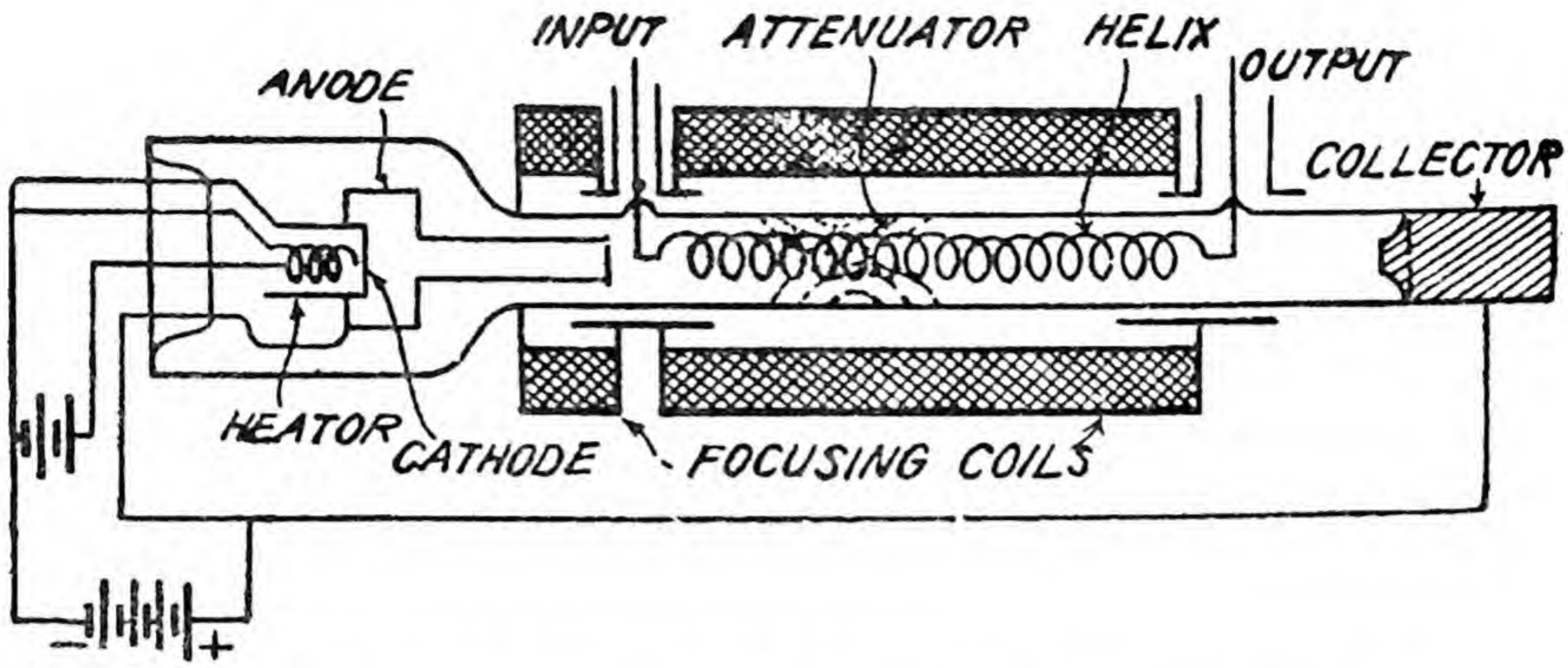


Fig. 52. Schematic diagram of the travelling wave amplifier.

from spreading and to guide it through the centre of helix, an axial magnetic focussing field, as shown in figure, is used. Electrons that move parallel to the magnetic field are not affected while those having a component of motion across the field experience a torque which tends to force them to move in the direction of magnetic field. The signal to be amplified is impressed upon the helix at the input end of the tube and takes the form of a wave travelling along the helix. The interaction of the electron stream and travelling wave results is an amplification of the signal as it progresses along the helix. The amplified signal is removed from the output of the tube. The helix exhibits the property that the waves tends to follow, with the velocity of light, the wire with which it is wound The waves produce a longitudinal component of the electric field on the axis of the helix, propagated axially with reduced velocity. Thus the helix may be considered as a waveguide in which slow waves are propagated. If the length of wire in helix is n times as long as the helix, the velocity of travelling wave is approximately $v=c/n$, where c is the velocity of light. The electrons enter the helix with a velocity slightly greater than that of the travelling wave and tend to become bunched as they travel the helix. Those electrons which are retarded by a.c. field of travelling wave slow down, while those accelerated by a c. field speed up and overtake the slower one. Thus electron bunching takes place. Now the energy transfer takes place from electrons to travelling wave and thereby building up the amplitude of the waves. The travelling wave tube can also serve as an oscillator if a portion of the output is fed back into the input. The tube may oscillate if reflections due to impedance mismatch occur at the output coupling causing a backward travelling wave which carries energy back to the input.

In order to prevent oscillation generated in a travelling wave tube, it is necessary to prevent internal feedback arising from reflections due to slight impedance mismatch at the terminals. This is done by introducing an attenuator at some place moderately near the input end of the tube.

Mechanism of operation : As the electrons are continuously injected into the field some of them are accelerated, while other are decelerated, When the tube is operated under appropriate conditions more electrons are decelerated than have been accelerated and hence net energy is delivered to the wave. The mechanism of energy conversion by which electrons deliver energy to the signal can by under-stood from figure 53.

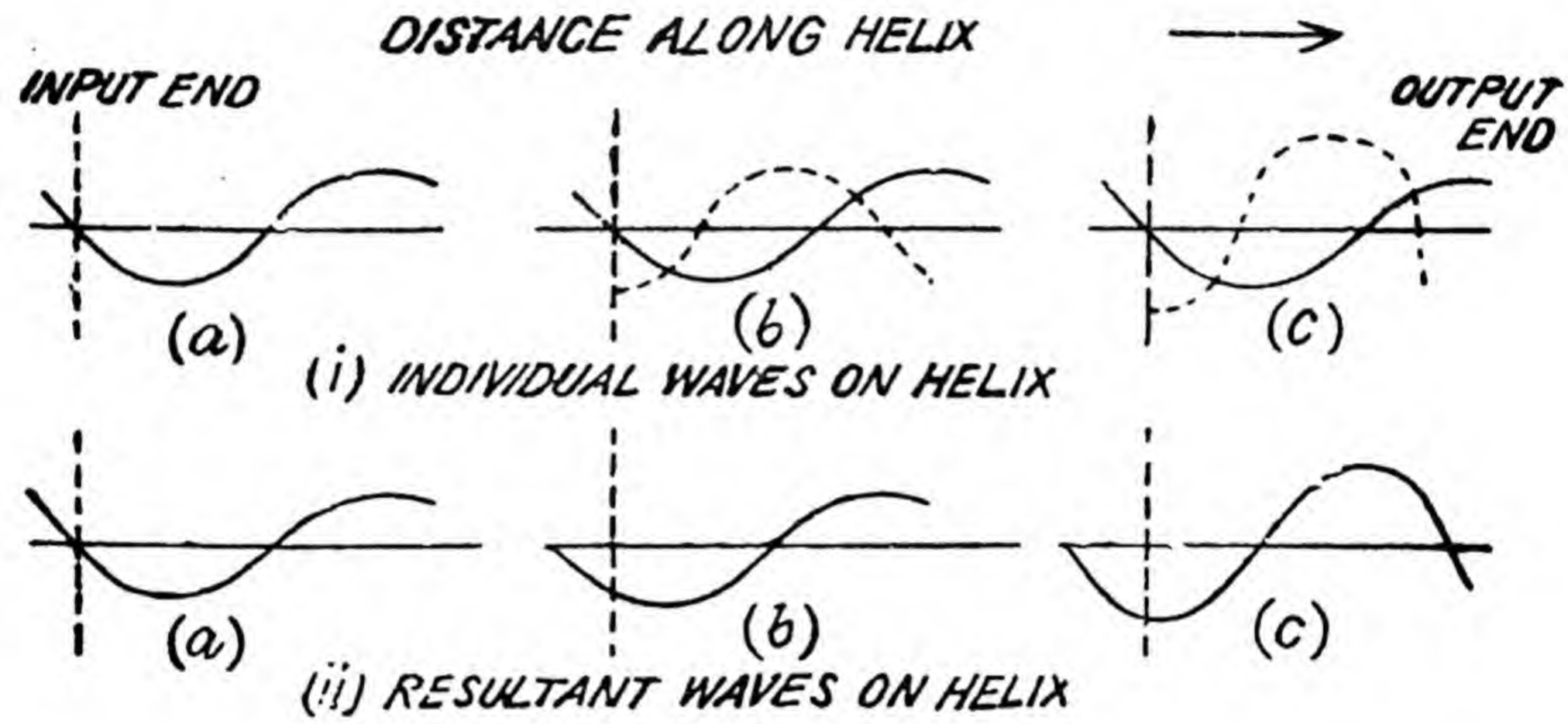


Fig. 53. Showing the electron bunch in a travelling wave.

Consider a group of electrons near the input end in the vicinity of *e* as shown in figure 53 (*iii*) *a*. Let the initial electric held be zero at this point and is negative towards the output end. The electron located exactly at *e* is unaffected by the signal of the helix as there is no axial electric field. However, electron *e''* just right to *e* encounters a decelerated field. This electron slows down and tends to be overtaken by the electron *e*. In a similar way, the electron *e'* just to left of *e* encounters positive axial field. This electron is accelerated and tends to catch up the electron *e*. The electrons centred about *e* are thus velocity modulated

When the velocity modulated electrons have travelled some distance, they begin to bunch about as shown in fig. 53 (*iii*) *b*. The bunch of electrons centred on *e* produces a second wave on the helix which produces an axial electric field. This field lags the solid curve of fig. 53 (*i*) by a quarter wavelength as shown in fig 53 (*i*) *b*. Now the field along the helix axis is the resultant of the two waves on the helix. This is shown in figure 53 (*ii*). It is

clear from the figure 53 (*ii*) *b* that the resultant field is located a small fraction of a wavelength closer to the gun end of the tube than the solid curve of figure 5 (*i*) *b* because the negative maximum of the induced wave is opposite to the centre of the electron bunch. The electrons in the bunch therefore, encounter a retarding field. The electrons thus deliver energy to the wave on the helix. The bunching continues because the electron *e* is no longer at the zero of the resultant wave on the helix.

As the electrons travel further along the helix, the situation shown by *b* in figure 53 changes to *c*. Now the bunching is more complete and the induced wave grows in amplitude. The phase shift of the resultant wave relative to the electron bunch is also increased because the induced wave in this case is larger. Each electron in the bunch now encounters a stronger retarding field A large and increasing amount of energy is thereby delivered by the electron bunch to the wave on the helix, which is now much larger than the original signal.

The motion of the individual electrons has been studied by Webber. He showed that when the bunches have been formed the electrons in the bunches do not simply coast with the wave but they oscillate longitudinally within the beam. The equations expressing the oscillatory motion of the electrons relative to a point within the beam moving with the initial velocity of the beam are of the same form as those for the motion of a pendulum. Unless the initial velocity of an electron exceeds a critical value, the oscillation takes place. If the initial velocity of an electron exceeds a critical value corresponding to injection angle (time in the cycle, expressed in radians, at which the electron is injected) its average velocity is not reduced to that of the wave but it continues to advance relative to the wave. Such electrons deliver energy to the wave while passing through the regions where they are retarded by the electric field of the wave but reabsorb energy where they are accelerated. The electrons deliver energy to the wave only when they leave the field at the instant when their velocity is lower than the initial velocity

Coupling of beam and slow-wave structure :

Now we consider that how the electron beam waves interact with the travelling electromagnetic wave on an adjacent slow wave guiding structure and produce amplification of the wave guide waves. The amplification can be shown by a relatively simple model in which coupling is introduced between the wave system of the beam and the travelling waves on the slow wave circuit. The coupling is provided by the displacement current induced between the charge concentrations travelling on the beam and the wave guiding structure. The assumed situation is shown in figure 54.

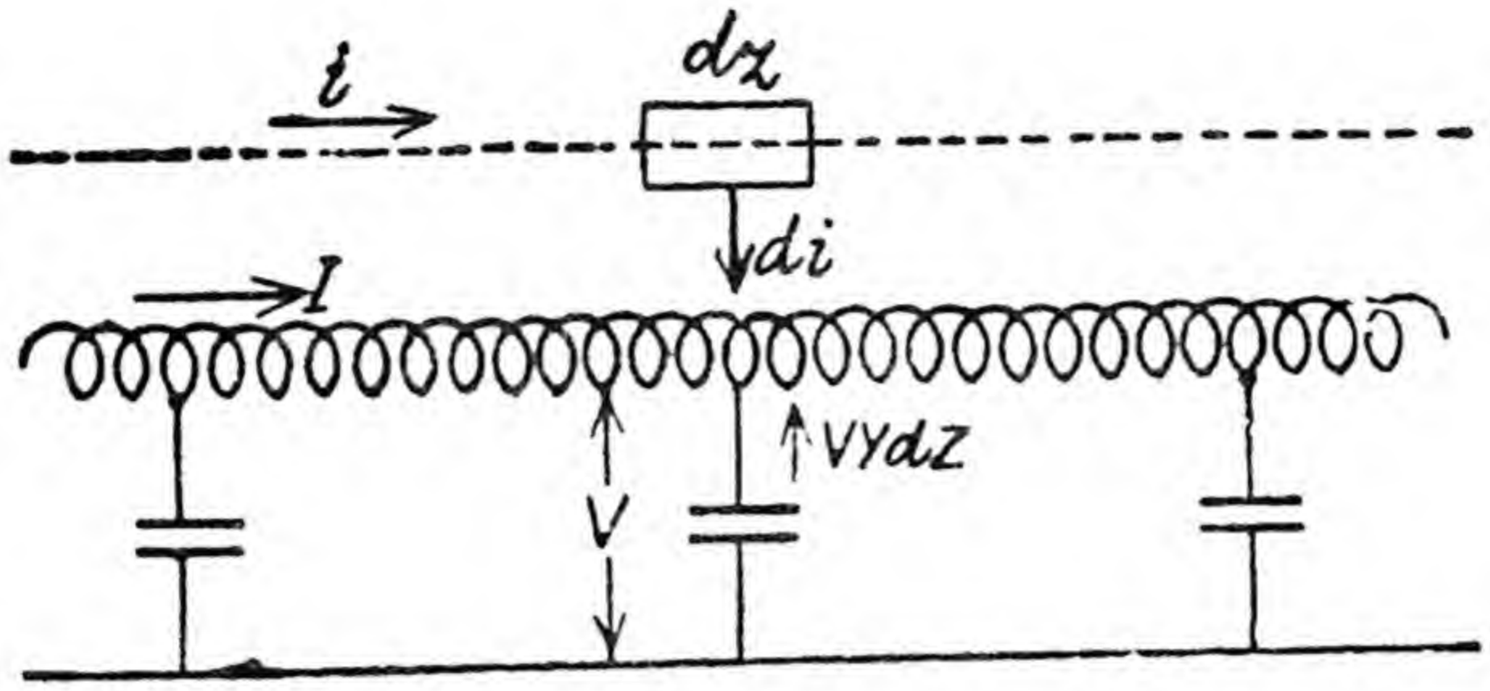


Fig. 54. Slow-wave structure and coupled beam.

The electron beam current i flows adjacent to a lossless transmission system in which the waveguide current and voltage are I and V, respectively. Let the transmission system has impedance $Z=jX$ and admittance $Y=jB$ per unit length In the absence of the beam, the waveguide current and voltage would be related by transmission line equations :

$$dV=-IZ\,dz \qquad \ldots(1)$$

$$dI=-VY\,dz. \qquad \ldots(2)$$

These equations represent a transmission line upon which waves travel with the propagation constant $j\beta_1=(ZY)^{1/2}$. Due to the presence of electron beam, the transmission line equations are modified. Let an additional current be introduced in dI. This current increment is set equal to the beam current which flows into the beam length dz and therefore is $di=\frac{\partial i}{\partial z}.\,dz$. Thus it is assumed that di flows as a Maxwellian displacement current from the beam to the slow wave circuit. The transmission line equations now become

$$dV=-IZdz \qquad \ldots(3)$$

$$dI=-VYdz-di. \qquad \ldots(4)$$

Replacing d/dz by $-j\beta$ in these equations, we have

$$-j\beta V=-IZ \qquad \ldots(5)$$

and

$$-j\beta I=-VY+j\beta i. \qquad \ldots(6)$$

Combining equations (5) and (6) and replacing ZY by $-\beta_1^2$, we obtain

$$(\beta^2-\beta_1^2)\,V=j\beta Zi. \qquad \ldots(7)$$

The value of i is given by the relation

$$i=-\frac{2\omega\beta_p^2}{\beta r_0}\times\frac{D_r}{(\beta-\beta_0)^2-\beta_p^2}. \qquad \ldots(8)$$

where r_0 = radius of cylindrical surface in which the electron beam is enclosed,

β = propagation constant,

D_r = radial component of electric displacement,

$\beta_p^2 = \omega_p^2/v_0^2$,

ω_p = plasma frequency of space charge of mean density ρ_0,

v_0 = mean velocity.

Substituting the value of i from equation (8), in equation (7) we have

$$(\beta^2 - \beta_1^2)\ V = -\frac{2j\beta_p^2 Z}{r_0} \times \frac{D_r \pi r_0^2}{(\beta-\beta_0)^2 - \beta_p^2}. \qquad \ldots(9)$$

The factor D_r is ϵ_0 times the radial component of electric field E_r, at the surface of the beam. We know that wave guide voltage V is proportional to transverse field, we assume that D_r is related to V by a constant of proportionality. We, therefore, write equation (9) as

$$\beta^2 - \beta_1^2 = -\frac{A}{(\beta-\beta_0)^2 - \beta_p^2}. \qquad \ldots(10)$$

where

$$AV = -\frac{2\omega\beta_p^2 \times \epsilon_0 E_r}{r_0} \pi r^2. \qquad \ldots(11)$$

In equation (11), the negative sign is introduced on *ad hoc* basis and we have used $Z = jX$.

For convenience, the various phase constants are termed as

$\beta_0 = \dfrac{\omega}{v_0}$ = phase constant of average beam velocity,

$\beta_p = \dfrac{\omega_p}{v_0}$ = phase constant of plasma frequency.

β_1 = propagation constant of slow waveguide alone.

The equation (10) may be written in the form

$$(\beta+\beta_1)(\beta-\beta_1)\ [(\beta-\beta_0)^2 - \beta_p^2] = -A. \qquad \ldots(12)$$

Equation (12) is fourth order in β It has four roots corresponding to four travelling waves in this system. Since β_p is smaller than β_0, hence there are three forward travelling wave solutions and one backward travelling wave which corresponds to the solution $(\beta+\beta_1)$. Here we are interested only in the interaction between forward travelling waves and space charge waves on the beam and hence the backward wave is of little interest. Replacing $(\beta+\beta_1)$ by $2\beta_1$ in equation (12), we have

$$(\beta-\beta_1)\ (\beta^2 - 2\beta\ \beta_0 + \beta_0^2 - \beta_p^3) = -A/2\beta_1. \qquad \ldots(13)$$

A particular solution of equation (13) will be considered when $\beta_1^2 = \beta_0^2 - \beta_p^2$. We can write equation (13) as

$$(\beta-\beta_1)^3 = -\frac{A}{2\beta_1} \text{ approximately since } \beta_0 \approx \beta_1. \qquad \ldots(14)$$

In solving equation (14). we consider the propagation constant β as complex and include the complex roots of right hand side.

$$\therefore \quad \beta - \beta_1 = \left(\frac{A}{2\beta_1}\right)^{1/3} [e^{j\ (2n+1)\ \pi}]^{1/3}. \qquad \ldots(15)$$

Let $(A/2\beta_1)^{1/3}=b$, then the three roots of equation (15) can be written as

$$\beta=\beta_1+b\,[\cos \pi/3 + j \sin \pi/3], \quad \ldots(16)$$

$$\beta=\beta_1+b\,[\cos \pi + j \sin \pi], \quad \ldots(17)$$

$$\beta=\beta_1+b\,[\cos (5\pi/3)+j \sin (5\pi/3)]. \quad \ldots(18)$$

The three forward travelling wave solution have the wave functions

$$\exp\,[-j\,(\beta_1+0{\cdot}5\,b)]\,z.\exp\,[-0{\cdot}866\,b]\,z, \quad \ldots(19)$$

$$\exp\,[-j\,(\beta_1-b)]\,z, \quad \ldots(20)$$

$$\exp\,[-j\,(\beta_1+0{\cdot}5\,b)]\,z.\exp\,(+0{\cdot}866\,b)\,z. \quad \ldots(21)$$

The second wave, equation (20) is an unattenuated wave which travels slightly faster than the characteristic of slow wave structure. The other two waves move slight slowly than the waves of the guide alone. Equation (19) shows that the wave is attenuated exponentially as it travels, while equation (21) shows that the wave grows exponentially as it travels.

The results derived above show the essential characteristic of travelling wave tube operation. By coupling the two wave systems, the slow waveguide, having two oppositely travelling waves at a given frequency, and the electron beam, with its pair of oppositely travelling space charge wave, we have a system of four possible wave motions characteristic of the system as a whole. Of the forward travelling waves, one is growing wave, it transfers the original kinetic energy of the beam to the electromagnetic fields in the tube. This results in the travelling wave amplification.

Variation of gain with beam voltage : The phase velocity v of the growing and attenuated waves is given approximately by the relation

$$v=\frac{v_0}{1+C/2}, \quad \ldots(22)$$

where v_0 is initial velocity of the electrons and C is related to the characteristic impedance Z_0 of the helix, the beam current I_a and the beam accelerating voltage V_a. The value of C is given by

$$C^3=\frac{Z_0\,I_0}{4V_0}. \quad \ldots(23)$$

The phase velocity of constant-amplitude wave is given approximately by relation

$$v=\frac{v_0}{1-C}. \quad \ldots(24)$$

The backward wave has constant amplitude and moves with the phase velocity given approximately by the relation

$$v=\frac{v_0}{1-C^3/4}. \quad \ldots(25)$$

As the typical values of C are ·02 or less, the phase velocity of growing and attenuated forward waves is one percent or less lower than the initial velocity of the electrons. Similarly the constant amplitude forward wave is two percent or less higher than the electron velocity. On the other hand, the phase velocity of backward wave is slightly higher than the electron velocity. The three forward waves have equal amplitude at the input end of the helix but if the tube is sufficiently long the amplitudes of attenuated and constant amplitude waves are negligible in comparison to the growing wave at output end.

Figure 55 shows a curve of gain versus beam accelerating voltage for a typical tube. The maximum gain and output are obtained when the beam accelerating voltage is adjusted in such a

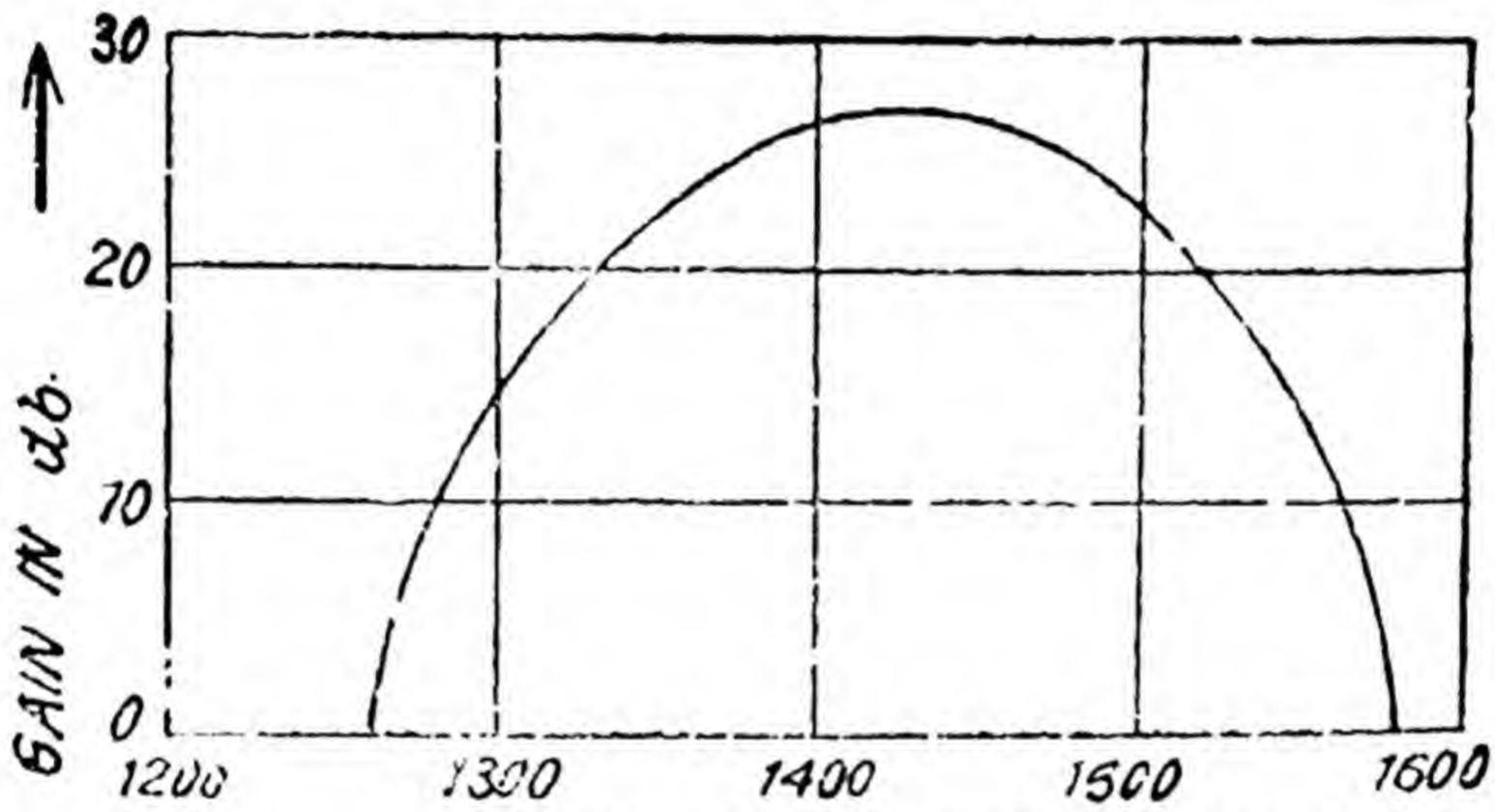


Fig. 55. Variation of gain with beam voltage

way that the initial velocity of the electrons is equal to the phase velocity of propagation in the absence of electron beam. It is clear from the figure that the gain is maximum when the beam accelerating voltage is 1430 volts. The gain, in decibels, of a tube that is sufficiently long so that the grownig wave predominates at the output end of the helix is given by

$$G = -9{\cdot}54 + 47{\cdot}3\ CN, \qquad \ldots(26)$$

where N is the length of the helix in wavelength. As the gain varies directly with the parameter C, this is known as gain parameter.

Under small signal operation, the power output of a travelling wave tube is given by the relation

$$P = CV_a\ I_a. \qquad \ldots(27)$$

Variation of small signal gain with wavelength : The small signal gain of 6 to 8 cm. tube is shown in fig. 56.

The gain decreases at longer wavelengths. This decrease is due to the following two facts :

(i) the phase velocity in the helix increases with increasing wavelength, and,

(ii) due to the deterioration of the matching of the helix to the waveguide.

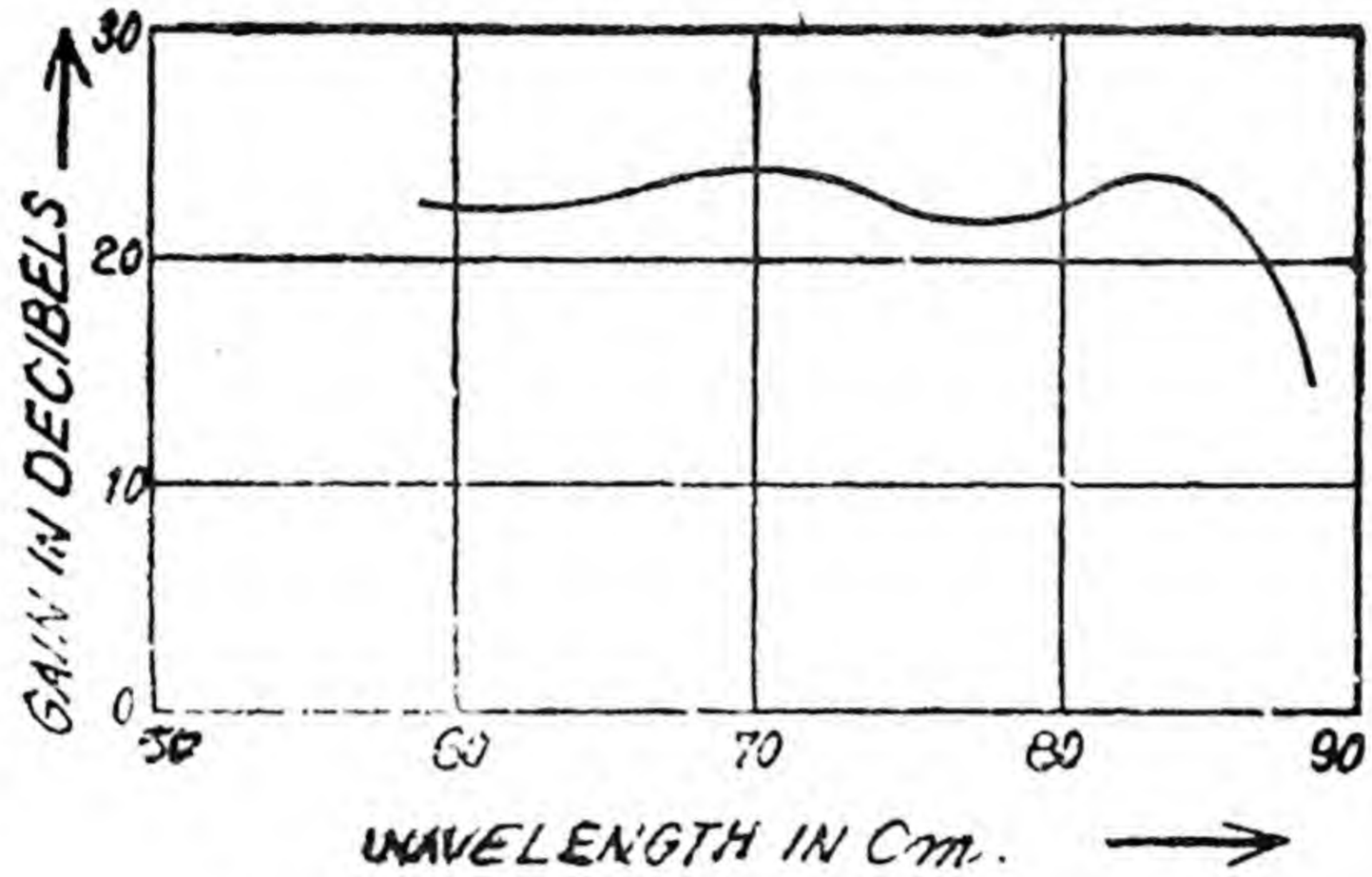


Fig. 56 Small signal gain of the tube.

Band width : The travelling wave tube is inherently a non-resonant device. Due to this fact, it can be made to have band-widths that are enormous compared with those obtainable from amplifiers involving resonant circuits. The amplification characteristic of a typical helix travelling wave tube is shown in fig. 57. From the figure it is observed that the amplification is nearly constant from 200 to 4000 mc/s.

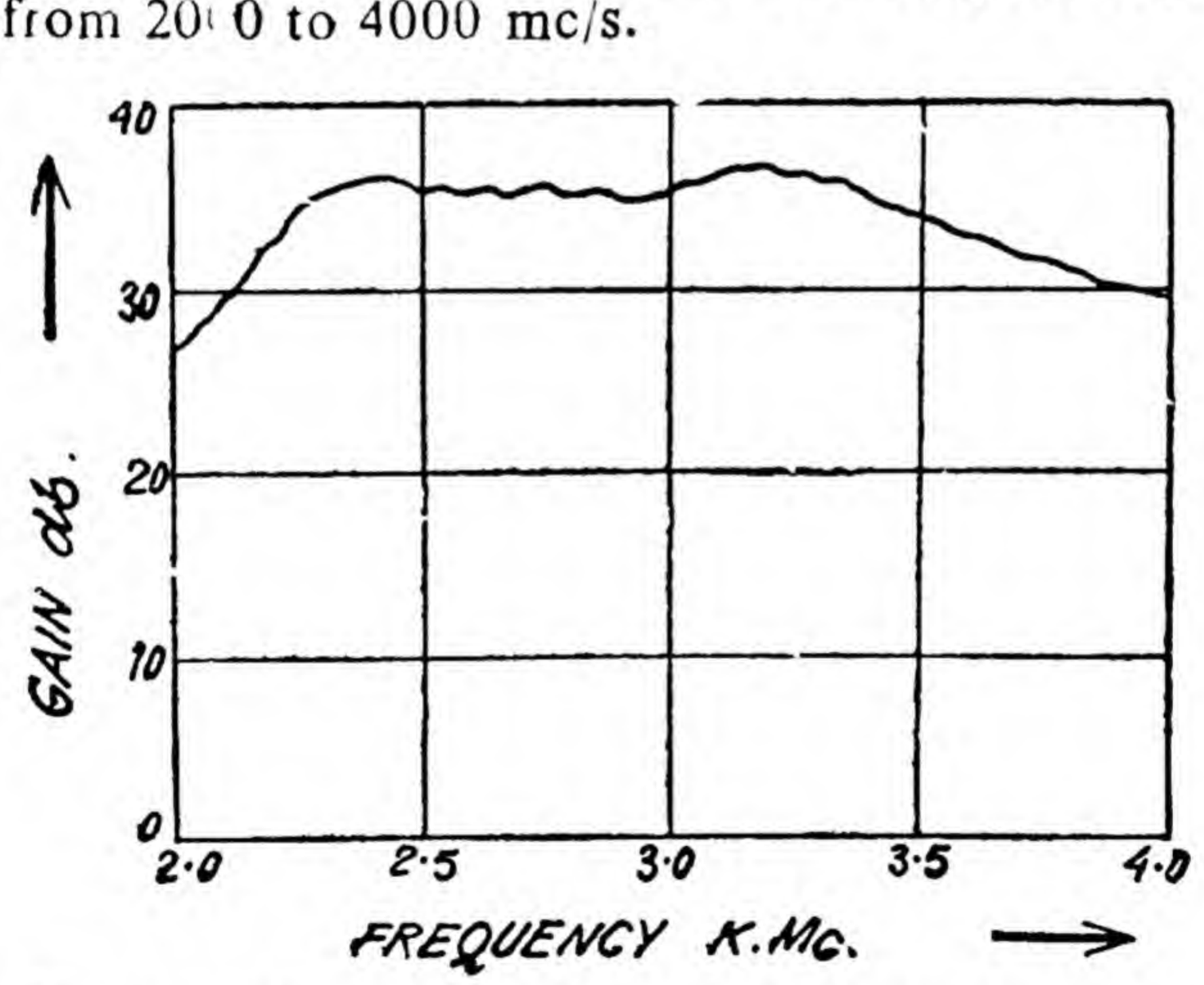


Fig. 57. Amplification of a travelling wave tube as a function of frequency.

The gain of a travelling wave tube varies with frequency. The reason of this variation is due to the following four factors :

(i) variation in the velocity of the electric field along the axis of the tube,

(ii) variation of length of the tube in wavelengths,

(iii) variation in the strength of the axial electric field as a function of frequency,

(iv) failure to match the terminal impedance of the tube accurately at all frequencies.

The helix form of the travelling wave tube is particularly suitable for achieving wideband operation. Over a wide frequency range, the velocity of axial field produced by the helix is substantially independent of frequency.

15·20. BACKWARD-WAVE OSCILLATOR (CARCINOTRON) :

The backward wave oscillator is shown in figure 58. The body of the tube consists of a folded transmission line. Alternately the body may be regarded as a waveguide operating in the TE_{10}, mode such that a wave travelling along the line winds itself back and forth, and in the process produces an axial component of electric field. An electron beam is directed along the axis through holes in the structure. When the electron beam has a suitable velocity, the interaction between the backward wave, *i.e.* a wave travelling from right to left and electron beam takes place. Now the average energy is delivered to the wave by electron beam. The collector end of the line is terminated with a matched load impedance for the purpose of absorbing any power that might be reflected at the out-put *i.e* , at the gun end as a result of an impedance mismatch at that point.

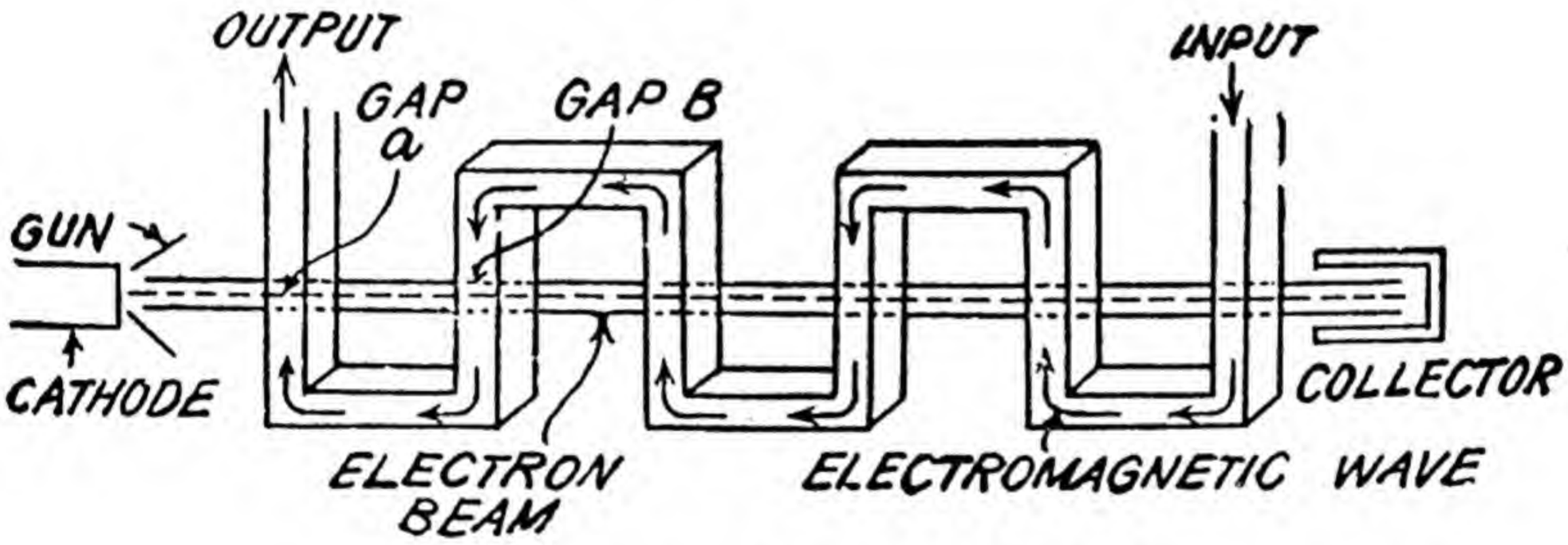


Fig. 58. Backward wave oscillator.

Mechanism of operation : In order to consider the mechanism of operation, we have to consider the interaction between the backward wave and electron steam. Here we assume that there is a backward wave in the structure. Let us first of all consider the situation existing at gap *a*, where the electrons enter from cathode side at a uniform rate. The electrons are subjected to an alternating axial field that varies with time. The electrons experience velocity modulation. A bunch of electrons is now formed about the reference electron that passes through the gap *a* when the alternating field is zero. Let the reference electron passes through the gap *b* from gap *a* after a time T_e. The delay corresponds to $N_e = T_e f$ cycles, where f is the frequency. The field at *a* lags $N_w = T_w f$ cycles, behind at the field *b*. Here T_w is time taken by the wave to travel from the gap *b* to gap *a*. Thus when the electron crosses the

gap *b* the field it encounters is N_e+N_w cycle behind the phase of the field it encounters at *a*. The geometry of folded line is such that if the electrons are given a velocity such that N_w+N_e is just less than a half cycle, then the field that the reference electron encounters at gap *b* will be as shown in the figure 59 (*b*). The field at *b* will vary with time as shown in figure 59 (*d*). In figure 59 (*d*), ΔT is the time such that $\Delta T \times f = N_\Delta$ is the fraction of a cycle which N_e+N_w fails to be exactly half cycle. If N_e+N_w is exactly half cycle $N_\Delta = 0$ and the electrons passing through *b* are further velocily modulated.

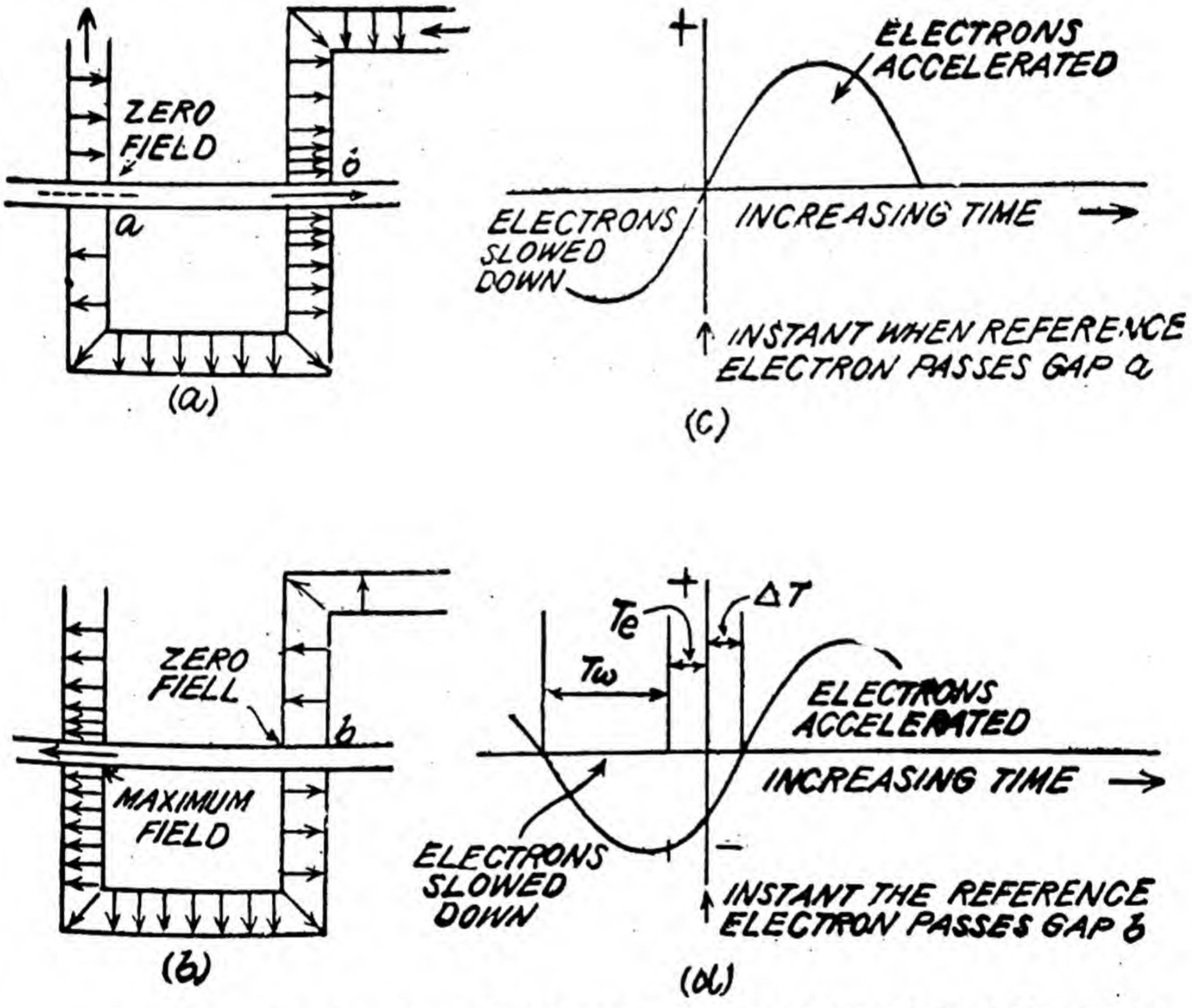


Fig. 59. Relationship of fields and electrons in the folded line backward wave tube.

The situation is repeated again and again as the electrons centred about the reference electron pass one gap after another. Thus as the electrons travel toward the collector, the interaction with the backward wave on the folded line causes them gradually to group together in bunches. If the electron velocity is so chosen that at the collector end the total phase advance is approximately half cycle, the electron bunches encounter relatively strong decelerating fields as they cross the gaps. The fields slow down the electrons and cause energy to be delivered to the backward wave. Since the electron beam conveys r.f. current from the output of the tube back to the input corresponds to the feed back path in the equivalent

amplifier. Furthermore, the r.f. current conveyed by the beam is small at the output of the tube, whereas it is large at the input, the percentage feed back in the equivalent amplifier may be much greater than 100%.

Frequency of oscillations : We have shown that in order to generate oscillations, the electron velocity must have a value such that $N_\Delta = n\Delta Tf$ approximately one half cycle, where n is number of gaps and other symbols have their usual meaning. Also $N_e + N_-$ is less than a half cycle by N_Δ, hence

$$N_e + N_\omega + N_\Delta = T_e f + T_w f + N_\Delta = 0{\cdot}5$$

$$N_\Delta \approx \frac{0{\cdot}5}{n}.$$

Hence

$$f \approx \frac{0{\cdot}5\,[1-(1/n)]}{T_e + T_w}$$

T_c, T_w and f are defined earlier.

As T_w is determined by the geometry of the tube and so is nearly constant for any given tube. The frequency generated by the backward wave tube is thus controlled by the transit time T_e, *i.e.* by the velocity of electron stream. Hence the frequency is a function of anode voltage. In this way, the backward wave oscillator has the unique property of being voltage tuned. Frequencies below 200 mc/s have been obtained from experimental tubes less than 2 ft. long. Frequencies as great as 100,000 mc/s have been achieved in structures that are physically large enough to be fabricated without undue difficulty.

15·21. ELECTRON WAVE TUBE :

This is essentially a travelling wave tube in which the guiding structure is avoided between input and output as shown in fig. 60.

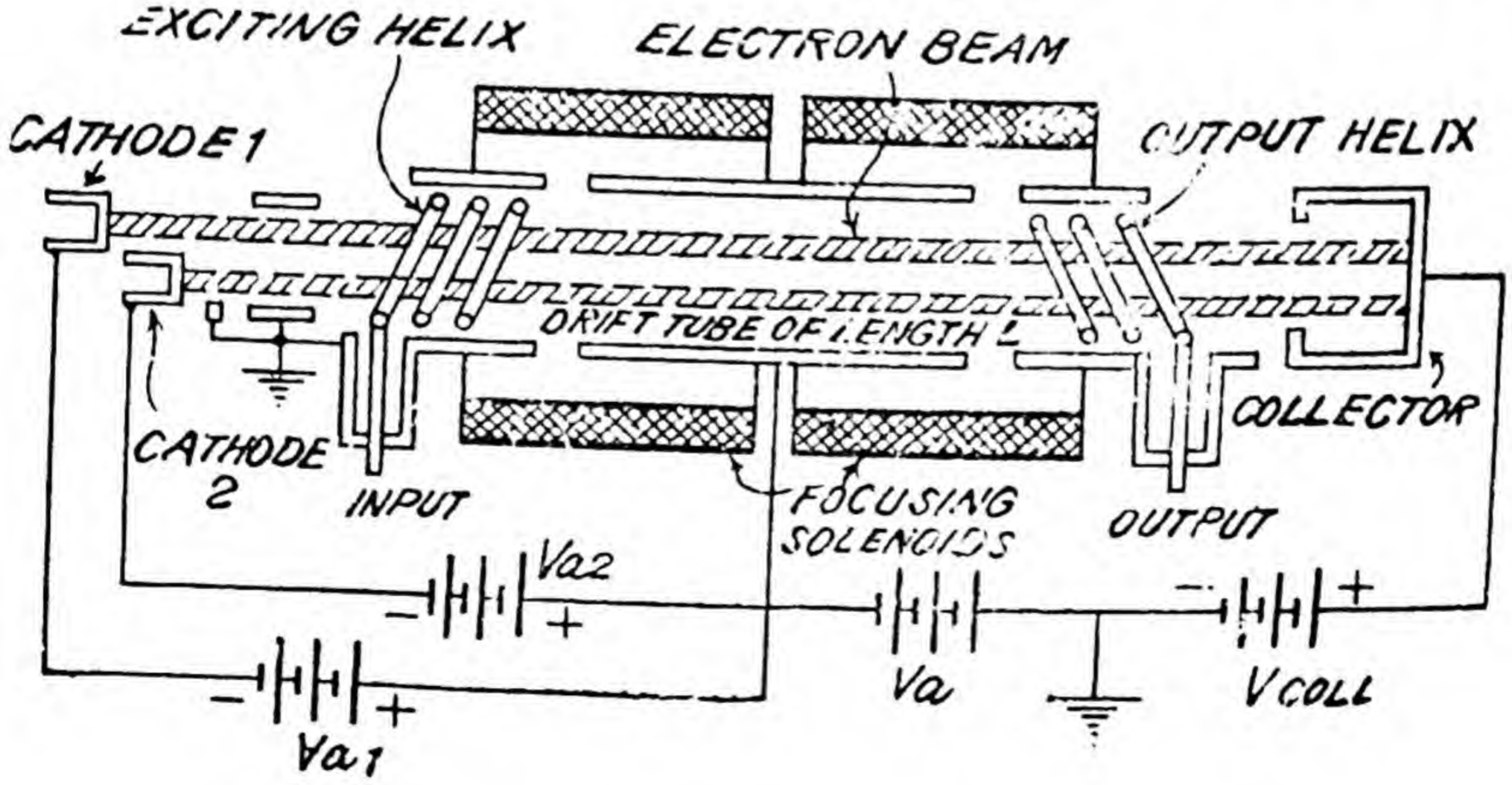
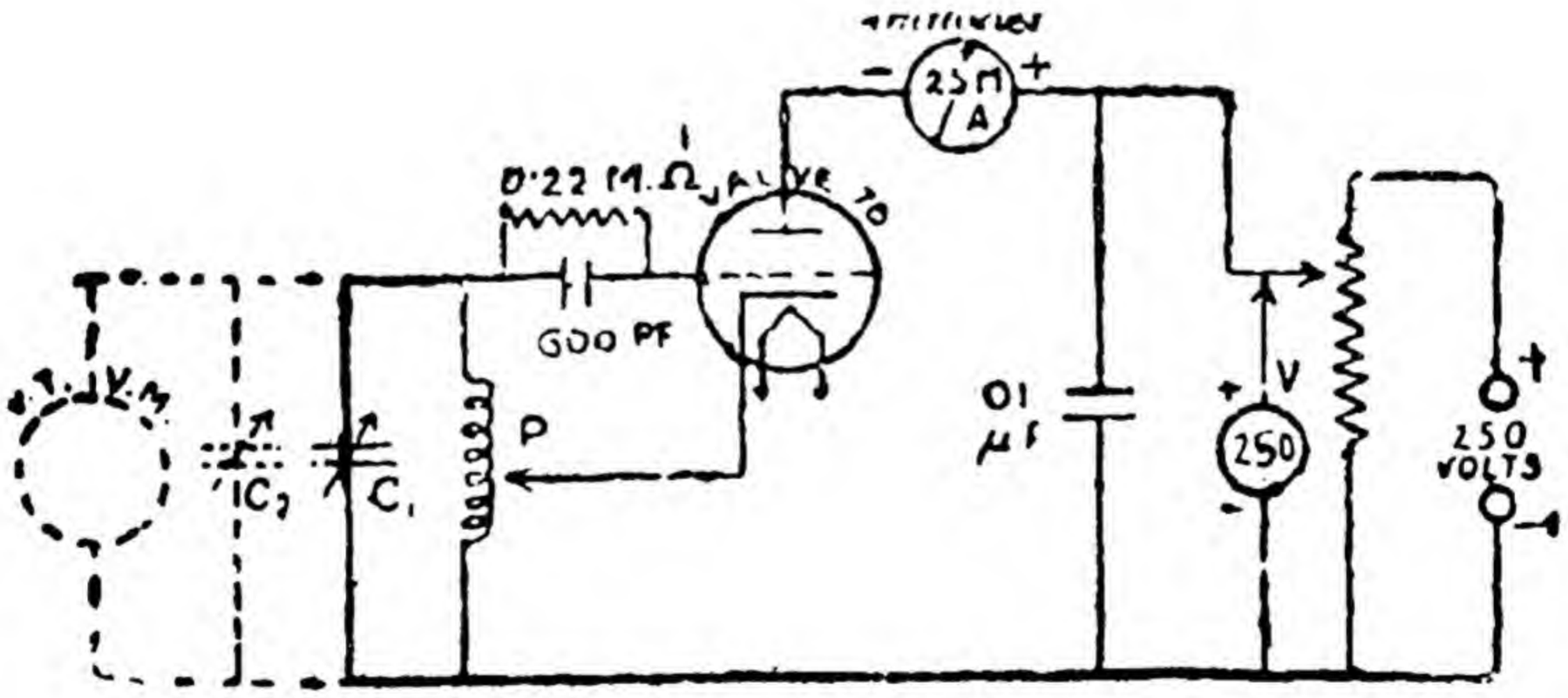


Fig. 60. Electron wave tube two beam type.

In this tube, there are two different electron guns *i.e.* two electronic beams are produced. These beams are accelerated through different direct accelerating potentials V_{a_1} and V_{a_2}. The average velocities of the individual electrons in the beams are different. The two beams are mixed together to form a single beam in which average electron velocities are inhomogeneous. This beam is subjected to an alternating accelerating potential in a manner similar to that in a klystron The inhomogeneous beam is passed through a short helix that is excited by r.f. energy supplied from a generator connected to the input terminals. The axial component of r.f. electric field in the helix causes the electrons in the homogeneous beam to have an alternating component of velocity that is superimposed upon their average velocities. The interaction between electrons that have different velocities takes place in the drift space which results in the production of bunches. The bunches travel through the drift space towards the output end. In the drift space the density of charge continually increases as the bunch travels. At the output, the energy is extracted from the beam by passing it through a suitable output structure such as short section of helix as shown in the figure. The tube exhibits considerable power gain due to mutual interaction of electrons having different velocities. The defocusing of the beam as a result of space charge effects is prevented by a focussing solenoid that provides an axial magnetic field along the drift space.

EXPERIMENT

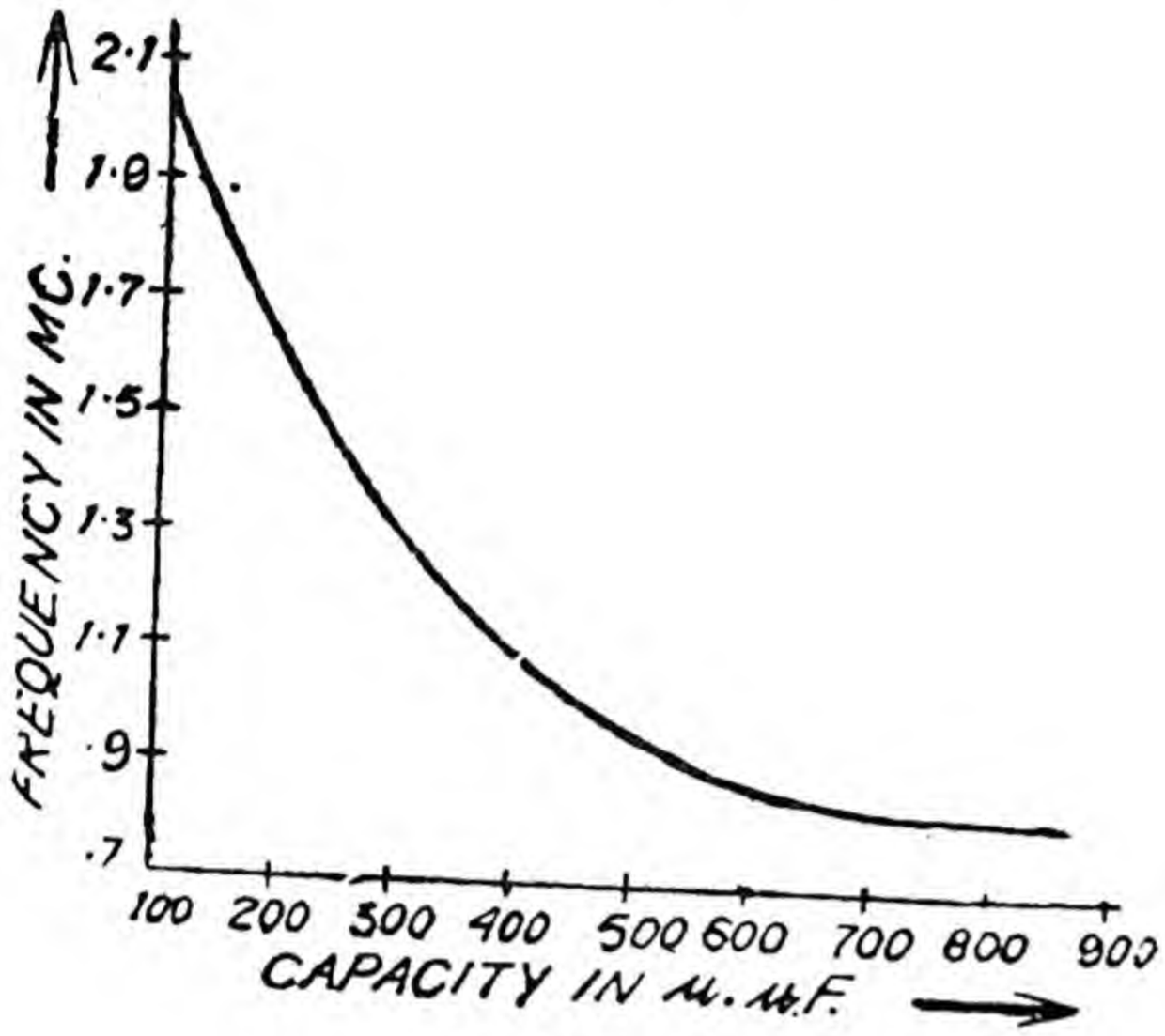
To verify the conditions for oscillations of a Hartley circuit and to determine the dielectric constant at various frequencies.



Hartley Oscillator

Procedure : Move the tapping key on the inductor *P* till the plate ammeter shows a distinct dip. The circuit is now oscillatory.

(i) Calibration of tank condensei C_1 : For different settings of variable condenser of oscillator, determine the frequency generated by noting the resonance in the wavemeter. Draw a graph between dial reading on the condenser and frequency.



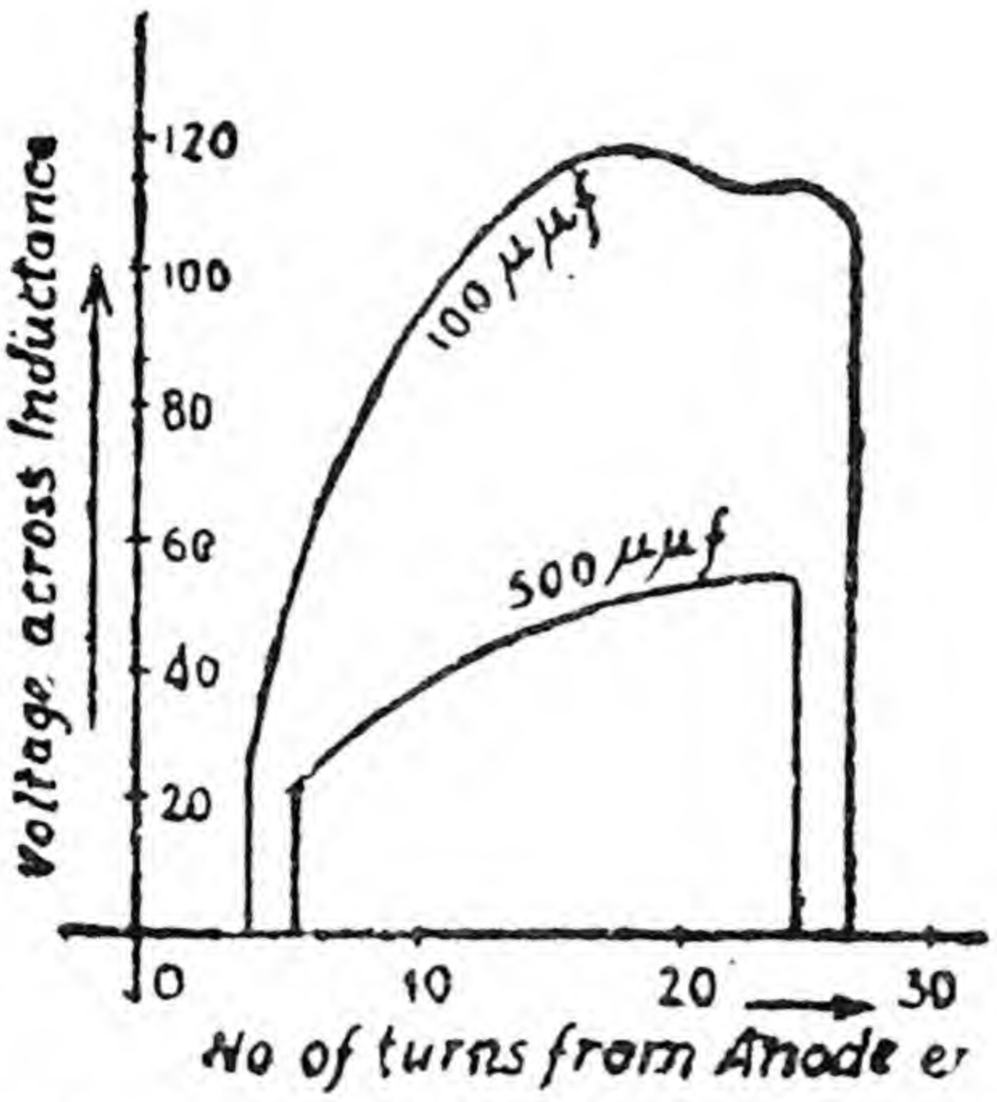
Capacity of C_1 at 100 V	Frequency on wave-meter
100	1·89 mc
200	...
300	...
400	...
500	...
600	910 kc/s.

Calibration of C_1 in terms of frequency graph (*i*).

(ii) Adjustment for Maximum Oscillations : The point *P* is moved by convenient increments along the coil towards the grid end, while the V.T.V.M. reading is noted for each position of *P*,

No. of turns	V.T.V.M. Reading	
	Dial reading of C_1= 500 $\mu\mu F$	Dail reading of C_1= 100 $\mu\mu F$
1	—	—
2	—	—
3	—	—
4	—	—
5	—	—
6	—	—
7	—	—
8	—	—
9	—	—
10	—	—
11	—	—
12	—	—

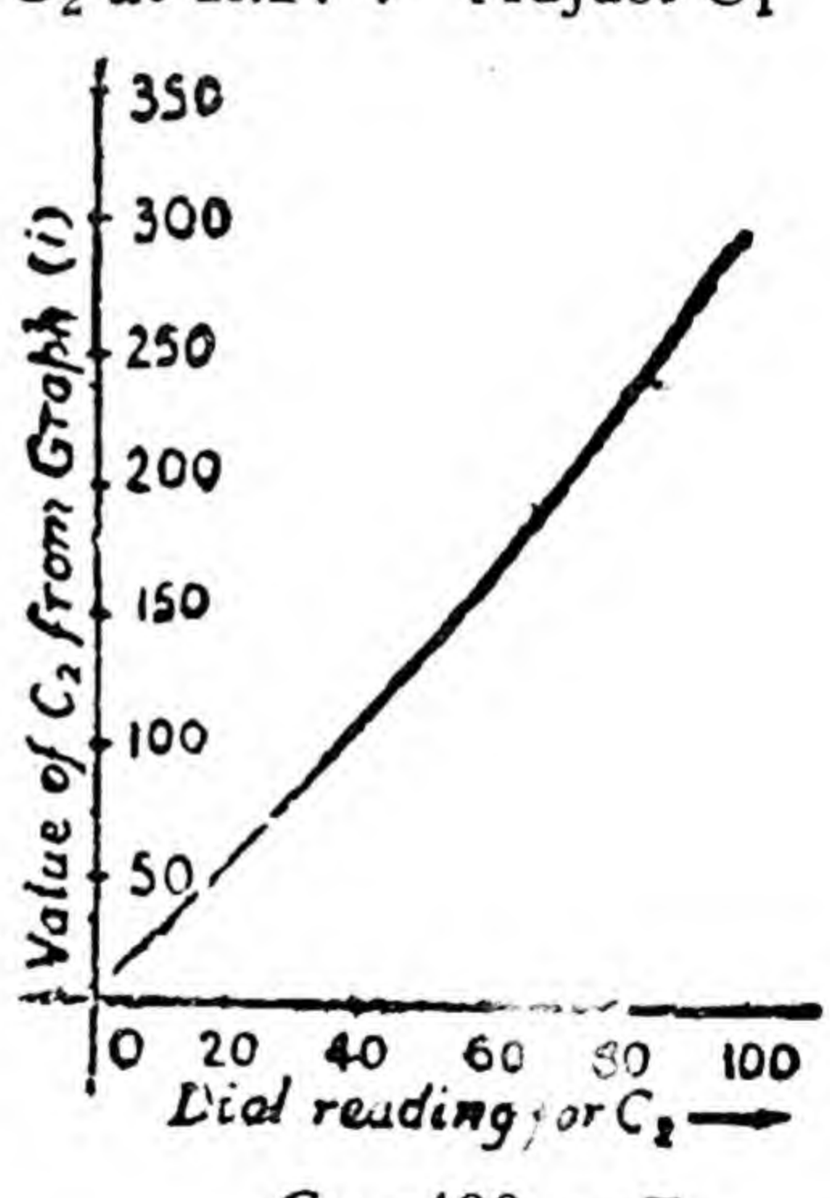
of course, for a fixed value of C_1. The V.T.V.M reading is proportional to the oscillatory current in the coil. The test is repeated for different values to C_1. The results are tabulated, and curves are drawn of voltage to a base of no. of turns on the coil from the anode end as shown in the figure.



Showing dependence on positions of tapping point.

It is clear from the graph that the tapping position for maximum oscillatory current moves towards the grid end of the coil as the value of C_1 increases.

(iii) Calibration of variable condenser C_2 at R.F. : Adjust C_1 at a particular value, say 400 $\mu\mu F$. Connect C_2 in parallel as shown in the figure and set it at the lower value. Adjust the wavemeter to indicate resonance and note the frequency. Vary C_2 for different dial readings and note the corresponding frequencies. Using graph (i) find out the total value of C_1 and C_2 with respect to these frequencies and by subtracting C_1 from this value, obtain the value of C_2. Draw a graph between dial reading and capacity of C_2. The graph so obtained is known as calibration curve for C_2.



C_1=400 $\mu\mu F$.

C_2 dial reading	Freq. in kc/s.	C_1+C_2 from graph (*i*)	C_2 in $\mu\mu F$
0	1100	430	40
20	975	530	135
40	950	60	160
60	...	...	...
80	...	...	...
100	...	...	...

(iv) Dielectric constant at R F. : Set C_1 at a particular value and find out the capacitance of the known plate condenser with

dielectric. Proceed with other frequencies by varying C_1 and find the value of dielectric constant using the following formula :

$$C = K \frac{A}{4\pi t} \text{ e.s.u.}$$

$$= K \frac{A}{36\pi t} \times 10^{-12} \text{ in practical units,}$$

where C is capacity of condenser, K is dielectric constant, A is area of the plate and t is thickness of dielectric medium between plates.

Capacity of C_1 in $\mu\mu F$	Frequency by wave-meter	C_2 + capacity of C_1 in $\mu\mu F$ from Graph (*i*)	Capacity of C_1 in $\mu\mu F$	Dielectric Constant
100	1·56 m.c.	205	105	3·89655
200	1·34 ,,	300	100	—
300	—	—	—	—
400	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—

NUMERICAL EXAMPLES

1 *A tuned anode oscillator employs a valve with* $\mu = 30$, $r_p = 10\ K\Omega$, *and a tuned circuit, comprising a coil of Q factor of 45·5, a capacitance of 350 pF and a resonant frequency of 50 kc's. Calculate (i) the value of minimum mutual inductance between grid and plate coil, and (ii) frequency of oscillations.*

$$Q = \frac{\omega_r L}{R} = \frac{1}{\omega_r C R}$$

$$\therefore \quad R = \frac{1}{\omega_r C Q} \text{ and } L = \frac{QR}{\omega}.$$

$$\text{Now } R = \frac{1}{\omega_r C Q} = \frac{1}{2 \times 3{\cdot}4 \times 50 \times 10^3 \times 350 \times 10^{-12} \times 45{\cdot}5}$$

$$= 200 \text{ ohms.}$$

$$L = \frac{QR}{\omega} = \frac{45{\cdot}5 \times 200}{2 \times 3{\cdot}14 \times 50 \times 10^3} = 29\ mH.$$

Agaih $\quad g_m = \dfrac{\mu RC}{\mu M - L}$, where $g_m = \dfrac{\mu}{r_p} = \dfrac{30}{10^4}$

$$\therefore \quad \frac{30}{10^4} = \frac{30 \times 200 \times 350 \times 10^{-12}}{30 \times M - 29 \times 10^{-3}}$$

or $\quad 30M = 200 \times 350 \times 10^{-12} \times 10^4 + 29 \times 10^{-3}$

$$M = \frac{200 \times 350 \times 10^{-8} + 29 \times 10^{-3}}{30}$$

$$= \frac{7\times10^{-4}+22\times10^{-3}}{30}$$

$$= \left(\frac{0{\cdot}7+29}{30}\right)\times10^{-3}$$

$$= 0\ 99\ mH.$$

(ii) The frequency f is given by

$$f = \frac{1}{2\pi\sqrt{(LC)}}\left(1+\frac{R}{r_p}\right)^{1/2}$$

$$= \frac{1}{2\pi\sqrt{(29\times10^{-3}+3\,50\times10^{-12})}}\left(1+\frac{200}{10^4}\right)^{1/2}$$

$$= \mathbf{50{\cdot}4\ kc/s.}$$

2. *A tuned anode oscillator has anode circuit resonant frequency of* 400×10^{-3} *radians/sec. and inductance* $L=40$ *millihenry. Frequency of oscillation is* ·05 *per cent above the resonant frequency of the tuned circuit. The electron tube used has dynamical plate resistance* $r_p=30{,}000$ *ohms and amplification factor* $\mu=30$. *Calculate the minimum value of mutual inductance* M *between grid and plate coils necessary for sustained oscillations. Assume the effect of inter-electrode capacitances to be negligible.*

$$\omega_0 = 400\times10^3 \text{ radians/sec}$$

$\omega = 0{\cdot}05$ percent above the resonant frequency.

Angular frequency of oscillations is given by

$$\omega = \omega_0\left(1+\tfrac{1}{2}\frac{R}{r_p}\right)$$

or
$$1+\frac{R}{2r_p}=\frac{\omega_0}{\omega}=1{\cdot}0005.$$

$$\frac{R}{r_p}=2\{1{\cdot}0005-1\}$$

$$=0{\cdot}001.$$

Hence $R=0{\cdot}001\times3{,}000=30\Omega.$

Mutual conductance

$$g_m=\frac{\mu}{r_p}=\frac{30}{30{,}000}=\mathbf{10^{-3}\ mho.}$$

Tuning capacitance

$$C=\frac{1}{\omega_0^2 L}=\frac{1}{(4\times10^5)^2\times40\times10^{-3}}$$

$$=\mathbf{156{\cdot}2\times10^{-12}\ Farad.}$$

We also know that $g_m=\dfrac{\mu RC}{\mu M-L}$,

$$10^{-3}=\frac{30\times30\times156{\cdot}2\times10^{-12}}{30\times M-40\times10^{-3}},$$

$$30M=40\times10^{-3}+\frac{40\times30\times156{\cdot}2\times10^{-12}}{10^{-3}},$$

$$30M = 40 \times 10^{-3} + 0{\cdot}1405 \times 10^{-3},$$
$$30M = 40{\cdot}1405 \times 10^{-3}$$
$$M = \frac{40{\cdot}1405 \times 10^{-3}}{30}$$
$$= \mathbf{1{\cdot}338 \times 10^{-3}\ H.}$$

3. *In a vacuum tube oscillator using a tuned anode circuit minimum mutual inductance between grid coil and the anode coil required for sustained oscillations is 200μH, when the load is equivalent to a resistance R=15 ohms in series with inductance L of the tuned circuit. If, however, the equivalent load resistance R is increased to 3·0Ω, the minimum value of mutual inductance M is required to be increased to 350μH. The electron tube used has amplification μ=12 and plate resistance, r_p=24,000 ohms. Find the inductance L and capacitance, C, of tuned plate circuit*

We know that $g_m = \frac{\mu}{r_p} = \frac{12}{24{,}000} = 3 \times 10^{-4}$ mho

and $$g_m = \frac{\mu\ RC}{\mu M - L}$$

or $$g_m L + \mu RC = g_m \mu M.$$

In first case,

$$5 \times 10^{-4}\ L + 12 \times 15C = 5 \times 10^{-4} \times 12 \times 200 \times 10^{-4}$$

or $$L + (36 \times 10^4)\ C = 24 \times 10^{-4}. \qquad \text{...(i)}$$

In second case,

$$5 \times 10^{-4}\ L + 12 \times 30C = 5 \times 10^{-4} \times 12 \times 350 \times 10^{-6}$$

or $$L + (72 \times 10^4)\ C = 42 \times 10^{-4}. \qquad \text{..(ii)}$$

Subtracting equation (i) from (ii), we get

$$36 \times 10^4 C = 18 \times 10^{-4},$$

or $$C = \frac{18 \times 10^{-4}}{36 \times 10^{-4}} = \mathbf{5000 \times 10^{-12}\ Farad.}$$

Substituting the value of C in equation (i), we get

$$\mathbf{L = 6 \times 10^{-6}\ Henry.}$$

4. *In a Hartley oscillator, the two sections of the coil have inductances of 90 mH and 30 mH, the latter being in the grid circuit. The tuning capacitance is of value 100 pf and the amplification factor is of the value 20. Find the approximate value of critical mutual inductance for maintaining oscillations. Also find the value of frequency of oscillations in that case.*

We know that $\mu = \left(\frac{L_1 + M}{L_2 + M}\right)$

or $$20 = \frac{(90 + M) \times 10^{-3}}{(30 + M) \times 10^{-3}}.$$

Solving it we have

$$M = -27\ mH.$$

Again
$$f=\frac{1}{2\pi}\frac{1}{\sqrt{[C(L_1+L_2-2M)]}}$$
$$=\frac{1}{2\pi\sqrt{[100\times10^{-12}\ (90+30-2\times27)\times10^{-3}]}}$$
$$=\frac{1}{2\pi}(17{\cdot}4\times10^{-12})^{-1/2}$$
$$=38\times10^3\ \text{c/s.}$$
$$=38\ \text{kc/s.}$$

5. ***In a resistance-capacitance tuned oscillator the regenerative network parameters are***

$R_1=4\times10^4$ *ohms*, $R_2=4\times10^4$ *ohms*,
$C_1=400$ *P.F.*, $C_2=100$ *pf.*

Determine

(*a*) *the frequency of oscillations, and*
(*b*) *voltage amplification to maintain oscillations.*

(*a*) We know that in case of RC tuned oscillators
$$f=\frac{1}{2\pi\sqrt{(K_1R_2C_1C_2)}}$$
$$=\frac{1}{2\pi\sqrt{(4\times10^4\times4\times10^4\times400\times100\times10^{-24})}}$$
$$=\frac{1}{2\pi\times8\times\sqrt{10^{-12}}}=\frac{\sqrt{10^{12}}}{16\pi}$$
$$=1{\cdot}9\times10^4\ \text{c/s.}$$

(*b*) Required gain to maintain oscillations is
$$=1+\frac{R_1}{R_2}+\frac{C_2}{C_1}$$
$$=1+1+4$$
$$=6.$$

6. ***In a phase shift oscillator the three networks are identical i.e.*** $R_1=R_2=R_3=R=6\times10^4$ ***ohms and*** $C_1=C_2=C_3=C=9$ ***pf. Calculate the angular frequency of oscillations.***

The angular frequency of oscillations is given by
$$\omega=\frac{1}{\sqrt{6.CR}}$$
$$=\frac{1}{\sqrt{6\times9\times10^{-12}\times6\times10^4}}$$
$$=\frac{1}{2{\cdot}441\times54\times10^{-8}}=\frac{10^8}{132{\cdot}246}$$
$$=75{\cdot}7\times10^4\ \text{radians/sec.}$$

7. ***A certain X-cut crystal is resonant at 450 kc/s. For this frequency, its equivalent inductance is 3·65 Henrys, its equivalent capacitance is 0·034 μμF., and its equivalent resistance 9,040 ohms. What is the Q of the crystal ?***

We know that Q of the crystal is given by

$$Q=\frac{2\pi fL}{R}$$

$$=\frac{6\cdot28\times450{,}000\times3\cdot65}{9040}$$

$$=\mathbf{1141}.$$

8. *A double resonator klystron is operating at a frequency of 3,000 megacycles. The drift space $S=2$ cms. The beam current is 25 mA and the ratio $V_2/V_0=0\cdot3$. Determine (a) the d.c. accelerating voltage (b) buncher voltage and (c) the efficiency, for the maximum power output condition taking the integer value of $n=1$. Assume the beam coupling coefficient as unity.*

(*a*) We know that d.c. accelerating voltage V_0 is given by the relation

$$V_0=0\cdot284\times10^{-11}\left[\frac{\omega S}{2\pi n-(\pi/2)}\right]^2$$

$$=0\cdot284\times10^{-11}\left[\frac{2\pi\times3000\times10^6\times2\times10^{-2}}{\pi\,(2n-\frac{1}{2})}\right]^2$$

where S is in metres.

$$=0\cdot284\times10^{-11}\times\left[\frac{12}{(3/2)}\right]^2\times10^{14}$$

$$=284\times64$$

$$=\mathbf{18176\ volts.}$$

(*b*) We know that

$$\frac{V_1}{V_0}=\frac{3\cdot68}{2n\pi-(\pi/2)}$$

or

$$V_1=\left[\frac{3\cdot68}{2\pi n-(\pi/2)}\right]\times V_0$$

$$=0\cdot78\times18176$$

$$=\mathbf{14177\cdot28\ volts.}$$

(*c*) $P_{output}=I_0V_2J_1(x)=25\times10^{-3}\times V_2\times0\cdot58,$

$P_{input}=I_0V_0=25\times10^{-3}\times V_0.$

$$\therefore\quad\eta=\frac{P_{output}}{P_{input}}=\frac{25\times10^{-3}}{25\times10^{-3}}\times\frac{V_2}{V_0}\times0\cdot58\times100$$

$$=1\times0\cdot3\times0\cdot58\times100$$

$$=\mathbf{17\cdot4\%}.$$

9. *Using equation* $\dfrac{V_0}{(V_R-V_0)^2}=\dfrac{[2n\pi-(\pi/2)]^2\,e}{(8\omega^2S^2m)}$ *derive an expression for* $\dfrac{df}{dV_R}$ *for reflex klystron. Assume that* $V_0=350$ *volt,* $S=0\cdot5$ *cm,* $f=5{,}000$ *megacycles, and* $n=4$. *Compute the frequency change for one volt change in reflector potential.*

The given expression is

$$\frac{V_0}{(V_R-V_0)^2}=\frac{[2\pi n-(\pi/2)]^2\, e}{8\omega^2 S^2 m}$$

From this expression, we have

$$\omega^2=\frac{[2n\pi-(\pi/2)]^2\, e}{8S^2m}\times\frac{(V_R-V_0)^2}{V_0}$$

or

$$\omega=\frac{[2n\pi-(\pi/2)]}{2\sqrt{2}S}\times\sqrt{\left(\frac{e}{m}\right)}\times\frac{(V_R-V_0)}{(V_0)^{1/2}}$$

$$\therefore\quad \frac{df}{dV_R}=\frac{[2n\pi-(\pi/2)]}{2\pi.2\sqrt{2}S}\times\sqrt{\left(\frac{e}{m}\right)}\times\frac{1}{\sqrt{V_0}}$$

According to the problem

$n=4$, $e/m=1{\cdot}76\times10^{11}$ e.m.u.

$V_0=350$ volts, $L=0{\cdot}5\times10^{-2}$ metres.

$$df=\frac{[8-(1/2)]\sqrt{(1{\cdot}76\times10^{11})}}{4\sqrt{2}\times0{\cdot}5\times10^{-2}\times\sqrt{350}}$$

$$=\mathbf{5{\cdot}928\ mc/s\ per\ unit\ reflector\ voltage.}$$

EXERCISES AND PROBLEMS

1. Define an oscillator as used in radio work. What is its importance ?
2. Draw the circuit diagram and explain the working of tuned anode oscillator. Using Barkhausen criterion, derive expression for g_m for sustained oscillations.
3. Derive condition for sustained oscillations in a tuned grid oscillator.
4. Draw a circuit diagram of a series fed, tuned anode Hartley oscillator and explain its action.
5. Draw a circuit diagram of Colpitt's oscillator and explain its action.
6. Deduce the frequency of oscillation for the following oscillators :
   (a) Dynatron oscillator,
   (b) phase shift oscillator.
   (c) Resistance-capacitance oscillators.
7. Explain how quartz crystal oscillator can be used in conjunction with a valve to generate high frequency currents. Discuss the advantages, difficulties and applications of such a generator.
8. Deduce the condition of maintenance of oscillations in a tuned plate oscillator. Describe and explain how a quartz crystal is used in the oscillator circuit for stabilising the frequency of the oscillation. Discuss the main cuts of the quartz crystal.
9. What limits the wavelength of the oscillation in a regenerative oscillator? Deduce expression for the wavelength of the Barkhausen Kurtz type of oscillations in a triode with cylindrical electrode.
10. Deduce the condition of maintenance of electrical oscillations in any standard type of oscillator. What precautions are described for the stability of the frequency and amplitude of oscillations ?

11. Draw a circuit diagram of multivibrator and explain how it works ? Show the typical waveform of a multivibrator output.
12. Give an account of the difficulties involved in producing U. H. F. oscillations by conventional feed-back methods using a triode. Explain, with the help of net diagrams, the mechanism of energy generation in a magnetron oscillating in the π mode ?
13. What is meant by the transit time effect? What are some of the common methods of overcoming this effect in an ordinary tube ?
14. Discuss why is it not possible to generate oscillations of frequency 3000 mc/s. with the help of conventional triodes.
    Draw a neat diagram of reflex klystron and explain its working and the phenomenon of a bunching of electrons.
15. What is meant by transit time effect ?
    Discuss why the conventional oscillator circuits cannot be used for generation of microwaves ?
    Show how oscillations are generated by a reflex klystron ?
16. Derive expressions for the output and efficiency of a double resonator klystron. In what respects do the double resonators klystron differ from the reflex klystron ?
17. Write short notes on :
    (i) Travelling wave tube,
    (ii) Electron wave tube, and
    (iii) Backward wave tube.

16 TRANSISTOR AMPLIFIERS

Transistor cannot be a direct substitute for vacuum tube. A vacuum tube is a voltage operated device in which the input voltage controls the output current or voltage or both. This type of amplifier works best with a constant voltage source. On the other hand, transistor is a current operated device in which the input current controls the output current. This type of amplifier works best with a constant current source.

The second difference can be made on the basis of isolation of input and output circuits. In transistor, these two circuits are not isolated ; therefore, output circuit parameters will affect the input circuit parameters and *vice versa*.

Third difference can be made with reference to output and input impedances. In vacuum tube circuits both these impedances are sufficiently high, while transistor circuits generally have a low to medium input impedance and moderate to high output impedance. The current flowing through these impedances determines the voltage or power gain of a transistor amplifier circuit. There are three basic types of transistor amplifier circuits :

(*i*) grounded emitter,

(*ii*) grounded base, and

(*iii*) grounded collector.

Choice of Parameters :

We have described Z, Y and h parameters in chapter 6. Z parameters require measurements under open circuit conditions, while Y parameter under short circuit conditions. The hybrid parameter h_{11} and h_{21} are measured with output short circuited, and h_{12} and h_{22} with input open circuited. It is convenient to short circuit the high impedance output of the transistor with a capacitor and to open circuit the low impedance input with an inductor. Due to this fact, at low frequencies, the hybrid parameters are more popular. The circuit analysis is easier with Y parameters at high frequencies. The Z parameters are most useful in describing low impedance devices. At low frequencies the Z parameters are mostly resistive so we can write $Z_{11}=r_{11}$, $Z_{12}=r_{12}$, $Z_{21}=r_{21}$ and $Z_{22}=r_{22}$. The parameters r_{11}, r_{12}, r_{21}, and r_{22} are known as r parameters.

16·1. EQUIVALENT CIRCUIT OF A TRANSISTOR :

We have discussed the equivalent circuit of a vacuum tube in two forms—one constant voltage generator and the other constant current generator. Here too a transistor can be replaced by an equivalent circuit.

An equivalent circuit can be traced simply by considering the three resistances r_b, r_e and r_o as shown in fig. 1 (*a*). But since a resistive network is attenuator, it is unable to amplify the input signal while the transistor do amplify. It shows that this passive circuit is not adequate for the transistor. A transistor must therefore, be represented by an active circuit containing one generator atleast with the circuit of fig. 1 (*a*). The active circuits are shown in fig. 1 (*b*) and fig. 1 (*c*).

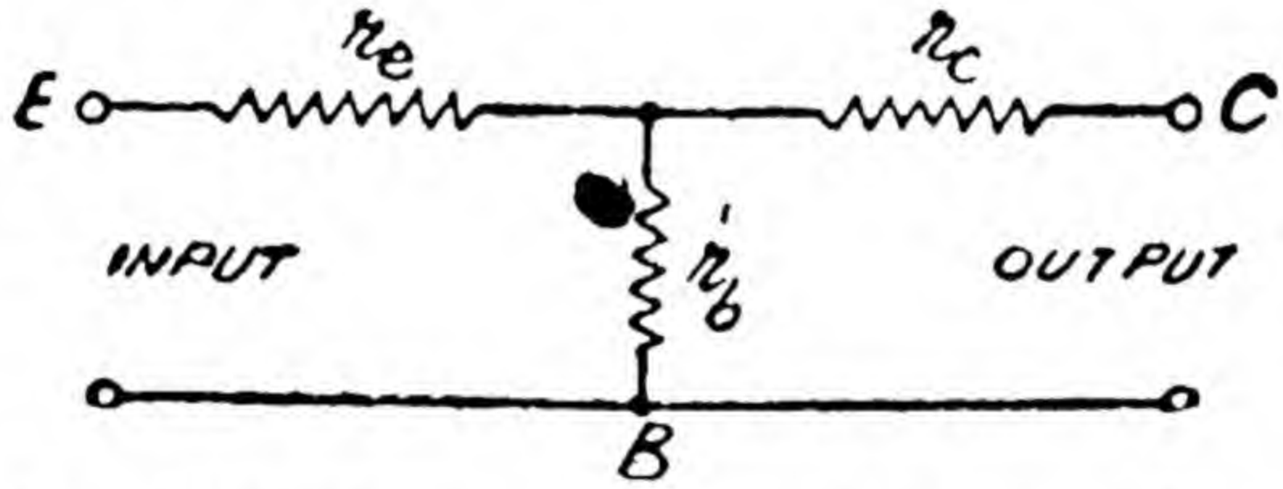
From fig. 1 (*b*), numerical value of current gain is

$$\left|\frac{i_o}{i_e}\right|_{output\ voltage=0} = \frac{r_m}{r_o},$$

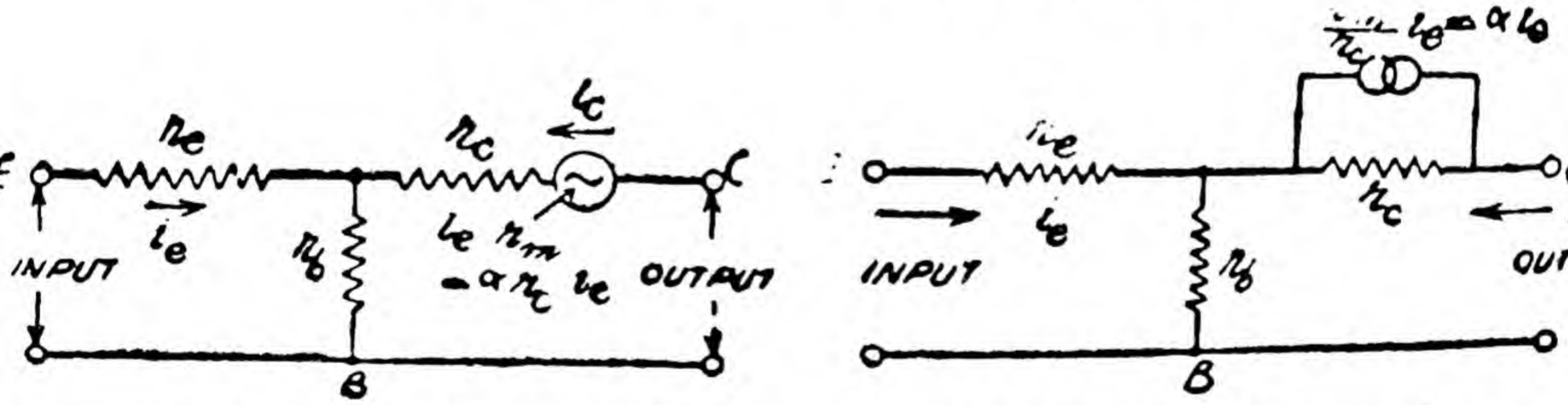
where r_m is the mutual resistance of the system. Above quantity is termed as alpha (α) for the transistor, *i.e*,

$$\alpha = \frac{r_m}{r_c}$$

$$\therefore \quad r_m = \alpha r_c.$$



(*a*) Preliminary equivalent circuit of a transistor.



(*b*) Constant voltage source form (*c*) Constant current source form

Fig. 1. Equivalent circuit of a transistor in common base connection

This generator can be shown as a constant current source fig. 1 (*c*), connected in parallel with the collector resistance r_o. The current thus applied is $\left(\frac{r_m i_e}{r_o}\right)$ or αi_e. In fig. 1 (*b*) generator is represented as constant voltage source connected in series with

r_0. This circuit will be adopted widely in further calcuations for output resistance, input resistance, voltage gain and current gain of the transistor in its three configurations.

16·2. GROUNDED EMITTER CIRCUIT :

The circuit of *PNP* type transistor with emitter grounded is shown in fig. 2 (*a*) and 2 (*b*).

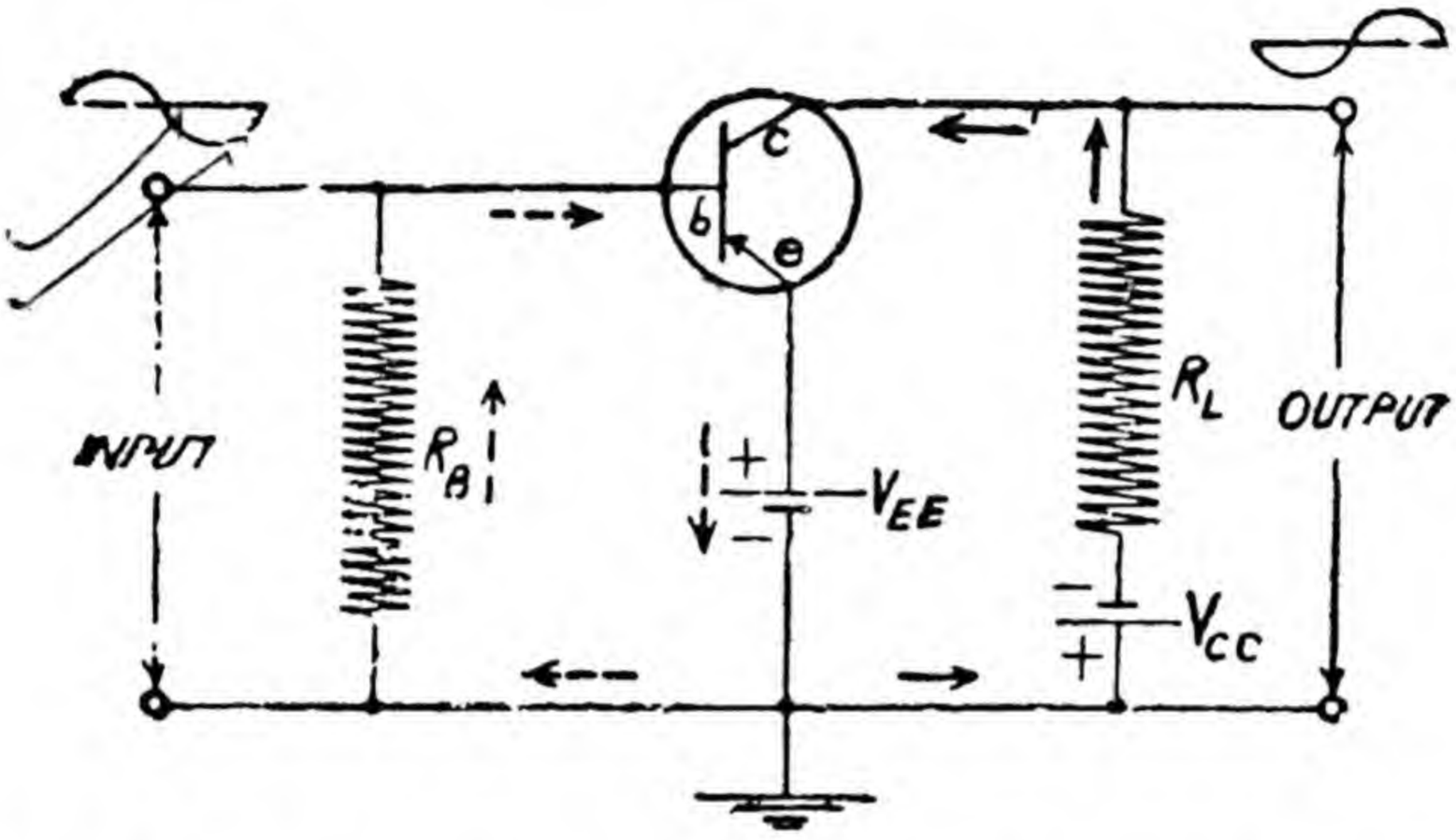


Fig. 2. (*a*) *PNP* grounded emitter trans stor-amplifier circuit.

Base is a driven element and the input signal is applied between the base and emitter. Output signal is drawn between collector and emitter. The emitter is biased in a forward direction by the voltage V_{EE} and the colleetor is biased in a reverse direction by the voltage V_{CC}.

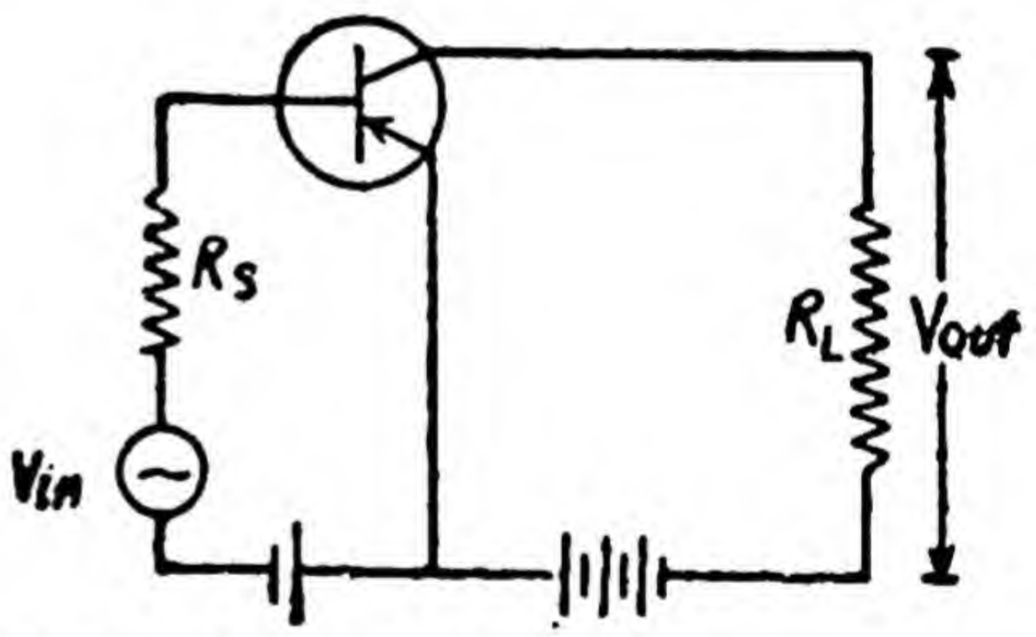


Fig 2. (*b*). Grounded emitter amplifier.

We have seen in chapter 6 that a small change in base current, due to input signal produces a large change in collector current which is responsible for the amplifying action of a grounded emitter transistor.

Suppose a positive signal is applied; then its polarity will be opposite to that of V_{EE}. Consequently, bias voltage V_{EE} decreases and causes a decrease in emitter and base currents. The decrease in emitter and base currents produces a decrease in the collector current with a corresponding decrease in the output voltage developed across R_L. As a result, the potential at the collector, with respect to ground will be negative. By the same reasoning a negative input signal will increase output voltage. Thus the input and output signals are 180° out of phase.

Now we shall proceed to calculate *input resistance*, *output resistance*, *voltage gain and current gain* of the transistor in the amplifying circuit.

Mathematical analysis of common emitter amplifier using r parameters :

Figure 3 (*a*) shows the one voltage generator equivalent circuit of a common emitter transistor treated as a four pole in *r* parameter, while 3 (*b*) shows the one voltage generator equivalent circuit derived on the basis of processes involved in the transistor.

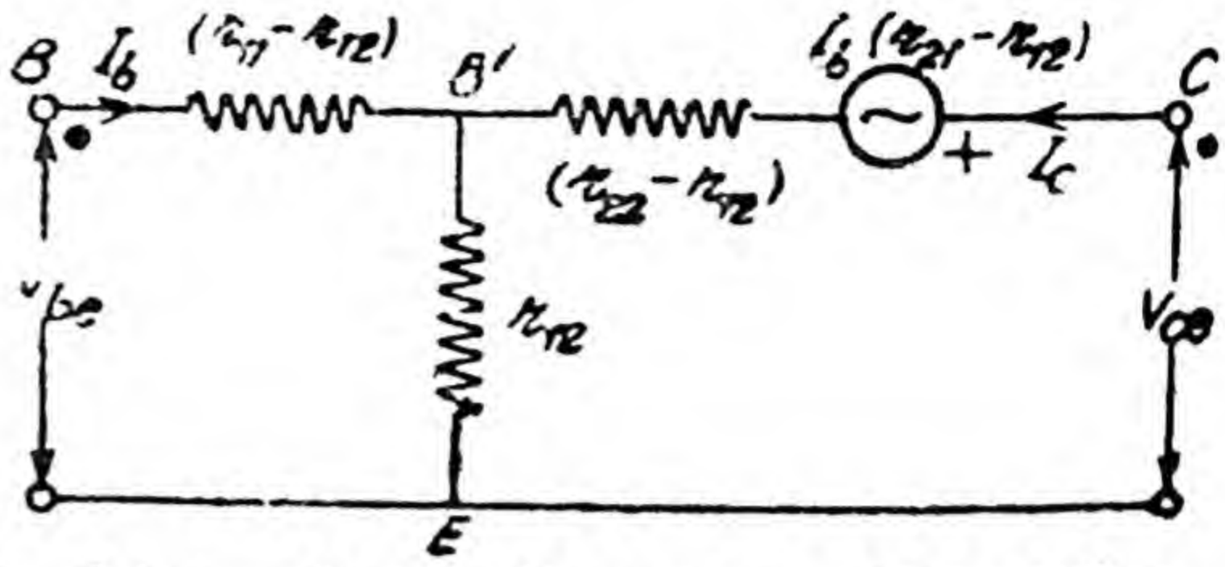


Fig. 3 (*a*). One voltage generator equivalent circuit of common emitter transistor.

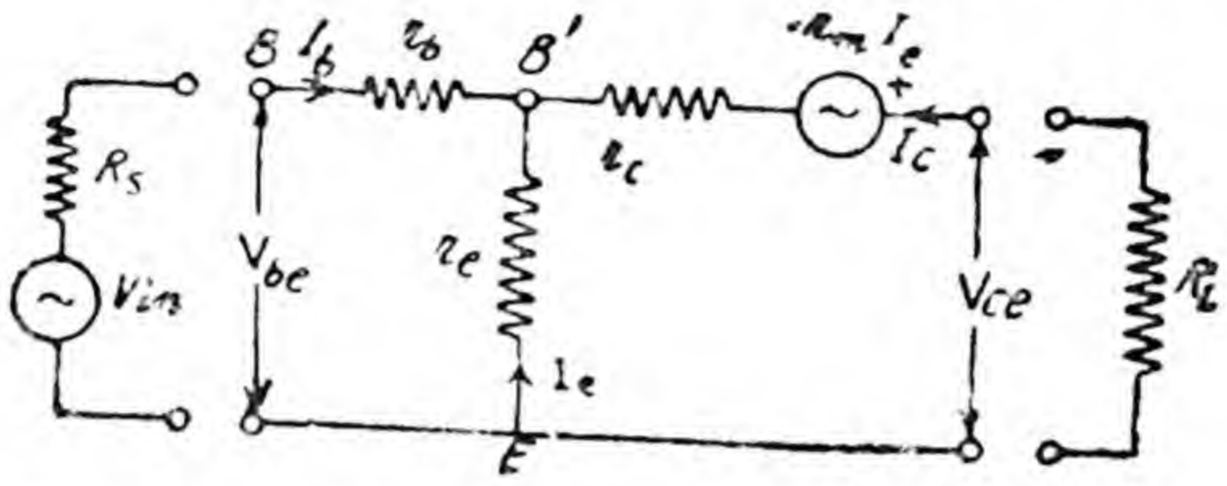


Fig. 3 (*b*). One voltage generator equivalent circuit of common emitter amplifier.

From figure 3 (*b*), at point *B'*, we have $I_b+I_e+I_c=0$ or $I_e=-I_b-I_c$. Here it is preferred to express I_c interms of base current. The circuit of figure 3 (*b*) is transformed to circuit shown in figure 3(*c*).

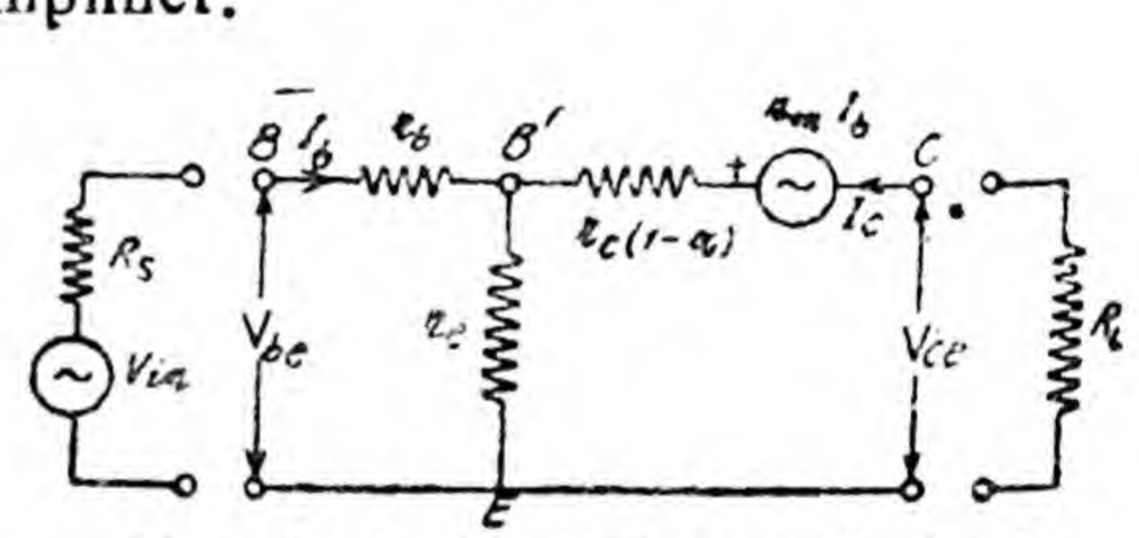


Fig. 3 (*c*). Modified equivalent circuit of figure 3 (*b*).

Network equations for figure 3 (*a*) are

$$V_{be}=I_b\,(r_{11}-r_{12})+(I_b+I_c)\,r_{12}=I_b\,r_{11}+I_c\,r_{12} \quad ...(1)$$

$$V_{ce}=I_b\,(r_{21}-r_{12})+I_c\,(r_{22}-r_{12})+(I_b+I_c)\,r_{12}=I_b\,r_{21}+I_c r_{22} \quad ...(2)$$

The network equations for figure 3 (*c*) are

$$V_{be}=I_b\,(r_b+r_e)+I_e\,r_e \quad ...(3)$$

$$V_{ce}=I_c\,r_c\,(1-\alpha)-r_m I_b+(I_b+I_e)\,r_e$$

or

$$V_{ce}=(r_c-r_m)+I_c\,[r_e+r_c\,(1-\alpha)] \quad ...(4)$$

Comparing equations (1) and (2) with equations (3) and (4) respectively, we get

$$r_{11}=r_b+r_e,\ r_{12}=r_e,\ r_{21}=r_e-r_m$$

and

$$r_{22}=r_e+r_c\,(1-\alpha)$$

Thus we can write

$$\begin{bmatrix} V_{be} \\ V_{ce} \end{bmatrix}=\begin{bmatrix} r_b+r_e, & r_e \\ r_c-r_m, & r_e+r_c\,(1-\alpha) \end{bmatrix}\begin{bmatrix} I_b \\ I_c \end{bmatrix} \quad ...(5)$$

In chapter 6 we have derived the following results,

Input impedance $Z_i=Z_{11}-\dfrac{Z_{12}Z_{21}}{Z_{22}+Z_L}$

Output impedance $Z_0 = Z_{22} - \frac{Z_{12}Z_{21}}{Z_{11}+Z_S}$ ...(6)

Voltage gain $A_v = \frac{Z_L Z_{21}}{(Z_{22}+Z_L) Z_i}$

Current gain $A_i = -\frac{Z_{21}}{Z_{22}+R_L}$,

where symbols have their usual meaning.

In this case, we shall find the following parameters.

(1) Input impedance $r_i = r_{11} - \frac{r_{12}\, r_{21}}{r_{22}+R_L}$

$$= (r_b + r_e) - \frac{r_e\,(r_e - r_m)}{[r_e + r_c\,(1-\alpha)] + R_L}$$

$$= r_b + r_e \left[1 - \frac{(r_e - r_m)}{\{r_e + r_c\,(1-\alpha)\} + R_L} \right]$$

$$= r_b + r_e \left[\frac{R_L + r_c}{R_L + r_e + r_c\,(1-\alpha)} \right]$$

In practice, r_e is much smaller as compared to other terms and is, therefore, negligible, *i.e.*

$$r_i = r_b + r_e \left[\frac{R_L + r_c}{R_L + r_c\,(1-\alpha)} \right]$$

Assuming R_L to be very small as compared to r_c and $r_c\,(1-\alpha)$, we get

$$r_i = r_b + r_e \frac{r_c}{r_c\,(1-\alpha)} = r_b + \frac{r_e}{(1-\alpha)} \qquad \text{...(a)}$$

If R_L is large as compared to r_c, then $r_i = r_b + r_e$. Variation of input resistance with load resistance is shown in figure 4. It is observed from the figure that input resistance is high for low values of load resistance and goes on decreasing with the increase of load resistance.

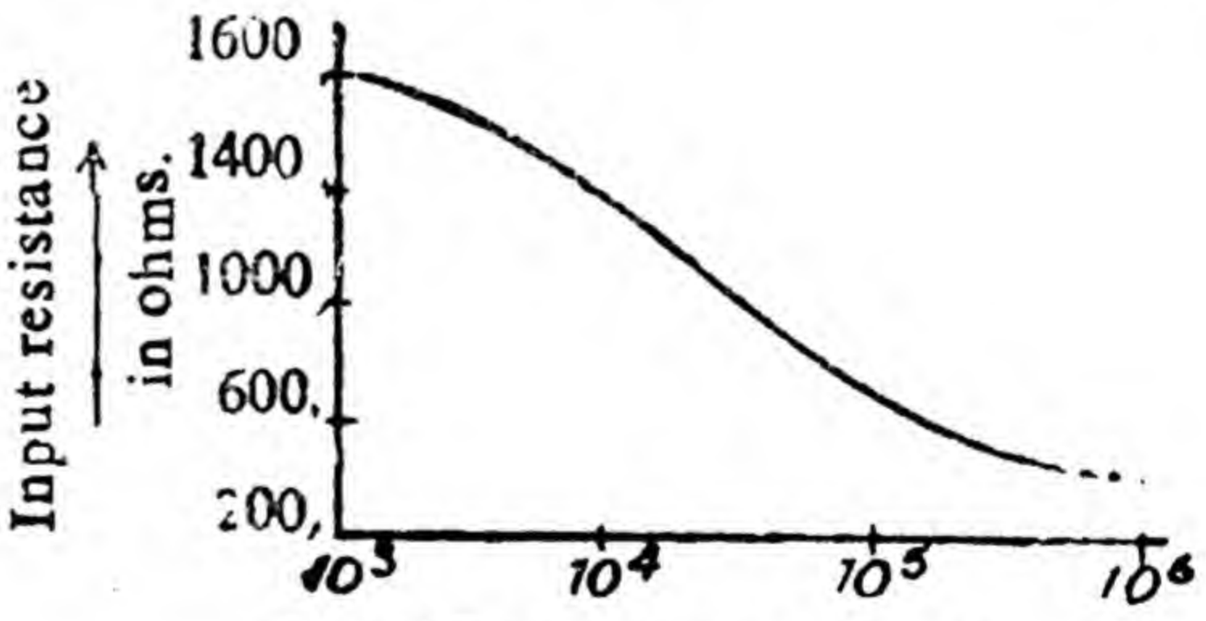


Fig. 4. Showing variation of input resistance with load.

(2) Output resistance $r_0 = r_{22} - \frac{r_{12}\, r_{21}}{r_{11}+R_S}$

$$= [r_e + r_c\,(1-\alpha)] - \frac{r_e\,(r_e - r_m)}{r_b + r_e + R_S}$$

$$= r_c\,(1-\alpha) + \frac{(r_b + r_m + R_S) r_e}{r_b + r_e + R_S}$$

In general $r_b \ll r_m$ and $R_S \ll r_m$, so that

$$r_0 \approx r_e\ (1-\alpha)+\frac{r_e r_m}{r_b+r_e+R_S} \qquad \ldots(b)$$

(3) Voltage gain $A_v=\dfrac{R_L\ r_{21}}{(r_{22}+R_L)\ r_i}$

$$=\frac{R_L\ (r_e-r_m)}{(r_e+r_c-r_m+R_L)\left[r_b+\dfrac{r_e(r_c+R_L)}{r_e+r_c(1-\alpha)+R_L}\right]}$$

$$=\frac{R_L}{r_b+\left[\dfrac{(r_b+r_e)(r_c+R_L)}{(r_e-r_m)}\right]}$$

As $r_b \ll r_m$ and $R_L \ll r_c$, we have

$$A_v=\frac{-\alpha R_L}{r_e+r_b(1-\alpha)} \qquad \ldots(c)$$

Figure 3 shows the variation of voltage gain with load.

(4) Current gain

$$A_i=\frac{-r_{21}}{r_{22}+R_L}$$

$$=-\frac{(r_e-r_m)}{r_e+r_c\ (1-\alpha)+R_L}$$

$$=-\frac{(r_e-r_m)}{r_e+r_c-r_m+R_L}$$

$$=-\frac{1}{1+\left(\dfrac{r_c+R_L}{r_e-r_m}\right)}$$

In practice $r_c \gg R_L$ and $r_e \ll r_m$

$$\therefore \quad A_i \approx \frac{-1}{1-(r_c/r_m)} \approx -\left(\frac{\alpha}{1-\alpha}\right) \qquad \ldots(d)$$

(5) Power gain

$$=A_v \times A_i \approx \frac{-\alpha}{(1-\alpha)} \times \left[\frac{-\alpha\ R_L}{r_e+r_b\ (1-\alpha)}\right]$$

$$\approx \frac{\alpha^2 R_L}{(1-\alpha)r_e+r_b\ (1-\alpha)^2} \qquad \ldots(e)$$

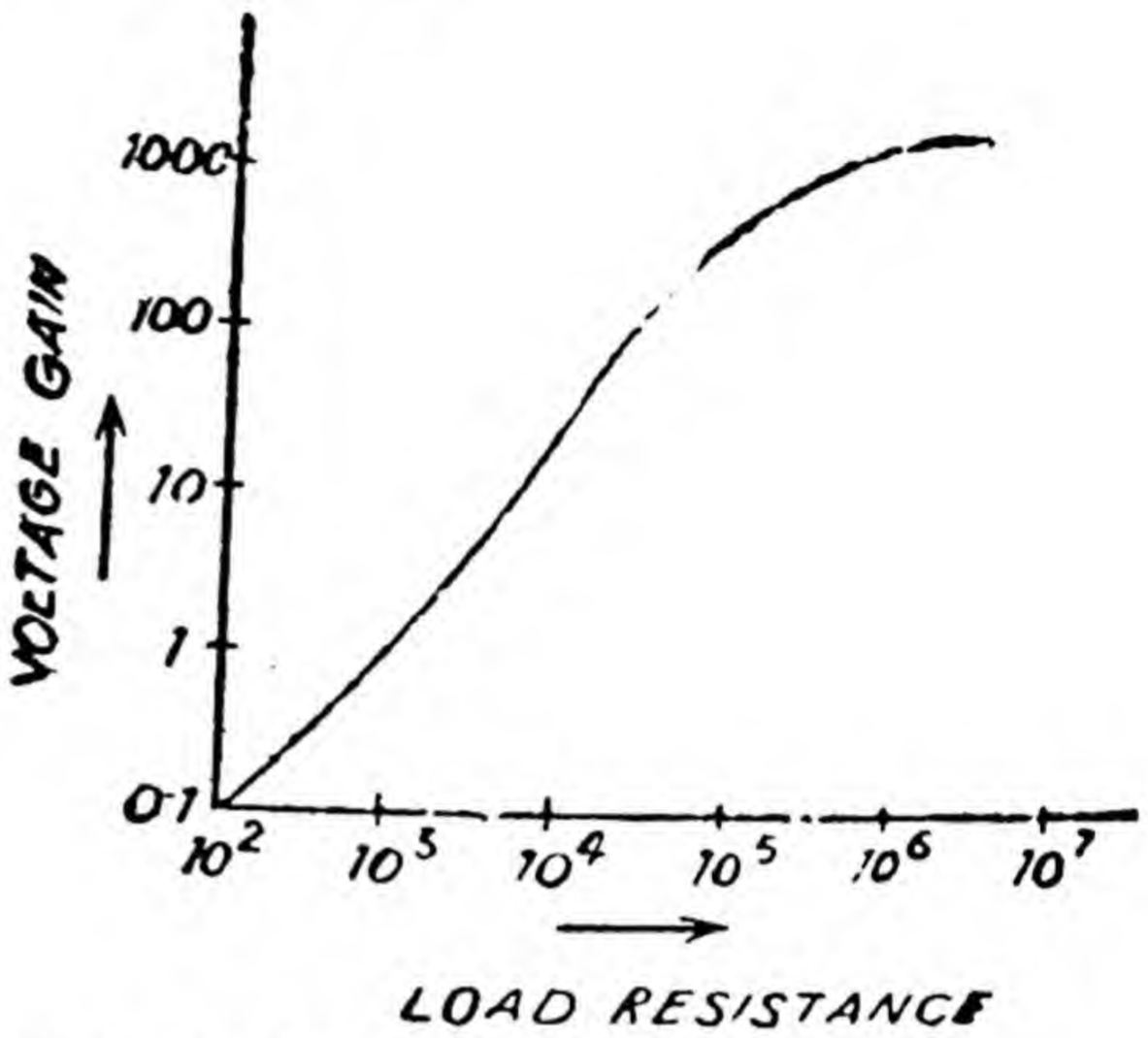


Fig. 5. Variation of voltage gain with load resistance.

COMMON EMITTER TRANSISTOR AMPLIFIER HYBRID EQUATIONS :

Hybrid equivalent circuit for a *CE* transistor is shown in fig. 6 (*a*). Refer to art. 6·39, where we have shown that, with equivalent circuit shown in fig., the voltage and current equations are

$$v_1=h_{11}\ \ i_1+h_{12}\ \ v_2 \qquad \ldots(7)$$

$$i_2=h_{21}\ \ i_1+h_{22}\ \ v_2 \qquad \ldots(8)$$

Here currents i_1, i_2 and voltages v_1, v_2 are incremental quantities. If the source voltage is assumed to be sinusoidal, these quantities are replaced by I_1, I_2 and V_1, V_2. The above equations (7) and (8) can be written as

$$v_1 = h_{ie}\, i_1 + h_{re}\, v_2 \qquad \ldots(7a)$$

and

$$i_2 = h_{fe}\, i_1 + h_{oe}\, v_2 \qquad \ldots(8a)$$

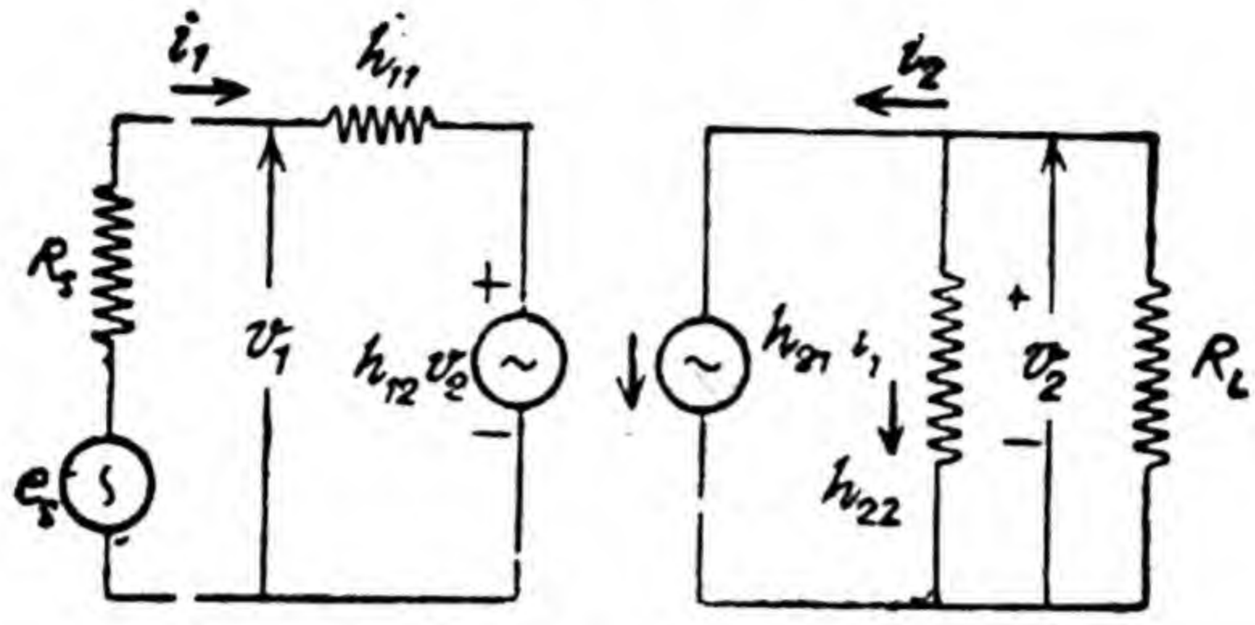


Fig. 6 (*a*).

The second two equations needed are

$$v_1 = e_S - i_1\, R_S, \qquad \ldots(9)$$

$$v_2 = -i_2\, R_L. \qquad \ldots(10)$$

(1) **Current gain :** Let us eliminate i_2 from equations (8) and (10). This gives

$$0 = h_{21}\, i_1 + v_2 \left(h_{22} + \frac{1}{R_L} \right)$$

or

$$v_2 = -\frac{h_{21}\, i_1}{h_{22} + \frac{1}{R_L}}$$

or

$$v_2 = -i_1 \frac{h_{21}\, R_L}{h_{22}\, R_L + 1}. \qquad \ldots(11)$$

Putting for v_2 from equation (10), we get

$$-i_2\, R_L = -i_1 \frac{h_{21}\, R_L}{h_{22}\, R_L + 1}$$

or

$$A_{ie} = \frac{-i_2}{i_1} = \frac{-h_{21}}{h_{22}\, R_L + 1} = \frac{-h_{fe}}{h_{oe}\, R_L + 1}. \qquad \ldots(12)$$

(2) **Input Resistance :** Putting equation (11) in equation (7), we get

$$v_1 = i_1 \left(h_{11} - h_{12} \frac{R_L\, h_{21}}{R_L\, h_{22} + 1} \right)$$

or

$$v_1 = i_1 \frac{h_{11}\, h_{22}\, R_L + h_{11} - h_{12}\, R_L\, h_{21}}{R_L\, h_{22} + 1} \qquad \ldots(13)$$

or

$$R_{ie} = \frac{v_1}{i_1} = \frac{R_L\,(h_{ie}\, h_{oe} - h_{re}\, h_{fe}) + h_{ie}}{h_{oe}\, R_L + 1}$$

$$=\frac{R_L\,\Delta+h_{ic}}{h_{oc}\,R_L+1}, \qquad \ldots(14)$$

where $\Delta=(h_{ic}\,h_{oc}-h_{rc}\,h_{fc})$.

Input resistance interms of current gain :

We know that

$$v_1=h_{ic}\,i_1+h_{rc}\,v_2$$

or $$v_1=h_{ic}\,i_1-h_{rc}\,i_2\,R_L \qquad \because \quad v_2=-i_2\,R_L$$

or $$v_1=h_{ic}\,i_1+A_{ic}\,i_1\,R_L\,h_{rc} \qquad \because \quad A_{ic}=-i_2/i_1$$

$$\therefore \qquad R_{ic}=\frac{v_1}{i_1}=h_{ic}+h_{rc}\,A_{ic}\,R_L \qquad \ldots(14a)$$

(3) **Voltage gain :** From equations (13) and (11), we get

$$A_{vc}=\frac{v_2}{v_1}=\frac{-h_{fc}\,R_L}{R_L\,(h_{ic}\,h_{oc}-h_{fc}\,h_{re})+h_{ic}} \qquad \ldots(15)$$

Voltage gain in terms of current gain :

$$A_{vc}=\frac{v_2}{v_1}=\frac{-i_2\,R_L}{v_1}=\frac{A_{ic}\,i_1\,R_L}{v_1}$$

or $$A_{vc}=\frac{A_{ic}\,R_L}{R_{ic}} \qquad \ldots(15a)$$

(4) **Power gain :** It is given by

$$A_{pc}=A_{ic}\,A_{vc}=\left(\frac{h_{21}}{h_{22}\,R_L+1}\right)\left(\frac{R_L\,h_{21}}{R_L\,\Delta+h_{11}}\right)$$

$$=\frac{h_{21}^2\,R_L}{(1+h_{22}\,R_L)\,(R_L\,\Delta+h_{11})}$$

$$=\frac{h_{fe}^2\,R_L}{(1+h_{oc}\,R_L)\,(R_L\,\Delta+h_{ic})}. \qquad \ldots(16)$$

(5) **Output resistance :** For the calculation of output resistance, we have to transfer the signal generator to the output terminals [fig. 6 (*b*)]. The equations are

$$0=(R_S+h_{11})\,i_1+h_{12}\,v_2,$$

$$i_2=h_{21}\,i_1+h_{22}\,v_2.$$

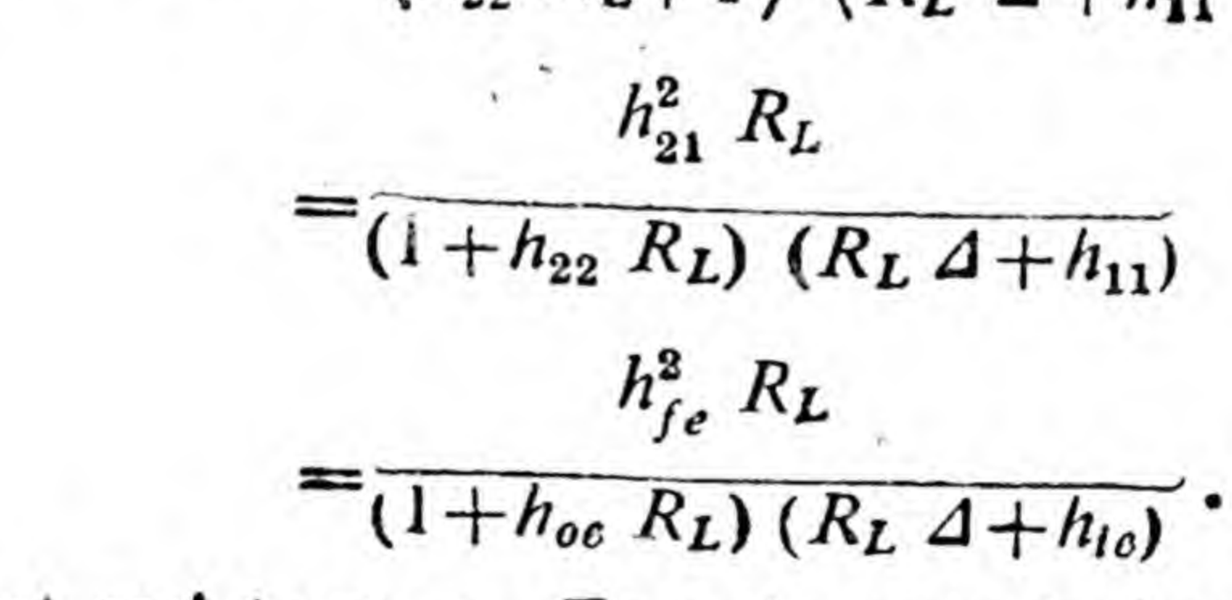


Fig. 6 (*b*).

Substituting i_1 from first equation into second, we get

$$i_2 = -\frac{h_{12}\, h_{21}\, v_2}{R_S + h_{11}} + h_{22}\, v_2$$

or
$$R_{oe} = \frac{v_2}{i_2} = \frac{R_S + h_{11}}{(R_S + h_{11})\, h_{22} - h_{12}\, h_{21}}$$
$$= \frac{R_S + h_{11}}{R_S\, h_{22} + \Delta}$$
$$= \frac{R_S + h_{ie}}{R_S\, h_{oe} + \Delta}. \qquad \ldots(17)$$

The output resistance can also be calculated in the following way :

The output admittance Y_{oe} is given by

$$Y_{oe} = \frac{i_2}{v_2} \text{ with } e_S = 0$$

But
$$i_2 = h_{fe}\, i_1 + h_{oe}\, v_2$$

$\therefore$
$$Y_{oe} = h_{fe}\frac{i_1}{v_2} + h_{oe}$$

From figure we have

$$R_S\, i_1 + h_{ie}\, i_1 + h_{re}\, v_2 = 0 \qquad (\because \quad e_S = 0)$$

or
$$\frac{i_1}{v_2} = -\frac{h_{re}}{h_{ie} + R_S}$$

$\therefore$
$$Y_{oe} = h_{oe} - \frac{h_{fe}\, h_{re}}{h_{ie} + R_S}$$

or
$$R_{oe} = [h_{oe} - h_{fe}\, h_{re}/h_{ie} + R_S]^{-1}$$

(6) Voltage gain considering source resistance R_S. The actual voltage applied to the input of the amplifier is e_S. Hence the overall voltage gain is given by

$$(A_{ve})_S = \frac{v_2}{e_S} = \frac{v_2}{v_1}.\frac{v_1}{e_S} = A_{ve}.\frac{v_1}{e_S} \qquad \ldots(18)$$

Corresponding to the input impedance R_{ie}, the value of v_1 is given by

$$v_1 = \frac{e_S.R_{ie}}{R_S. + R_{ie}} \qquad \ldots(19)$$

Hence
$$(A_{ve})_S = A_{ve}.\frac{R_{ie}}{R_S + R_{ie}} \qquad \ldots(20)$$

If $R_S = 0$, $(A_{ve})_S = A_{ve}$.

(7) Current gain considering the source resistance R_S. Let a current source i_S with internal resistance R_S be at the input side, then overall current gain is given by

$$(A_{ie})_S = \frac{i_2}{i_1}.\frac{i_1}{i_S} = A_{ie}.\frac{i_1}{i_S}$$

Now
$$i_1 = \frac{i_S\, R_S}{R_{ie} + R_S}$$

$$\therefore \qquad (A_{ie})_S = A_{ic} \cdot \frac{R_S}{R_S + R_{ie}}$$

If $R_S = \infty$, then $(A_{ie})_S = A_{ie}$.

The subscript e affixed to each parameter indicates that these parameters belong to a common emitter transistor.

Important formulae for common emitter transistor amplifier using h parameters :

Quantity	Expression
Current gain	$A_{ie} = -h_{fe}/(h_{oc} R_L + 1)$
Input resistance	$R_{ie} = h_{ie} + h_{re} A_{ie} R_L$
Voltage gain	$A_{ve} = A_{ie} R_L / R_{ie}$
Power gain	$A_{pe} = A_{ve} A_{ie}$
Output resistance	$R_{oe} = [h_{oe} - h_{fe} h_{re}/(h_{ie} + R_S)]^{-1}$

Example 1. *A common emitter transistor amplifier is driven by a voltage e_S of internal resistace $R_S = 800\,\Omega$. The load impedance is a resistor $R_L = 2000\Omega$. The parameters are : $h_{ie} = 1100\Omega$, $h_{re} = 2{\cdot}5 \times 10^{-4}$, $h_{fe} = 50$. $h_{oe} = 25\mu A/V$. Compute the current gain, voltage gain. overall current gain, overall voltage gain, input Resistance, output resistance and operating power.*

(i) Current gain $A_{ie} = \dfrac{-h_{fe}}{h_{oe} R_L + 1}$

$$= \frac{-50}{25 \times 10^{-6} . 2000 + 1} = -47{\cdot}6.$$

(ii) Input resistance $R_{ie} = h_{ie} + h_{re} A_{ie} R_L$

$$= 1100 + (2{\cdot}5 \times 10^{-4})\,(-47{\cdot}6)\,(2000)$$
$$= 1076\Omega$$

(iii) Voltage gain $A_{ve} = \dfrac{A_{ie} R_L}{R_{ie}}$

$$= \frac{-47{\cdot}6 \times 2000}{1076} = -88{\cdot}47.$$

(iv) Overall voltage gain $(A_{ve})_S = A_{ve} \dfrac{R_{ie}}{R_S + R_{ie}}$

$$= (-88{\cdot}47)\,\frac{1076}{800 + 1076}$$
$$= -50{\cdot}75.$$

(v) Overall current gain $(A_{ic})_S = A_{ic} \dfrac{R_S}{R_S + R_{ic}}$

$$= (-47{\cdot}6) \frac{800}{800+1076}$$

$$= -20{\cdot}30.$$

(vi) Output Resistance $R_{oc} = 1/Y_{oc}$

where $$Y_{oc} = h_{oc} - \frac{h_{fc}\, h_{nc}}{h_{ic} + R_S}$$

$$= (25 \times 10^{-6}) - \frac{50 \times 2{\cdot}5 \times 10^{-4}}{(1100+800)}$$

$$= 18{\cdot}42 \times 10^{-6}\ \Omega$$

$$\therefore \quad R_{oc} = \frac{1}{18{\cdot}42 \times 10^{-6}} = \frac{10^6}{18{\cdot}42}$$

$$= \mathbf{54{\cdot}3 \times 10^3}\ \Omega$$

(vii) Power gain $\rightleftharpoons A_{vc} . A_{ic} = (88{\cdot}47)\ (47{\cdot}6)$

$$= \mathbf{4211}.$$

16·3. COMMON BASE AMPLIFIER :

PNP transistor with grounded base is shown in fig. 7 (*a*). In this circuit, emitter is driven element and the input signal is applied between the emitter and base ; the output signal is produced between the collector and base. Thus the base is common to both the input and output circuits. The emitter is biased in forward direction by the voltage V_{BB} and the collector is biased in a reverse direction by the voltage V_{CC}. The input resistance of the transistor is very low while output resistance is very high.

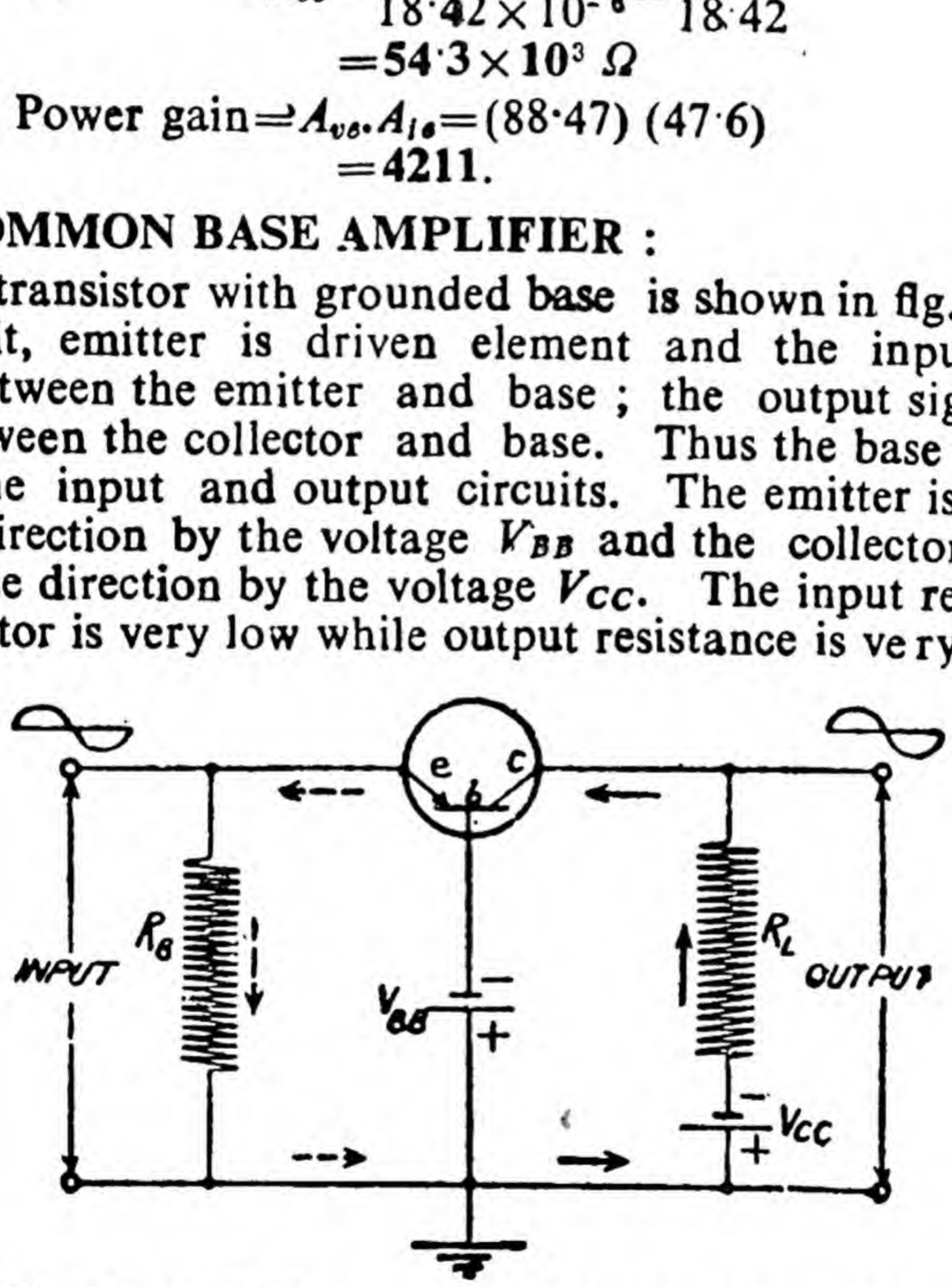


Fig. 7. (*a*) *PNP* grounded base transistor amplifier.

It is obvious from characteristics (chap. 6) that a small change in emitter current, due to input signal, produces large change in collector current and hence the greater output voltage. This explains the amplifying action of grounded base transistor.

When the input signal is positive, its polarity is the same as V_{BB}; consequently the bias voltage V_B will increase and causes an increase in emitter and base currents which, in turn, will

increase output voltage. Thus input and output signals are in the same phase.

Now we shall proceed to calculate the *input resistance*, *output resistance*, *voltage gain* and *current gain* of this amplifier circuit.

A simplified common base amplifier circuit is shown in figure 7 (*b*).

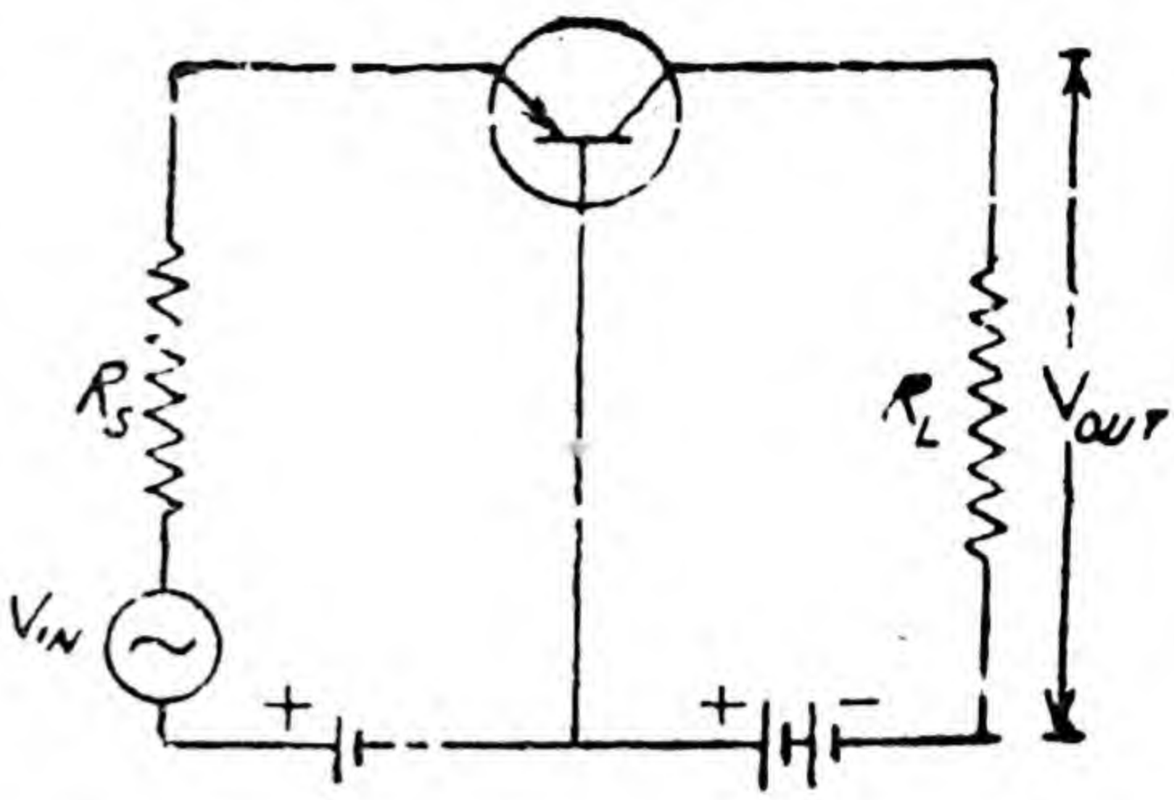


Fig. 7. (b) Common base amplifier.

Mathematical analysis of common base amplifier using r parameters :

Figure 8 (*a*) shows the one current generator equivalent of common base transistor using *r* parameters while figure 8 (*b*) shows the one current generator equivalent circuit derived on the basis of processes involved in the transistor.

From figure 8 (*a*), we have

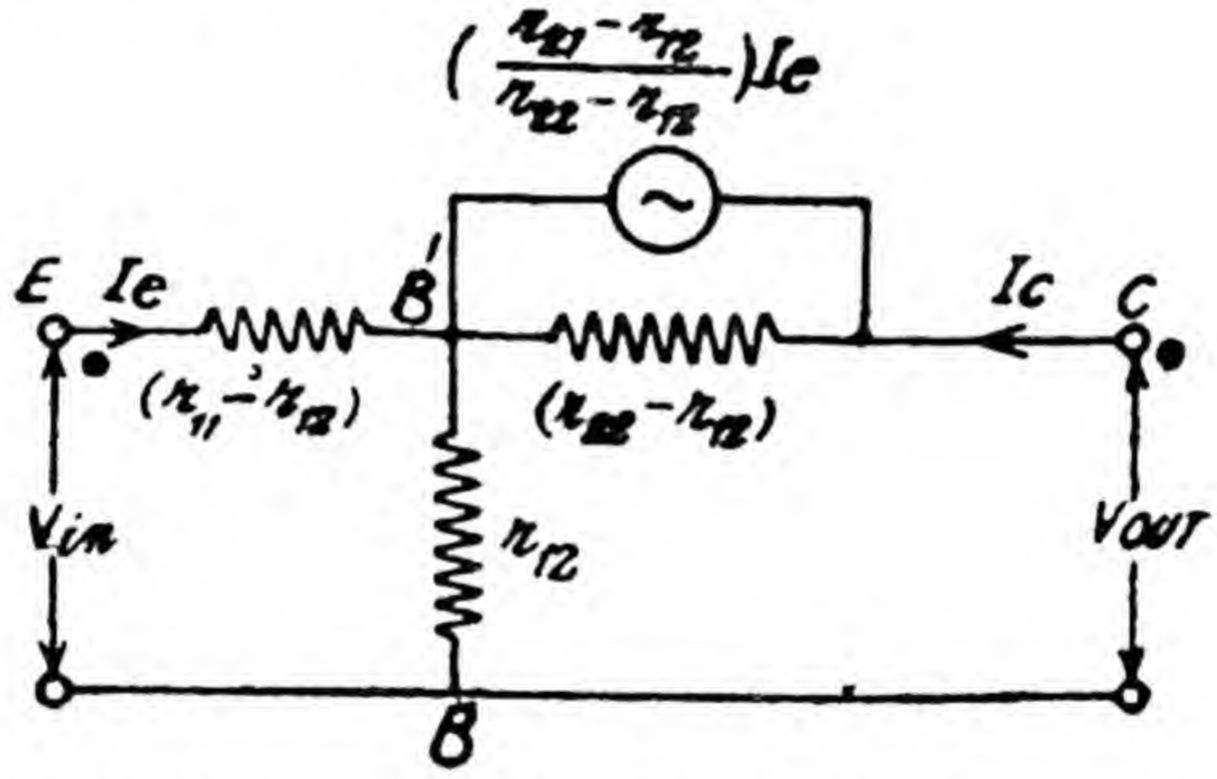


Fig. 8. (*a*) One current generator equivalent circuit of common base transistor as a four pole.

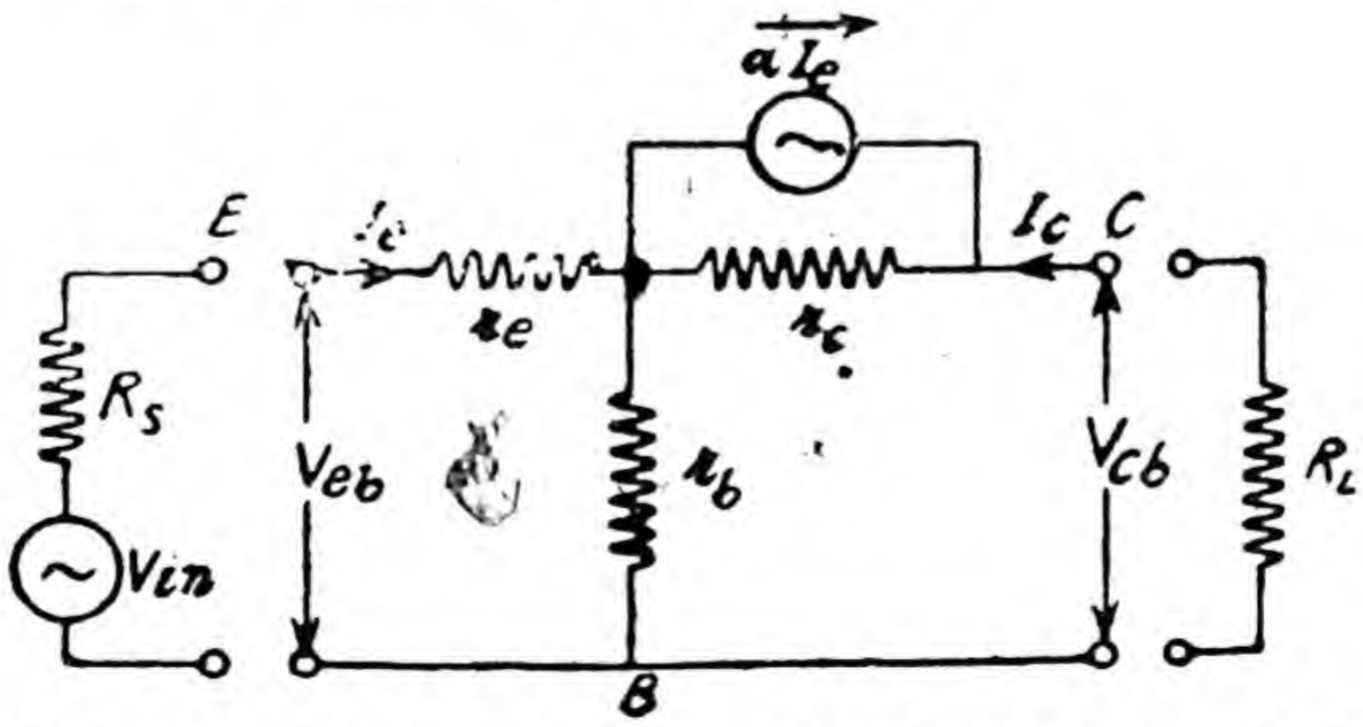


Fig. 8. (*b*) One current equivalent circuit of a common base transistor drived on the basis of processes involved in transistor.

$$V_{eb}=I_e\,(r_{11}-r_{12})+(I_e+I_c)\,r_{12}=I_e\,r_{11}+I_c\,r_{12}, \qquad \ldots(1)$$

$$V_{cb}=(I_e+I_c)\,r_{12}+(r_{22}-r_{12})\,I_c+I_c\,(r_{21}-r_{12})=I_e r_{21}+I_c r_{22}. \qquad \ldots(2)$$

Similarly from figure 8 (*b*), we have

$$V_{ob}=I_e(r_e+r_b)+I_c r_b, \qquad ...(3)$$

$$V_{ob}=I_e(r_b+r_m)+I_c(r_b)+r_c, \text{ where } r_m=\alpha r_c. \qquad ...(4)$$

Comparing equations (1) and (2) with equations (3) and (4) respectively, we get

$$r_{11}=r_e+r_b,\ r_{12}=r_b,\ r_{21}=r_b+r_m \text{ and } r_{22}=r_b+r_c. \qquad ...(5)$$

For accurate circuit considerations it is seen that

$$\alpha=\frac{r_m+r_b}{r_c+r_b}=\frac{r_{21}}{r_{22}}. \qquad ...(6)$$

In Chapter 6 we have derived the following results :

Input impedance $Z_i=Z_{11}-\dfrac{Z_{12}Z_{21}}{Z_{22}+R_L}$.

Output impedance $Z_0=Z_{22}-\dfrac{Z_{12}Z_{21}}{Z_{11}+Z_S}$. ...(7)

Voltage gain $A_v=\dfrac{Z_L Z_{21}}{(Z_{22}+Z_L)Z_i}$

and power gain $A_p=A_i^2\dfrac{Z_L}{Z_i}$,

where the symbols have their usual meaning.

(I) Input impedance $r_i=r_{11}-\dfrac{r_{12}r_{21}}{r_{22}+R_L}$

$$=(r_b+r_e)-\frac{r_b(r_b+r_m)}{r_b+r_c+R_L}$$

$$=r_e+\frac{r_b(r_c-r_m+R_L)}{r_b+r_c+R_L}$$

r_c and αr_c are both large compared to r_b and R_L is usually chosen to be small hence

$$r_i \simeq r_e+\frac{r_b(r_c-r_m)}{r_c} \approx r_e+r_b(1-\alpha). \qquad ...(a)$$

Variation of input resistance with load resistance is shown in figure 9.

It is obvious from figure that the input resistance increases with load—the increment being small for low values of load resistance while large for higher values.

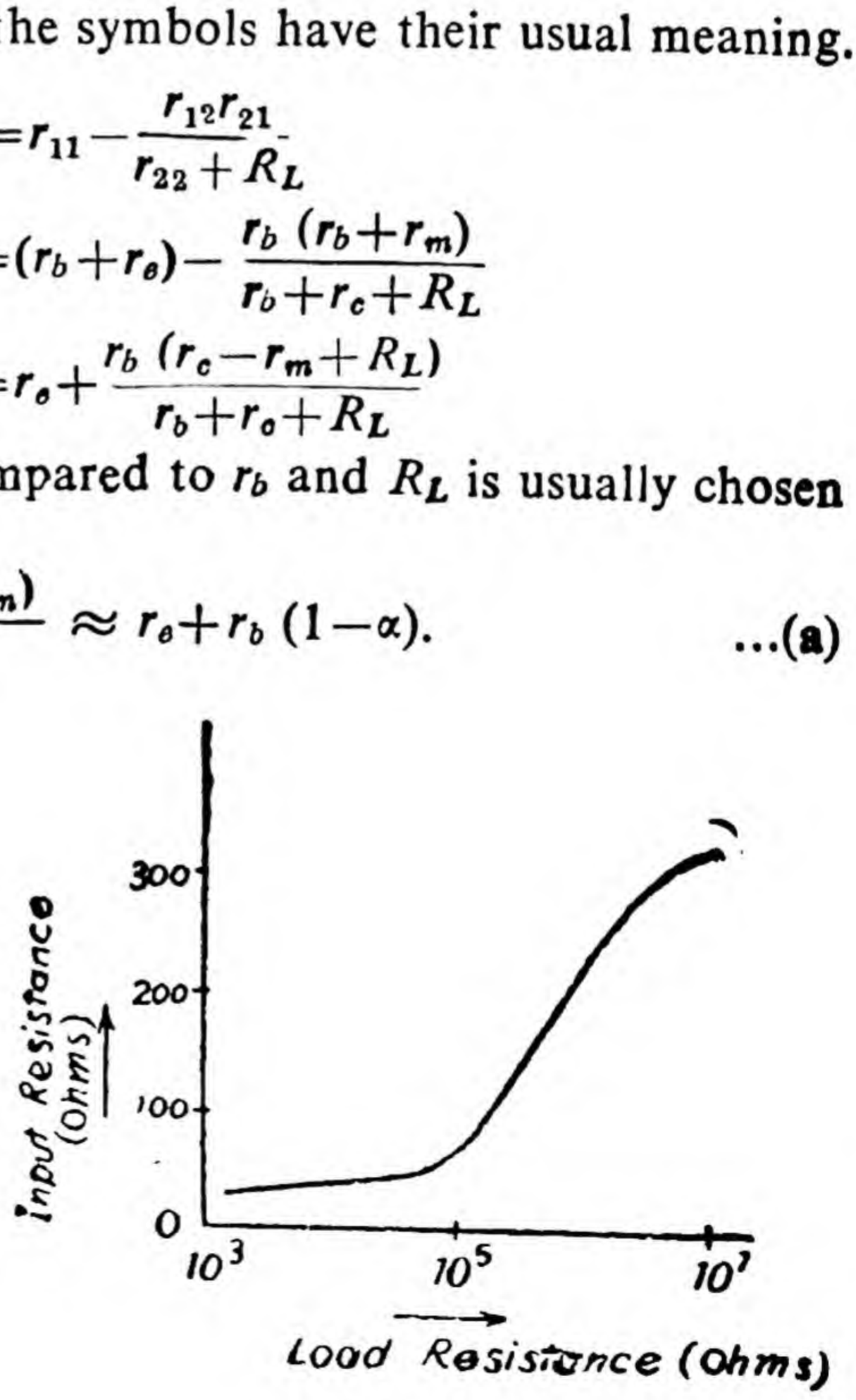


Fig. 9.

(2) Output resistance $r_0 = r_{22} - \dfrac{r_{12} r_{21}}{r_{11} + R_S}$

$$= (r_b + r_c) - \frac{r_b (r_b + r_m)}{(r_b + r_c) + R_S}$$

This third term is small as compared with r_c, hence

$$r_0 \approx r_b + r_c. \qquad \ldots(b)$$

(3) Voltage gain

$$A_0 = \frac{r_{21} R_L}{(r_{22} + R_L)\, r_i}$$

$$= \frac{(r_b + r_m)\, R_L}{(r_b + r_c + R_L) \times \left| r_e + \dfrac{r_b (r_c - r_m + R_L)}{r_c + r_b + R_L} \right|}$$

$$= \frac{(r_b + r_m)\, R_L}{r_b r_e + r_c [r_e + r_b (1 - \alpha)] + (r_e + r_b) R_L}$$

$$\approx \frac{\alpha R_L}{r_e + r_b (1 - \alpha)} \qquad \ldots(c)$$

as first and last terms in denominator are small compared to the second term.

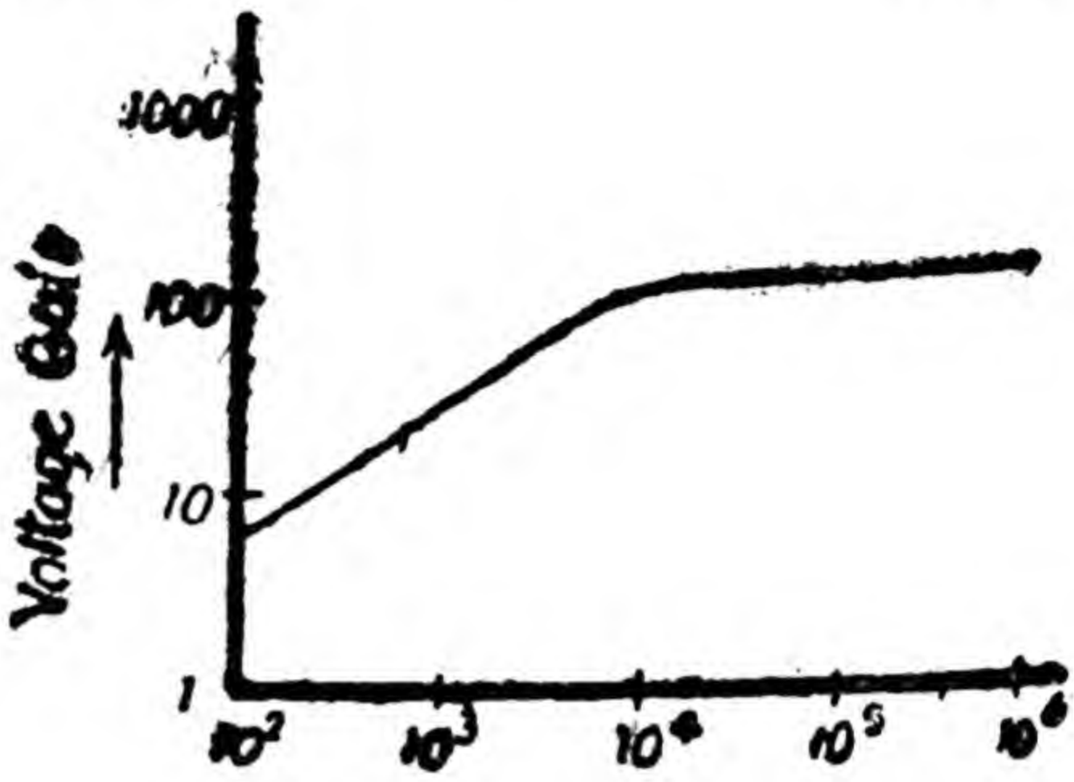


Fig. 10. Variation of voltage gain with load resistance.

The variation of voltage gain with load resistance is shown in fig. 10.

(4) Current gain $A_i = \dfrac{-r_{21}}{r_{22} + R_L} = \dfrac{-(r_b + r_m)}{r_e + r_b + R_L} \qquad \ldots(d)$

$$\approx \alpha.$$

(5) Power gain $= A_i \times A_v = \dfrac{\alpha^2 R_L}{r_e + r_b (1 - \alpha)}. \qquad \ldots(e)$

COMMON BASE TRANSISTOR AMPLIFIER HYBRID EQUATIONS :

Hybrid equivalent circuit of a common base transistor amplifier is shown in fig. 11 (*a*).

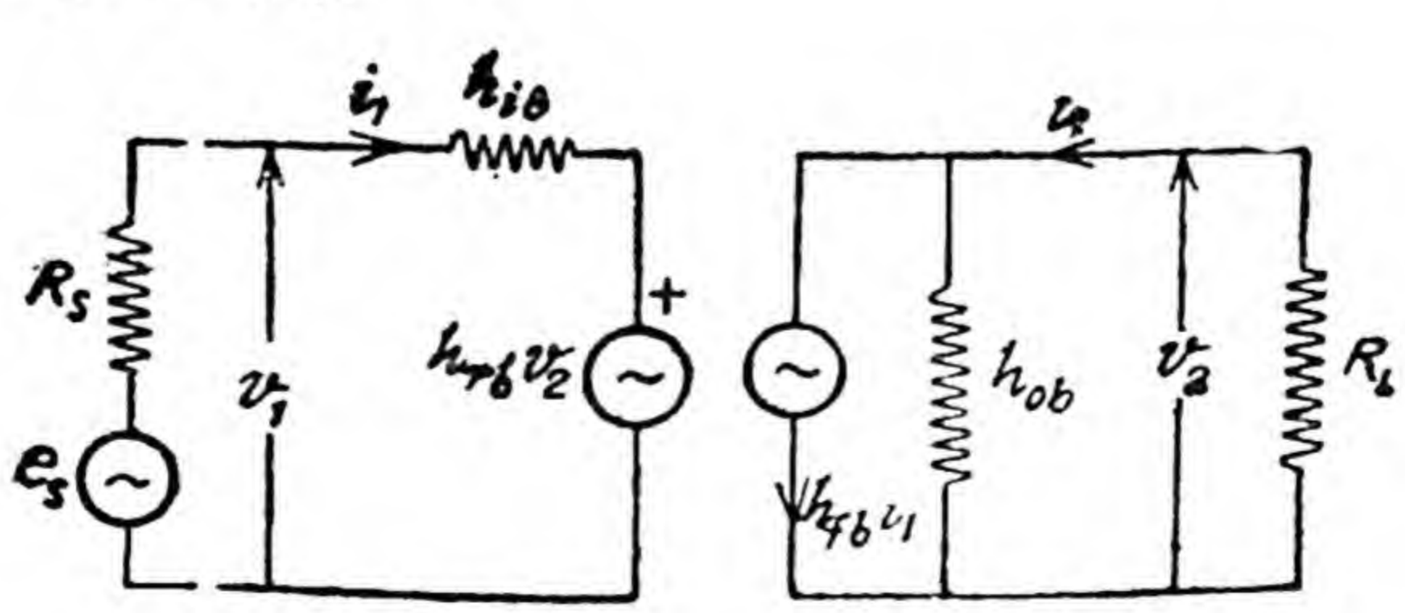


Fig. 11. (*a*).

The equations are

$$v_1 = h_{ib}\ i_1 + h_{rb}\ v_2 \qquad \ldots(8)$$

$$i_2 = h_{fb}\ i_1 + h_{ob}\ v_2. \qquad \ldots(9)$$

The second two equations needed are

$$v_1 = e_s - i_1 R_S \qquad \ldots(10)$$

$$v_2 = -i_2 R_L. \qquad \ldots(11)$$

(1) Current gain : Let us eliminate i_2 from equations (9) and (11).

$$0 = h_{fb} i_1 + \left(h_{ob} + \frac{1}{R_L} \right) v_2$$

or

$$v_2 = \frac{-h_{fb} i_1}{\left(h_{ob} + \frac{1}{R_L} \right)}$$

$$= -i_1 \frac{h_{fb} R_L}{1 + h_{ob} R_L}. \qquad \ldots(12)$$

Putting for v_2 from equation (11), we get

$$-i_2 R_L = -i_1 \frac{h_{fb} R_L}{1 + h_{ob} R_L}.$$

Therefore current gain is

$$A_{ib} = \frac{-i_2}{i_1} = \frac{-h_{fb}}{1 + h_{ob} R_L}. \qquad \ldots(13)$$

(2) Input Resistance : Putting equation (12) in equation (8), we get

$$v_1 = i_1 \left(h_{ib} - \frac{h_{rb} h_{fb} R_L}{1 + h_{ob} R_L} \right)$$

$$= i_1 \frac{h_{ib} + h_{ib} h_{ob} R_L - h_{rb} h_{fb} R_L}{1 + h_{ob} R_L} \qquad \ldots(14)$$

or

$$R_{ib} = \frac{v_1}{i_1} = \frac{R_L (h_{ib} h_{ob} - h_{fb} h_{rb}) + h_{ib}}{1 + h_{ob} R_L}$$

$$= \frac{R_L \Delta + h_{ib}}{1 + h_{ob} R_L}, \qquad \ldots(15)$$

where $\Delta = (h_{ib} h_{ob} - h_{rb} h_{rb})$.

Input resistance interms of current gain :

We know that

$$v_1 = h_{ib}\, i_1 + h_{re}\, v_2$$

or $v_1 = h_{ib}\, i_1 - h_{rb} i_2 R_L \qquad \because \quad v_2 = -i_2 R_L$

or $v_1 = h_{ib}\, i_1 + A_{ib}\, i_1 R_L\, h_{rb} \qquad \because \quad A_{ib} = -i_2/i_1$

$$\therefore \quad R_{ib} = \frac{v_1}{i_1} = h_{ib} + h_{rb}\, A_{ib}\, R_L \qquad \ldots(15a)$$

(3) Voltage Gain : From equations (12) and (14), we get

$$A_{vb} = \frac{v_2}{v_1} = \frac{-h_{fb} R_L}{h_{ib} + R_L \Delta}. \qquad \ldots(16)$$

Voltage gain interms of current gain :

$$A_{vb} = \frac{v_2}{v_1} = \frac{-i_2 R_L}{v_1} = \frac{A_{ib} i_1 R_L}{v_1}$$

or

$$A_{vb} = \frac{A_{ib} R_L}{R_{ib}} \qquad \ldots(16a)$$

(4) Power Gain : It is given by

$$A_{pb}=A_{ib}A_{vb}$$

$$=\frac{h^2_{fb}R_L}{(1+h_{ib}R_L)\,[h_{ib}+R_L\Delta]}. \qquad ...(17)$$

(5) Output Resistance : For the calculation of output resistance, we have to transfer the signal generator to the output terminals. At the same time, we place only R_S across the input terminals fig. 11 (*b*).

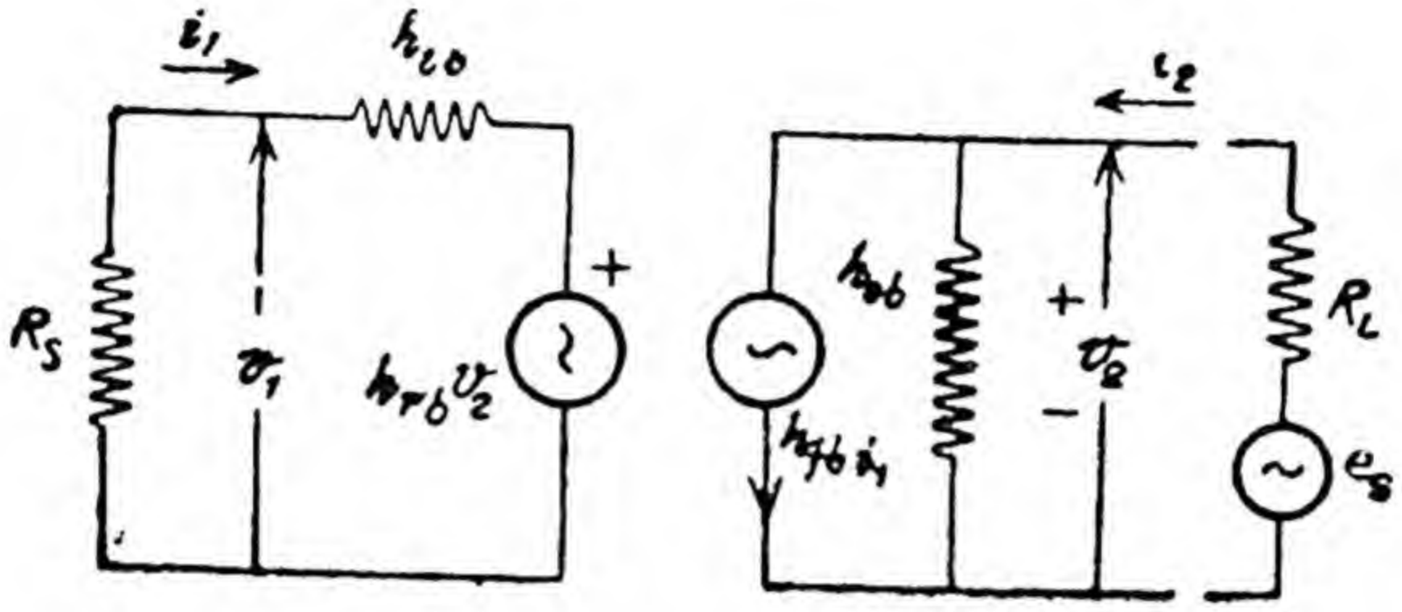


Fig. 11. (*b*)

The equations are

$$0=(R_S+h_{ib})\,i_1+h_{rb}v_2.$$

$$i_2=h_{fb}i_1+h_{ob}v_2.$$

Putting i_1 from first equation into second, we get

$$i_2=-\frac{h_{fb}h_{rb}}{(R_S+h_{ib})}\,v_2+h_{ob}v_2$$

or

$$R_{ob}=\frac{v_2}{i_2}=\frac{R_S+h_{ib}}{(R_S+h_{ib})\,h_{ob}-h_{fb}h_{rb}}$$

$$=\frac{R_S+h_{ib}}{R_S\,h_{ob}+h_{ib}h_{ob}-h_{fb}h_{rb}}$$

$$=\frac{R_S+h_{ib}}{R_S\,h_{ob}+\Delta} \qquad ...(18)$$

The output resistance can also be calculated in the following way :

The output admittance Y_{ob} is given by

$$Y_{ob}=\frac{i_2}{v_2} \text{ with } e_S=0$$

But $$i_2=h_{fb}\,i_1+h_{ob}\,v_2$$

$$\therefore \quad Y_{ob}=h_{fb}\,\frac{i_1}{v_2}+h_{ob}$$

But $R_S\,i_1+h_{ib}\,i_1+h_{rb}\,v_2=0$ $\qquad (\because \quad e_S=0)$

or $$\frac{i_1}{v_2}=-\frac{h_{rb}}{h_{ib}+R_S}$$

$$\therefore \quad Y_{ob}=h_{ob}-\frac{h_{fb}\,h_{rb}}{h_{ib}+R_S}$$

or $$R_{ob}=\frac{1}{Y_{ob}}=\left[\ h_{ob}-\frac{h_{fb}\,h_{rb}}{h_{ib}+R_S}\right]^{-1} \qquad ...(18a)$$

(6) Voltage gain considering source resistance R_S (overall voltage gain) : The overall voltage gain of the amplifier is given by

$$(A_{vb})_S=\frac{v_2}{e_S}=\frac{v_2}{v_1}.\frac{v_1}{e_S}=A_{vb}\,\frac{v_1}{e_S} \qquad ...(19)$$

corresponding to input impedance R_{ib}, the value of v_1 is given by

$$v_1=\frac{e_S\,R_{ib}}{R_S+R_{ib}}$$

Hence $$(A_{vb})_S=A_{ib}\,\frac{R_{ib}}{R_S+R_{ib}} \qquad ...(20)$$

If $R_S=0$ $\qquad (A_{vb})_S=A_{vb}$

(7) Current gain considering the source resistance R_S (overall current gain) : Let a current source is with internal resistance R_S be at the input side, then overall current gain is given by

$$(A_{ib})_S=\frac{i_2}{i_1}\cdot\frac{i_1}{i_S}=A_{ib}.\frac{i_1}{i_S}$$

Now $$i_1=\frac{i_S\,R_S}{R_{ib}+R_S}$$

$$\therefore\quad (A_{ib})_S=A_{ib}.\frac{R_S}{R_S+R_{ib}} \qquad ...(21)$$

If $R_S=\infty$, then $(A_{ib})_S=A_{ib}$.

Important formulae for common base transistor amplifier using h parameteres

Quantity	Expression
Current gain	$A_{ib}=-h_{fb}/(h_{ob}\,R_L+1)$
Input resistance	$R_{ib}=h_{ib}+h_{rb}\,A_{ib}\,R_L$
Voltage gain	$A_{vb}=A_{ib}\,R_L/R_{ib}$
power gain	$A_{pb}=A_{vb}.A_{ib}$
Output resistance	$R_{ob}=[h_{ob}-h_{fb}h_{rb}/(h_{ib}+R_S)]^{-1}$
Overall voltage gain	$(A_{vb})_S=A_{vb}.R_{ib}/(R_S+R_{ib})$
Overall current gain	$(A_{ib})_S=A_{ib}.R_S/(R_S+R_{ib})$

Example 1. *A common base amplifier is driven by a voltage es of internal resistance $R_S=1K\Omega$. The load impedance is a resistor $R_L=1K\Omega$. The transistor parametets are : $h_{ib}=22\Omega$, $h_{rb}=2{\cdot}9\times10^{-4}$, $h_{fb}=-0{\cdot}98$ and $h_{ob}=0{\cdot}50\ \mu A/V$. Compute current gain, voltage gain,*

overall current gain, overall voltage gain, input resistance, output resistance and operating power gain.

(i) Current gain $A_{ib}=-\dfrac{h_{fb}}{(h_{ob}R_L+1)}$

$$=-\frac{(-0\cdot98)}{(0\cdot50\times10^{-6}\times10^{3}+1)}\approx 0\cdot98$$

(ii) Input resistance $R_{ib}=h_{ib}+h_{rb}A_{ib}R_L$

$$=22+(2\cdot9\times10^{-4})\ (0\cdot98)\ (10^{3})$$

$$=22\cdot28\ \Omega$$

(iii) Voltage gain $A_{vb}=A_{ib}\ R_L/R_{ib}$

$$=\frac{0\cdot98\times10^{3}}{22\cdot28}=43\cdot99$$

(iv) Overall voltage gain $(A_{vb})s=A_{vb}.\dfrac{R_{ib}}{(R_S+R_{ib})}$

$$=(43\cdot99)\ \frac{22\cdot28}{10^{3}+22\cdot28}=0\cdot96$$

(v) Overall current gain $(A_{ib})s=A_{ib}.\dfrac{R_S}{R_S+R_{ib}}$

$$=(0\cdot98)\ \frac{10^{3}}{10^{3}+22\cdot88}=0\cdot96$$

(vi) Output resistance $R_{ob}=[h_{ob}-h_{fb}.h_{rb}/(h_{ib}+R_S)]^{-1}$

$$=\left[\ 0\cdot50\times10^{-6}-\frac{(-0\cdot98)(2\cdot9\times10^{-4})}{22+1000}\right]^{-1}$$

$$=1\cdot28\ \text{M}\Omega.$$

(vii) Power gain $=A_{vb}.A_{ib}=(43\cdot99)\ (0\cdot98)=43\cdot11.$

16·4. GROUNDED COLLECTOR CIRCUIT :

A grounded collector transistor amplifier circuit is similar to cathode follower in vacuum tube amplifier circuit. Figure 12 (*a*) shows a grounded collector *PNP* transistor circuit. The collector is at a.c. ground, while d c. ground is prevented by means of a capacitor (not shown in fig.) or in other words, collector is common to base and emitter for a.c signals. Emitter is biased in forward direction by the voltage V_{EE} and collector is biased in reverse direction by the voltage V_{CC}.

The input signal will vary base current about its average value which produces same variations of greater amplitude in emitter current. Therefore, output voltage will be larger in variation than input signal.

When input signal is positive then bias voltage V_C will increase because of the polarity of input signal and that of V_{CC} being the same. Due to increase in V_C, the emitter and base currents will decrease. Consequently, the output voltage across R_L will decrease or in other words the potential at the emitter will become relatively more positive with respect to the collector. In a similar man-

ner, the negative input signal will make emitter less positive with respect to collector. Thus we see that there is no phase reversal between the input and output signals.

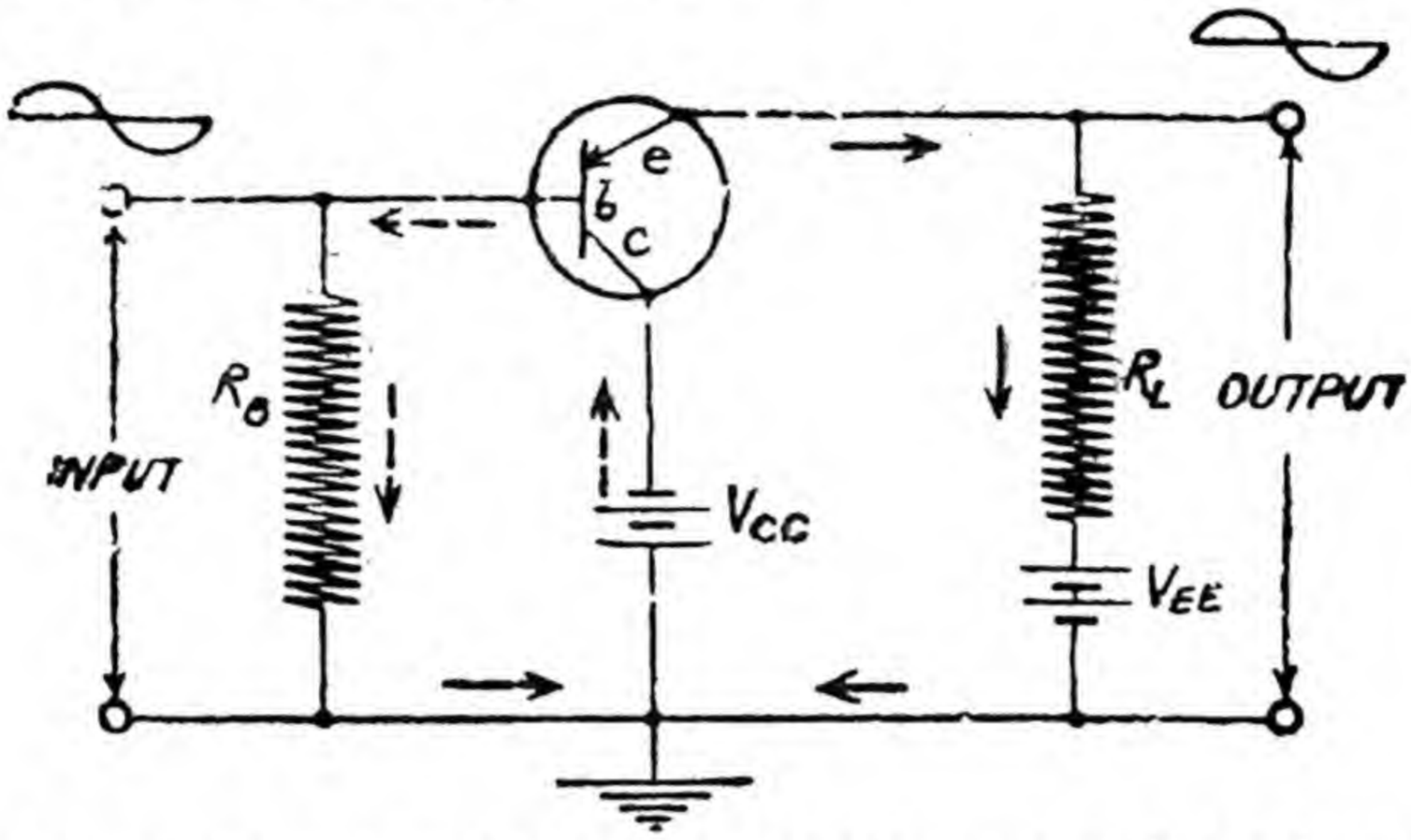


Fig. 12 (*a*) *PNP* ground collector transistor amplifier circuit.

In this type of amplifier, the value of load impedance applied in the circuit is determined by its input impedance. Because of the low value of the required output impedance, the voltage gain is always less than unity.

Now we shall produced to calculate *input resistance, output resistance, voltage gain and current gain of common collector amplifier.*

Mathematical analysis of common collector amplifier using r parameters :

Fig. 12 (*b*) and 12 (*c*) show the one voltage generator form of equivalent circuit of a common collector (*i*) treated as a four pole and (*ii*) derived on the basis of the processes involved within the transistor respectively. The equivalent circuit is obtained simply by changing the position of emitter arm and collector arm in a common emitter amplifier.

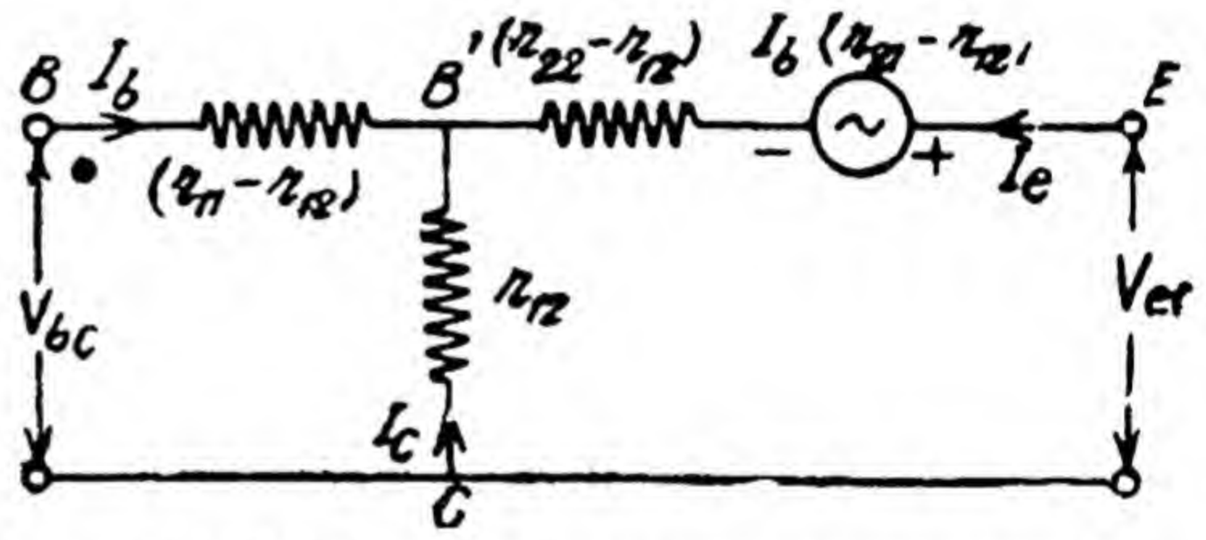


Fig. 12 (*b*) One voltage generator equivalent circuit of common transistor treated as a four pole.

Network equations for figure 12 (*b*) are

$$V_{bc}=I_b r_{11}+I_e r_{12} \quad \ldots(1)$$

$$V_{ec}=I_e(r_{22}-r_{12}) + I_b(r_{21}-r_{12})+(I_b+I_e)\, r_{12}$$

or $$V_{ec}=I_b r_{21}+I_e r_{22} \quad \ldots(2)$$

Network equations for figure 12 (*c*) are

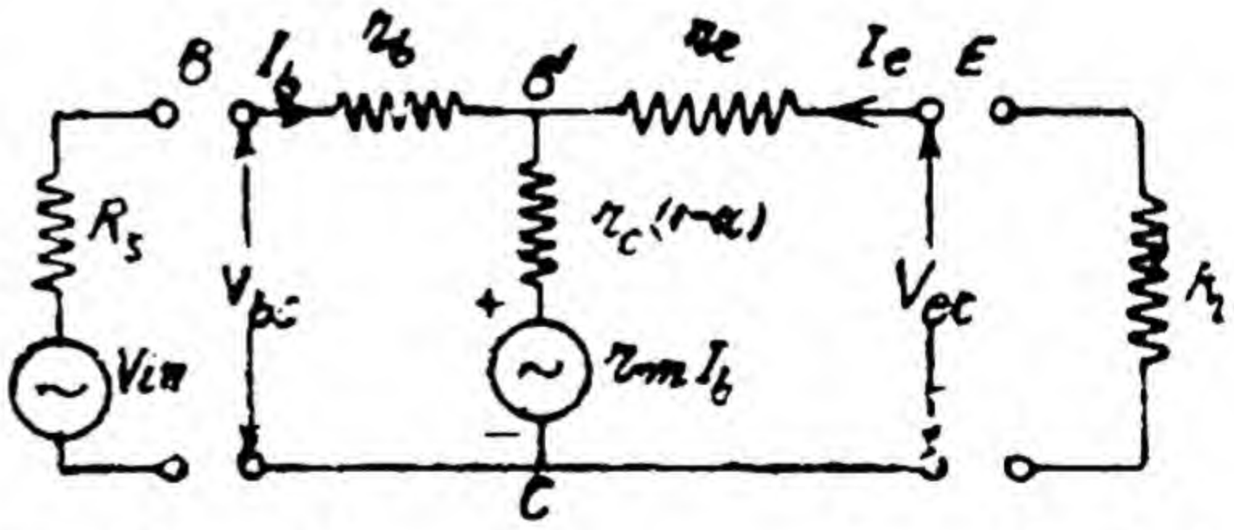


Fig. 12 (*c*). One voltage generator equivalent circuit of common collector amplifier derived on the basis of process s involved in the transistor.

$$V_{bc}=I_b r_b+(I_b+I_c)\,[r_c\,(1-\alpha)]+r_m I_b$$

or $$V_{bc}=I_b\,(r_b+r_c)+I_e r_c\,(1-\alpha) \qquad \ldots(3)$$

$$V_{ec}=I_c r_e+(I_b+I_e)\,[r_c\,(1-\alpha)]+r_m I_b$$

or $$V_{cc}=I_b r_c+I_e\,[r_e+r_c\,(1-\alpha)]. \qquad \ldots(4)$$

Comparing equations (1) and (2) with equations (3) and (4) respectively, we have

$$r_{11}=r_b+r_c,\ r_{12}=r_c\,(1-\alpha),\ r_{21}=r_c$$

and $$r_{22}=r_e+r_c\,(1-\alpha).$$

Thus the network equations reduce to

$$\begin{bmatrix} V_{bc} \\ V_{cc} \end{bmatrix}=\begin{bmatrix} r_b+r_c & r_c\,(1-\alpha) \\ r_c & r_e+r_c\,(1-\alpha) \end{bmatrix}\begin{bmatrix} I_b \\ I_e \end{bmatrix}. \qquad \ldots(5)$$

(1) The input resistance : $r_i=r_{11}-\dfrac{r_{12}r_{21}}{r_{22}+R_L}$

$$=r_b+r_c-\frac{r_c\,(1-\alpha)\,r_c}{r_e+r_c\,(1-\alpha)+R_L}$$

$$=r_b+\frac{r_c\,(r_e+R_L)}{r_e+r_c\,(1-\alpha)+R_L}$$

If $r_c\,(1-\alpha) \gg r_e$ and R_L, then

$$r_i=r_b+\frac{r_e\,(r_c+R_L)}{r_c\,(1-\alpha)}=r_b+\frac{r_e+R_L}{(1-\alpha)}$$

If $r_e \gg R_L$, then $r_i=r_b+\dfrac{r_e}{(1-\alpha)}$, .. **(a)**

and if $r_e \ll R_L$, then $r_i=r_b+\dfrac{R_L}{(1-\alpha)}$.

The variation of input resistance with load is shown in fig. 12 (*d*).

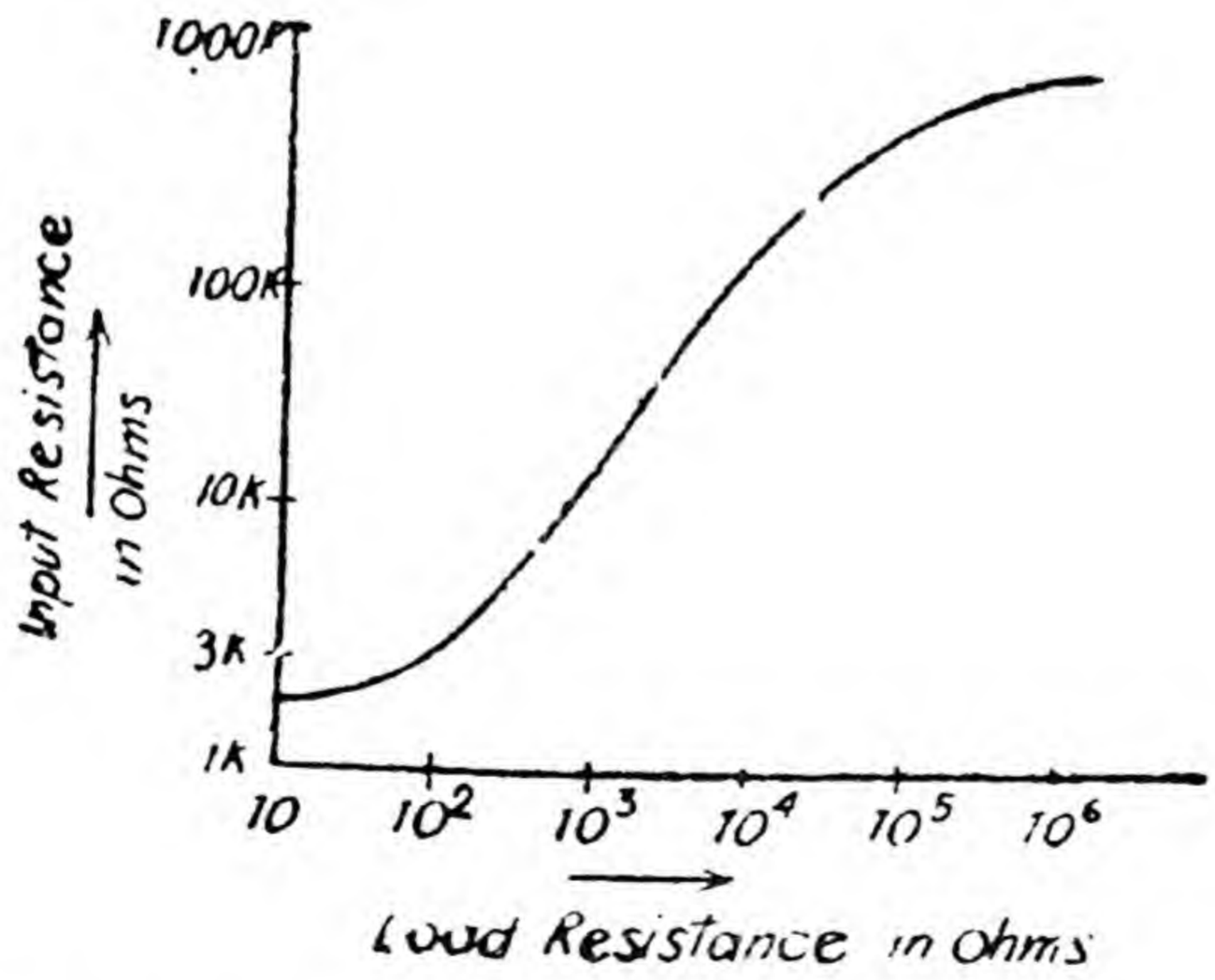


Fig. 12 (*d*). Variation of input resistance with load.

(2) The output resistance $r_0=r_{22}-\dfrac{r_{12}r_{21}}{r_{11}+R_S}$.

$$=r_e+r_c(1-\alpha)-\frac{r_c(1-\alpha)\,r_c}{r_b+r_c+R_S}$$

$$=r_e+\frac{r_c(1-\alpha)(r_b+R_S)}{r_b+r_c+R_S}$$

If $r_c \gg r_b+R_S$, then

$$r_0 \approx r_e+(1-\alpha)(r_b+R_S) \qquad \ldots\text{(b)}$$

The output resistance is very small.

(3) Voltage gain $A_v=\dfrac{r_{21}R_L}{(r_{22}+R_L)r_i}$

$$=\frac{r_cR_L}{[r_e+r_c(1-\alpha)+R_L]}\times\frac{1}{\left[r_b+\dfrac{(r_e+R_L)\,r_c}{r_e+r_c(1-\alpha)+R_L}\right]}$$

$$=\frac{r_c\,R_L}{r_b[r_e+r_c(1-\alpha)+R_L]+(r_e+R_L)\,r_c}$$

$$\approx\frac{1}{1+\dfrac{r_b(1-\alpha)}{R_L}} \qquad \ldots\text{(c)}$$

when $r_c \gg r_br_c$ and R_L.

Equation (c) shows that the voltage gain reduces to unity for extremely large load.

(4) Current gain $A_i=-\dfrac{r_{21}}{r_{22}+R_L}=\dfrac{-r_c}{r_e+r_c(1-\alpha)+R_L}$

$$=\frac{-1}{1+\dfrac{r_c-r_m+R_L}{r_c}}$$

$$\approx\frac{-1}{(1-\alpha)}. \qquad \ldots\text{(d)}$$

when $r_m \gg r_e$ and R_L.

The negative sign indicates that there is no phase reversal in common collector amplifier.

(5) Power gain $=A_i\times A_v$

$$\approx\frac{1}{(1-\alpha)}\times\frac{1}{1+\dfrac{r_b(1-\alpha)}{R_L}} \qquad \ldots\text{(e)}$$

COMMON COLLECTOR TRANSISTOR AMPLIFIER HYBRID EQUATIONS

Hybrid equivalent circuit of a common collector transistor amplifier is shown in fig. 12 (*e*).

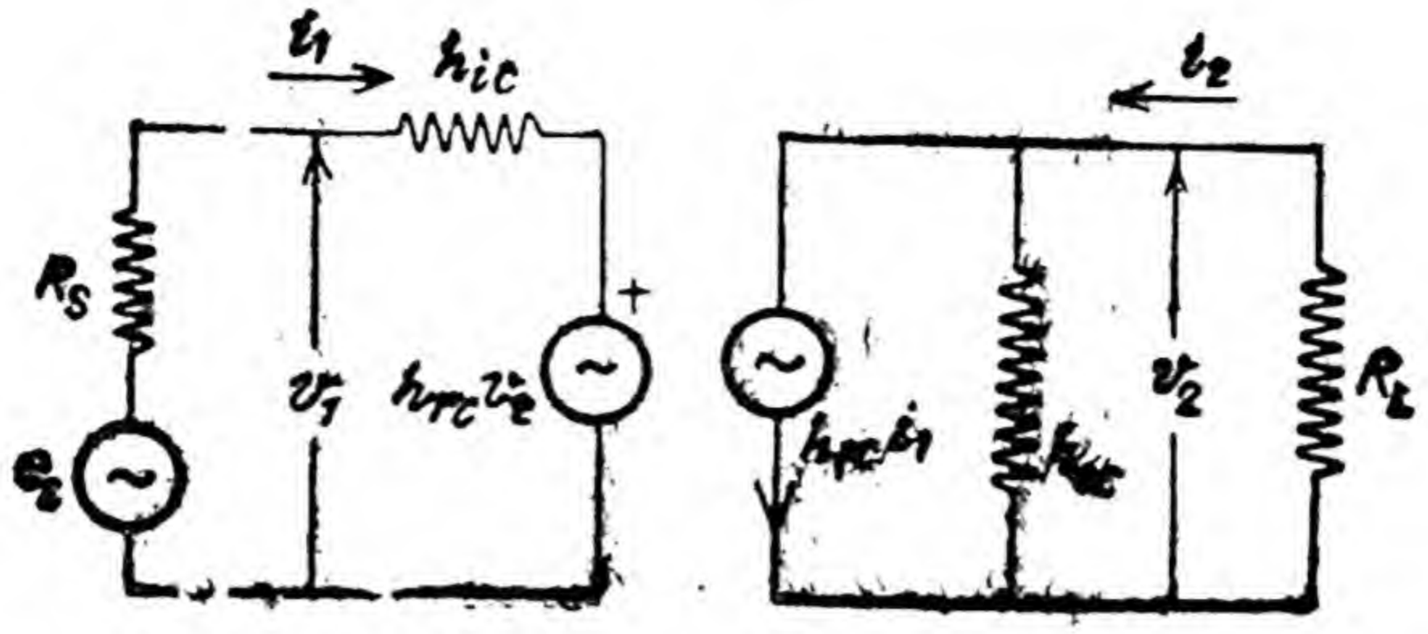


Fig. 12 (*e*)

The equations are

$$v_1 = h_{ic} i_1 + h_{rc} v_2 \quad \ldots(6)$$

$$i_2 = h_{fc} i_1 + h_{oc} v_2 \quad \ldots(7)$$

The second two equations needed are

$$v_1 = e_S - i_1 R_S \quad \ldots(8)$$

$$v_2 = -i_2 R_L \quad \ldots(9)$$

(1) Current Gain : Let us eliminate i_1 from equations (7) and (9). This gives

$$0 = h_{fc} i_1 + \left(h_{oc} + \frac{1}{R_L} \right) v_2$$

or

$$v_2 = \frac{-h_{fc} i_1}{\left(h_{oc} + \frac{1}{R_L} \right)}$$

$$= -i_1 \frac{h_{fc} R_L}{1 + h_{oc} R_L} \quad \ldots(10)$$

Putting for v_2 from eq. (9), we get

$$-i_2 R_L = -i_1 \frac{h_{fc} R_L}{1 + h_{oc} R_L}$$

Therefore current gain is

$$A_{ic} = \frac{-i_2}{i_1} = \frac{-h_{fc}}{1 + h_{oc} R_L} \quad \ldots(11)$$

(2) Input Resistance : Putting eq. (10) in equation (6), we get

$$v_1 = i_1 \left(h_{ic} - \frac{h_{fc} h_{rc} R_L}{1 + h_{oc} R_L} \right)$$

$$= i_1 \frac{h_{ic} + h_{ic} h_{oc} R_L - h_{fc} h_{rc} R_L}{1 + h_{oc} R_L} \quad \ldots(12)$$

so that

$$R_{ic} = \frac{v_1}{i_1} = \frac{R_L \Delta + h_{ic}}{1 + h_{oc} R_L}, \quad \ldots(13)$$

where

$$\Delta = (h_{ic} h_{oc} - h_{fc} h_{rc})$$

Input resistance interms of current gain.

We know that

$$v_1 = h_{ic} i_1 + h_{rc} v_2$$

or $v_1 = h_{ic} i_1 - h_{rc} i_2 R_L$ $\because v_2 = -i_2 R_L$

or $v_1 = h_{ic} i_1 + A_{ic} i_1 R_L h_{rc}$ $\because A_{ic} = -i_2/i_1$

$$\therefore \quad R_{ic} = \frac{v_1}{i_1} = h_{ic} + h_{rc} A_{ic} R_L \qquad (13a)$$

(3) Voltage Gain : From equations (10) and (12), we get

$$A_{vc} = \frac{v_2}{v_1} = \frac{-h_{fc} R_L}{h_{ic} + R_L \Delta} \qquad ...(14)$$

Voltage gain interms of current gain

$$A_{vc} = \frac{v_2}{v_1} = \frac{-i_2 R_L}{v_1} = \frac{A_{ic} i_1 R_L}{v_1}$$

or $$A_{vc} = \frac{A_{ic} R_L}{R_{ic}} \qquad ...(14a)$$

(4) Power Gain : It is given by

$$A_{pc} = A_{ic} A_{vc}$$

$$= \frac{h^2_{fc} R_L}{(1 + h_{oc} R_L)(h_{ic} + R_L \Delta)} \qquad ...(15)$$

(5) Output Resistance : For the calculation of output resistance, we have to transfer the signal generator to the output terminals. At the same time, we place only R_S across the input terminals fig. 12 (f). The equations are

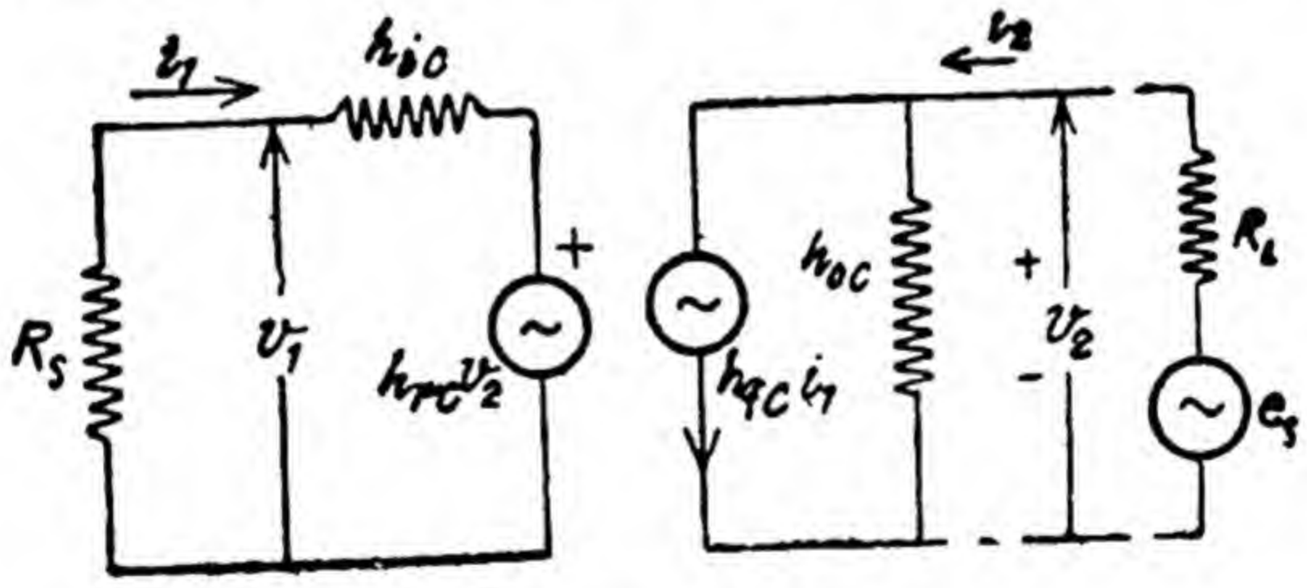


Fig. 12 (f)

$$0 = (R_S + h_{ic})\, i_1 + h_{rc} v_2$$

$$i_2 = h_{fc} i_1 + h_{oc} v_2$$

Putting i_1 from first equation into second, we get

$$i_2 = -\frac{h_{fc} h_{rc} v_2}{(R_S + h_{ic})} + h_{rc} v_2$$

or $$R_{oc} = \frac{v_2}{i_2} = \frac{R_S + h_{ic}}{(R_S + h_{ic})\, h_{rc} - h_{fc} h_{rc}}$$

$$= \frac{R_S + h_{ic}}{R_S h_{oc} + \Delta} \qquad ...(16)$$

The output resistance can also be calculated in the following way :

The output admittance Y_{oc} is given by

$$Y_{oc}=\frac{i_2}{v_2} \text{ with } e_S=0$$

But $$i_2=h_{fc}i_1+h_{oc}v_2$$

$$\therefore \quad Y_{oc}=h_{fc}\frac{i_1}{v_2}+h_{oc}$$

But $$R_Si_1+h_{ib}i_1+h_{rb}v_2=0 \qquad (\because\ e_S=0)$$

or $$\frac{i_1}{v_2}=\frac{-h_{rc}}{h_{ic}+R_S}$$

$$\therefore \quad Y_{oc}=h_{oc}-\frac{h_{fc}\ h_{rc}}{h_{ic}+R_S}$$

or $$R_{oc}=\frac{1}{Y_{oc}}=\left[\ h_{oc}-\frac{h_{fc}h_{ro}}{h_{ic}+R_S}\ \right]^{-1} \qquad ...(16a)$$

(6) **Voltage gain considering source resistance (overall voltage gain) :** The overall voltage gain of the amplifier is given by

$$(A_{vo})_S=\frac{v_2}{e_S}=\frac{v_2}{v_1}.\frac{v_1}{e_S}=A_{vc}\frac{v_1}{e_S}$$

Corresponding to input impedance R_{ic}, the value of v_1 is given by

$$v_1=\frac{e_S\ R_{ic}}{R_S+R_{ic}}$$

Hence $$(A_{vc})_S=A_{vo}\frac{R_{ic}}{R_S+R_{ic}} \qquad ...(17)$$

If $R_S=0$ $\qquad (A_{vo})_S=A_{vo}$

(7) **Current gain corresponding the source resistance R (overall current gain).** Let a current source i_S with internal resistance R_S be at the input side, then overall current gain is given by

$$(A_{ic})_S=\frac{i_2}{i_1}.\frac{i_1}{i_S}=A_{ic}\ \frac{i_1}{i_S}$$

Now $$i_1=\frac{i_SR_S}{R_{ic}+R_S}$$

$$\therefore \quad (A_{iv})_S=A_{ic}\frac{R_S}{R_S+R_{ic}} \qquad ...(18)$$

If $R_S=\infty$, $(A_{ic})_S=A_{ic}$

Important formulae for common collector transistor amplifier using h parameters

Quantity	Expression
Current gain	$A_{ic}=-h_{fc}/(h_{oc}R_L+1)$
Input Resistance	$R_{ic}=h_{io}+h_{rc}A_{ic}R_L$
Voltage gain	$A_{vc}=A_{ic}R_L/R_{ic}$
Power gain	$A_{pc}=A_{vc}.\ A_{ic}$
Output Resistance	$R_{oc}=[h_{oc}-h_{fc}h_{rc}/(h_{ic}+R_S)]^{-1}$
Overall voltage gain	$(A_{vc})_S=A_{vc}.R_{ic}/(R_S+R_{ic})$
Overall current gain	$(A_{ic})_S=A_{ic}\ R_S/(R_S+R_{ic})$

Example 1. *A common collector amplifier is driven by a voltage e_S of internal resistance $R_S=1\ K\Omega$. The load impedance is a resistor $R_L=1\ K\Omega$. The transistor parameters are : $h_{ic}=1{\cdot}1K\Omega$, $h_{ro}=1$, $h_{fc}=-51$ and $h_{oc}=25\,\mu A/V$. Compute current gain, voltage gain, overall current gain, overall voltage gain, input resistance, output resistance and operating power gain.*

(i) Current gain $A_{ic}=-\dfrac{h_{fc}}{(h_{oc}R_L+1)}=-\dfrac{(-51)}{(25\times10^{-6})(10^3)+1}$

$=\mathbf{49{\cdot}76.}$

(ii) Input resistance $R_{ic}=h_{ic}+h_{ro}A_{ic}R_L$

$\therefore\quad R_{ic}=1{\cdot}1\times10^3+1.\ (49{\cdot}76)(1\times10^3)$

$=\mathbf{50{\cdot}86\times10^3\,\Omega}$

(iii) Voltage gain $A_{vc}=\dfrac{A_{ic}R_L}{R_{ic}}=\dfrac{49{\cdot}76\ (1\times10^3)}{50{\cdot}86\times10^3}$

$=\mathbf{0{\cdot}98.}$

(iv) Overall voltage gain $(A_{vc})_S=\dfrac{A_{vc}\ R_{ic}}{R_S+R_{ic}}$

$=\dfrac{0{\cdot}98\times(50{\cdot}86\times10^3)}{10^3+(50{\cdot}86\times10^3)}$

$=\mathbf{0{\cdot}96.}$

(v) Overall current gain $(A_{ic})_S=A_{ic}\dfrac{R_S}{R_S+R_{ic}}$

$\therefore\quad (A_{ic})_S=(49{\cdot}76)\dfrac{1\times10^3}{(1\times10^3)+(50{\cdot}86\times10^3)}$

$=\mathbf{0{\cdot}96.}$

(vi) Output resistance $R_{oc}=[h_{oc}-h_{fc}h_{rc}/(h_{ic}+R_S)]^{-1}$

$=\left[25\times10^{-6}-\dfrac{(-51)\ (1)}{1{\cdot}1\times10^3+10^3}\right]^{-1}$

$=\mathbf{41{\cdot}17\times10^3\,\Omega}$

(vii) Power goin $A_{pc}=A_{vc}.\ A_{ic}$

$=(0{\cdot}98)(49{\cdot}76)=\mathbf{48{\cdot}76.}$

16·5. COMPARATIVE STUDY OF THREE TYPES OF AMPLIFIER CIRCUITS:

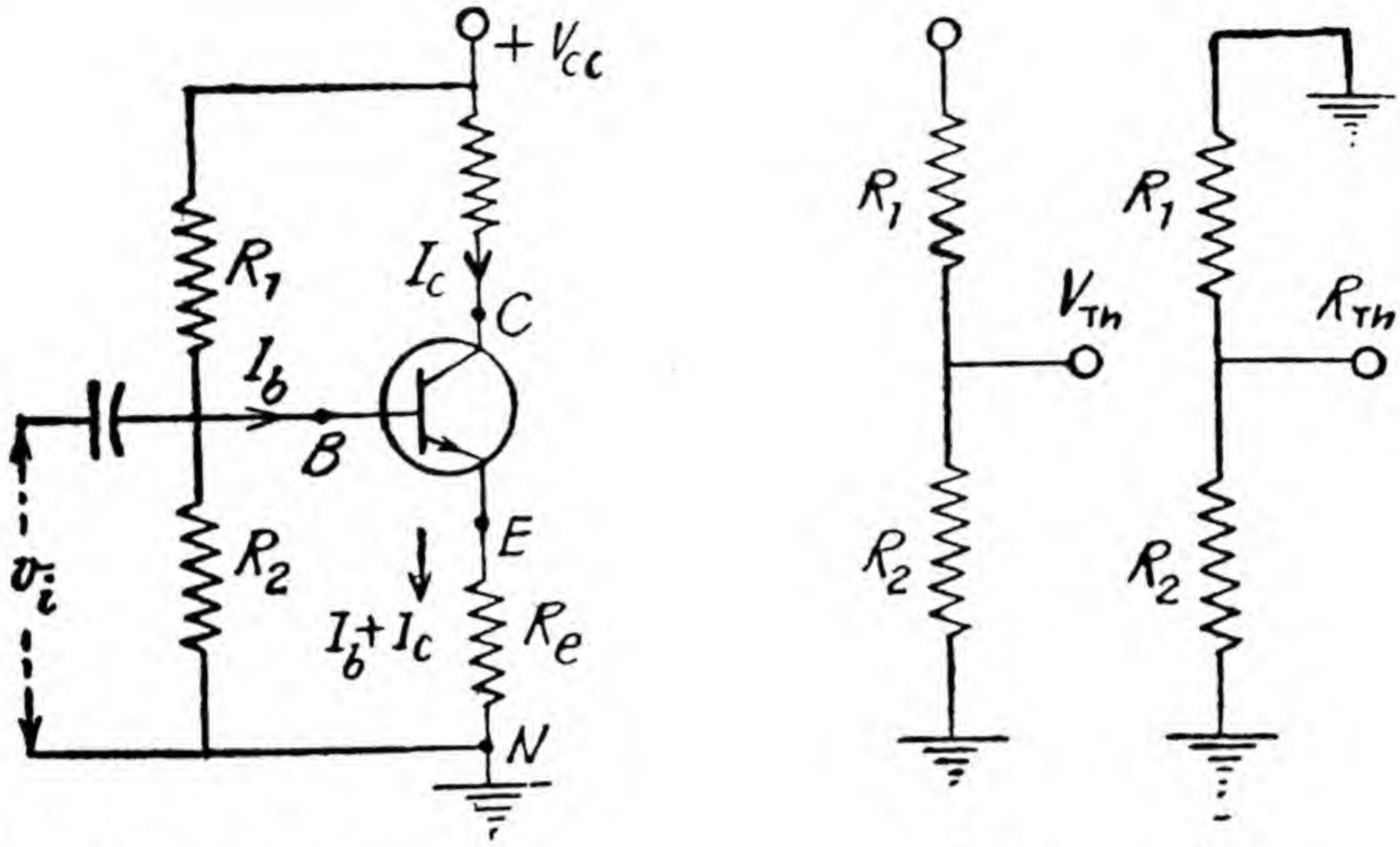
Property	Common Emitter	Common base	Common collector
Transistor resistance	Output resistance of the transistor is of the order of $50K\Omega$ and input resistance is 1 to $2K$ ohms.	Output resistance of the transistor is very high, of the order of 1 to 2 megohm. Input resistance of transistor is very low of the order of 20 to 50 ohms.	Output resistance of the transistor is very low, of the order of 100 to 1000 ohms. The input resistance is very high of the order of 150 K ohms to 600 ohms.
Current gain	Large current gain 20 to 200	Approximately no current gain, 0 85 to 0·995	Current gain is high 20 to 200
Voltage gain	Large voltage gain	Voltage gain is fair because of resistance gain between base and collector	Voltage gain is always less than unity
Power gain	Highest Power gain	Power gain is less	Power gain is less than other types of circuits
Phase of input and output signals	Input and output signals are 180° apart	No phase reversal	No phase reversal

From the comparative study we conclude that common emitter amplifier is alone capable of providing both current gain and voltage gain, hence most popularly used.

16·6. SELF-BIAS OR EMITTER BIAS

Of all the biased circuits, self bias or voltage divider bias is most important. Fig. (13 *a*) shows the circuit of a self bias. In the figure a resistance R_e is connected in the emitter lead and a combination of R_1 and R_2 form a voltage divider in the base circuit. When a current flows through R_e, there is a voltage drop in a direction which reverse-bias the emitter junction. But the emitter junction must be forward biased for a transistor to operate in active region of characteristics. For this purpose a base voltage

is obtained from the common supply V_{CC} through R_1R_2 combination.



(a) Self biased circuit (b) (c)

(d) Equivalent circuit

Fig. (13) Self biased circuit and its equivalent circuit.

Now the net forward bias across the emitter junction is equal to base voltage [V_b fig. 13 (d)] minus the d.c. voltage drop across R_e. In order to Thevenizing the base circuit, we first break the base lead as shown in fig. (13 b). The open circuit voltage is

$$V_{Th}=\frac{R_2\,V_{CC}}{R_1+R_2}\ (V_{Th}\text{ is generally denoted by } V_b)$$

Again we reduce V_{CC} to zero as shown in fig. (13 c). Looking into the voltage divider, the Thevenin resistance is

R_{Th} = a parallel combination of R_1 and R_2

or
$$R_{Th}=\frac{R_1R_2}{R_1+R_2}$$

This is generally denoted by R_b.

The equivalent circuit is shown in fig. (13 d).

The reason that how the operating point stability is improved can be understood as follows :

Let us consider that the collector current I_c tend to increase due to the increase in I_{Co} (leakage current) because of temperature rise. Now this increased current I_c causes an increased d.c. voltage drop across R_e. This reduces the net emitter to base forward bias. As a result the base current I_b is reduced. The reduction in base current causes a reduction in collector current. Thus the self biasing resistor R_e reduces the increase in I_c and improves the operating point stability.

16·7. BIAS STABILISATION (Operating point stability)

The application of proper bias in transistor amplifying circuits is essential since it specifies the operating point and hence the desired mode of operation.

The properties of a transistor, the latter being of semiconducing materials, are temperature dependent and thus transistor parameters and characteristics are markedly temperature-sensitive. Unless bias arrangement is correctly designed these variations will shift the operating point and hence the desired mode of operation. It is, therefore, necessary to stabilise the bias in various transistor circuits.

Due to the rise in temperature, thermal breakages of covalent bonds is resulted which gives rise to a leakage current. This leakage current forms the one component of collector cur ent of a junction transistor, the other component of which is dependent on base current and contributes all to the ampl fying properties. Leakage current is denoted by I_{Co}, the zero emitter collector current, and it varies exponentially with temperature.

There is a further increase of this leakage current on account of the heating of collector junction. The transistor will be damaged if this current reaches an abnormal value far greater than safe collector dissipation. The process of developing heat by a transistor to an extent that it is destroyed is called thermal runaway. It is therefore, essential to limit the collector current to a safe value by providing protective circuits.

16·7-1. BIAS INSTABILITY IN COM MON BASE AMPLIFIER :

A common base amplifier c rcuit is shown in figure 14. When emitter circuit is open, collector current consists only the component fo current due to reverse current of the diohe formed by the oscillator base junction. It si called leakage current, represented by I_{Co}. Tdis total collector current I_c is I_{Co}. But when emitter circuit is completed, usefull current αI_e is also in-

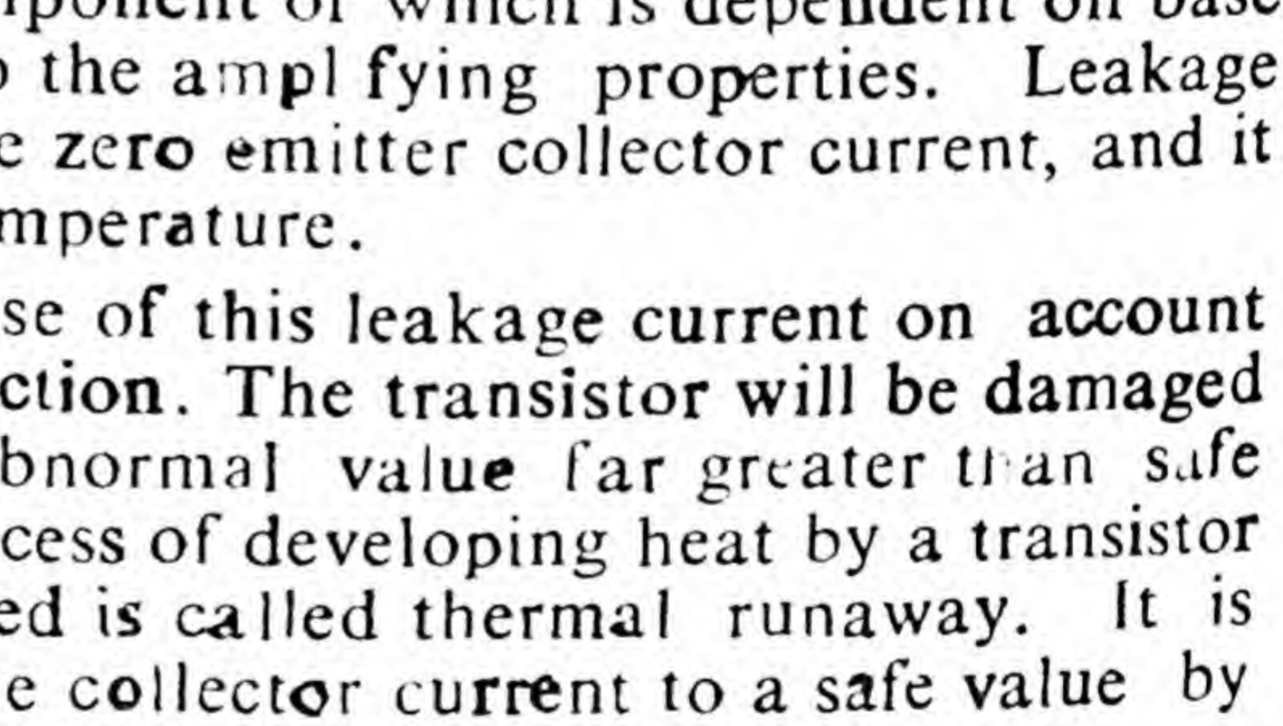


Fig. 14. Common base configuration.

cluded in I_o, *i.e.*,

$$I_c = \alpha I_e + I_{Co}. = I_C \text{ (ideal)} + I_{Co}$$

Relation shows that a change in emitter current, caused by an input signal, will produce an α change in collector current but I_{Co} stands alone. Though I_{Co} varies exponentially with temperature and even if it doubles with each 10°C rise in temperature for germanium and 6°C rise in temperature for Silicon, it is still negligible as compared to αI_e. I_{Co} for such configuration is always of the order of μA while I_c may be of the order of mA. For example, at 25°C, I_{Co} is μA which is negligibie as compared to I_o order.

Since I_{Co} is small, bias instability is least in common base configuration.

It is worth noting that at room temperature I_{Co} in silicon transistors are much smaller than in germanium transistors, and so are widely used.

16·7-2. BIAS INSTABILITY IN COMMON EMITTER AMPLIFIER :

The case of common emitter circuit is most important to be considered since it is widely adopted in practice particularly because of its high power gain.

We know that emitter current is equal to the sum of the base and collector currents *i.e.*,

$$I_e = I_b + I_c;$$

The collector current is

$$I_c = \alpha I_e + I_{Co}$$

$$I_c = \alpha (I_b + I_c) + I_{Co}$$

$$I_c (1 - \alpha) = \alpha I_b + I_{Co}$$

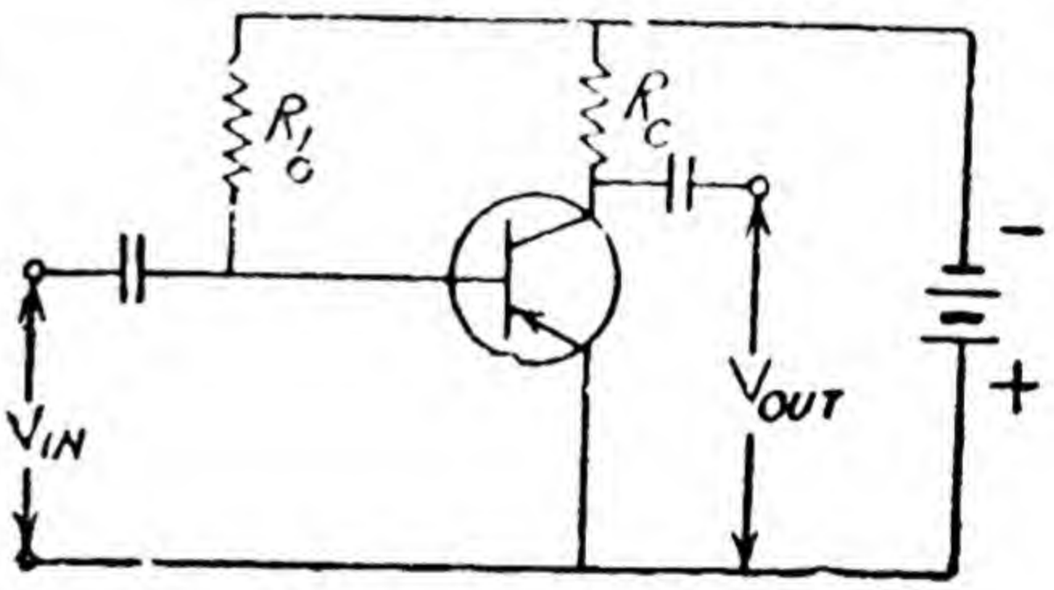


Fig. 15. Common emitter amplifier with fiixed bias.

$$\therefore \quad I_c = \frac{\alpha}{(1-\alpha)} I_b + \frac{1}{(1-\alpha)} I_{Co}$$

$$= \frac{\alpha}{(1-\alpha)} I_b + \left[1 + \frac{\alpha}{(1-\alpha)}\right] I_{Co}$$

$$= \beta I_b + (1+\beta) I_{Co}$$

$$= \beta I_b + I'_{Co},$$

where β is current gain and I'_{Co} is the total leakage current in collector circuit.

The bias instability depends upon I'_{Co} which in such type of amplifier, can assume a value of the order of 300μA at 25°C rising to 3 mA. at 60 °C. Since it is comparatively higher than signal current, bias stabilisation is essential in common emitter configuration. For this purpose, protective circuit should be provided.

16·8. STABILITY FACTOR :

A stability factor, S, will be developed which relates the varia-

tion in output current to the variation of I_{C_0}. This is defined as the rate of change of collector current I_o with respect to leakage current I_{C_o} keeping β and V_{BE} constant *i.e.*,

$$S=\frac{\partial I_o}{\partial I_{C_o}}\simeq\frac{dI_c}{dI_{C_o}}\approx\frac{\Delta I_o}{\Delta I_{C_o}};\ \beta,\ I_b \text{ constant.}$$

Smaller the value of S, less is the thermal instability. Hence S should be kept as small as possible.

For stabilised circuit, $dI_o < dI_{C_o}$ and $S < 1$; while $dI_o > dI_{C_o}$ and $S > 1$ for an unstabilised circuit.

Now we shall proceed to determine the thermal stability of a particular amplifying circuit.

16·9. STABILISATION BY COLLECTOR BASE RESISTANCE (Self Operation) :

The collector current is given by

$$I_c=\beta I_b+(1+\beta)\ I_{C_o}$$

If temperature rises leakage current I'_{C_o} will increase. An increase in I'_{C_o} will develop more voltage across R_c and hence the collector voltage will become more positive. The result of this increase in collector voltage is to reduce the base current I_b and hence the useful component βI_b. This decrease in total collector current I_c will decrease the voltage across R_o and the previous conditions are again restored. This explains stabilising action.

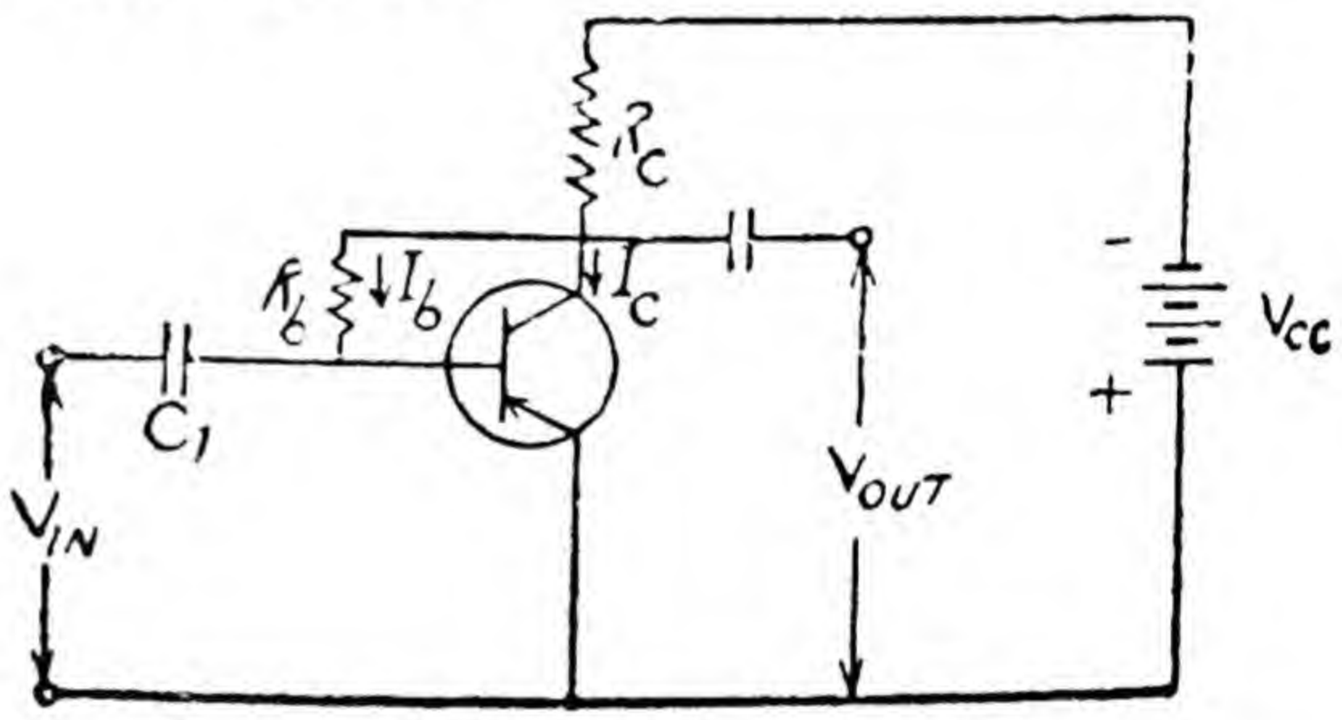


Fig. 16. Self bais operation.

Mathematically conditions can be derived as follows :

Collector supply voltage

$$V_{cc}=I_bR_b+(I_b+I_c)\ R_c.$$

But $$I_c=\beta I_b+(1+\beta)\ I_{C_o}$$

or $$I_b=\frac{I_c-(1+\beta)\ I_{C_o}}{\beta},$$

$$\therefore\quad V_{cc}=I_b\ (R_b+R_c)+I_cR_o$$

$$=\frac{[I_c-(1+\beta)\ I_{C_o}]}{\beta}\ (R_b+R_c)+I_cR_o$$

$$=I_c\left[\frac{R_b+R_c}{\beta}+R_c\right]-(1+\beta)\ I_{C_o}\ \frac{R_b+R_o}{\beta}$$

$$\therefore\quad I_c\left[R_c+R_c\ (1+\beta)\right]=\beta V_{cc}+(1+\beta)\ I_{C_o}\ (R_o+R_c),$$

Differentiating with respect to I_{C_o}

$$\frac{\partial I_o}{\partial I_{Co}}\,[R_b+R_c\,(1+\beta)]=(R_b+R_c)\,(1+\beta)$$

$$\therefore \quad S=\frac{\partial I_o}{\partial I_{Co}}=\frac{(R_b+R_c)\,(1+\beta)}{R_b+R_c\,(1+\beta))}=\frac{(1+\beta)}{1+\dfrac{\beta R_c}{R_b+R_c}}$$

Obviously, from above equation $S < 1$, which is for stabilised circuit. For better stability, β should have high value, R_c large and R_b small. Values of R_b and R_c are, however not chosen purely by the condition of bias stabilisation, but R_b (base bias resistance) is determined by collector supply voltage and bias current, and R_c is determined by collector supply voltage and collector current

Example 1. *A PNP with* $\beta=25$ *is used in a common emitter amplifier with collector to base bias. The collector circuit resistance* $R_c=2{\cdot}5K\Omega$ *and* $V_{cc}=10V$. *Assuming* $V_{BE}=0$, *find* (*a*) *suitable value of* R_b *so as to make quiescent collector to emitter voltage* $V_{CE}=-5V$ (*b*) *the stability factor S.*

(a) The collector current I_o flowing through R_c is given by

$$I_c=\frac{-V_{cc}-V_{co}}{R_c}=\frac{-10-(-5)}{2{\cdot}5\times10^3}=-2\ mA.$$

Now $I_b=\dfrac{I_c}{\beta}=\dfrac{-2}{25}=-80\ \mu A$ (I_{Co} is assumed to be negligibly small)

According to the given problem, $V_{be}=0$, hence

$$R_b=\frac{V_{co}-V_{be}}{I_b}=\frac{-5-0}{-80\times10^{-6}}=\mathbf{62{\cdot}5}K\Omega$$

(b) We know that $S=\dfrac{(1+\beta)}{1+\beta\,[R_c/(R_b+R_c)]}$

Substituting the given values

$$S=\frac{1+25}{1+25\,[2{\cdot}5\times10^3/(62{\cdot}5\times10^3+2{\cdot}5\times10^3)]}$$

$$=\frac{26}{1+25\times(2{\cdot}5/65)}=\frac{26}{1{\cdot}96}=\mathbf{13{\cdot}32.}$$

16·10. STABILISATION BY POTENTIAL DIVIDER AND EMITTER RESISTANCE :

In fig. 17, R–C coupled common emitter amplifier is shown, R_1 and R_2 form a potential divider arrangement. R_c is a resistance in emitter lead.

Emitter voltage is practically equal to the base voltage the latter remains constant so long as I_b is constant. When temperature rises, leakage current increases and emitter voltage tends to approach base voltage, so that base emitter voltage becomes zero.

This reduces both I_b and I_c and each checks the increase in collector current due to I'_{C0}.

In practice base voltage is not constant because of the variation in I_b and hence complete stabilisation of plate current is not possible.

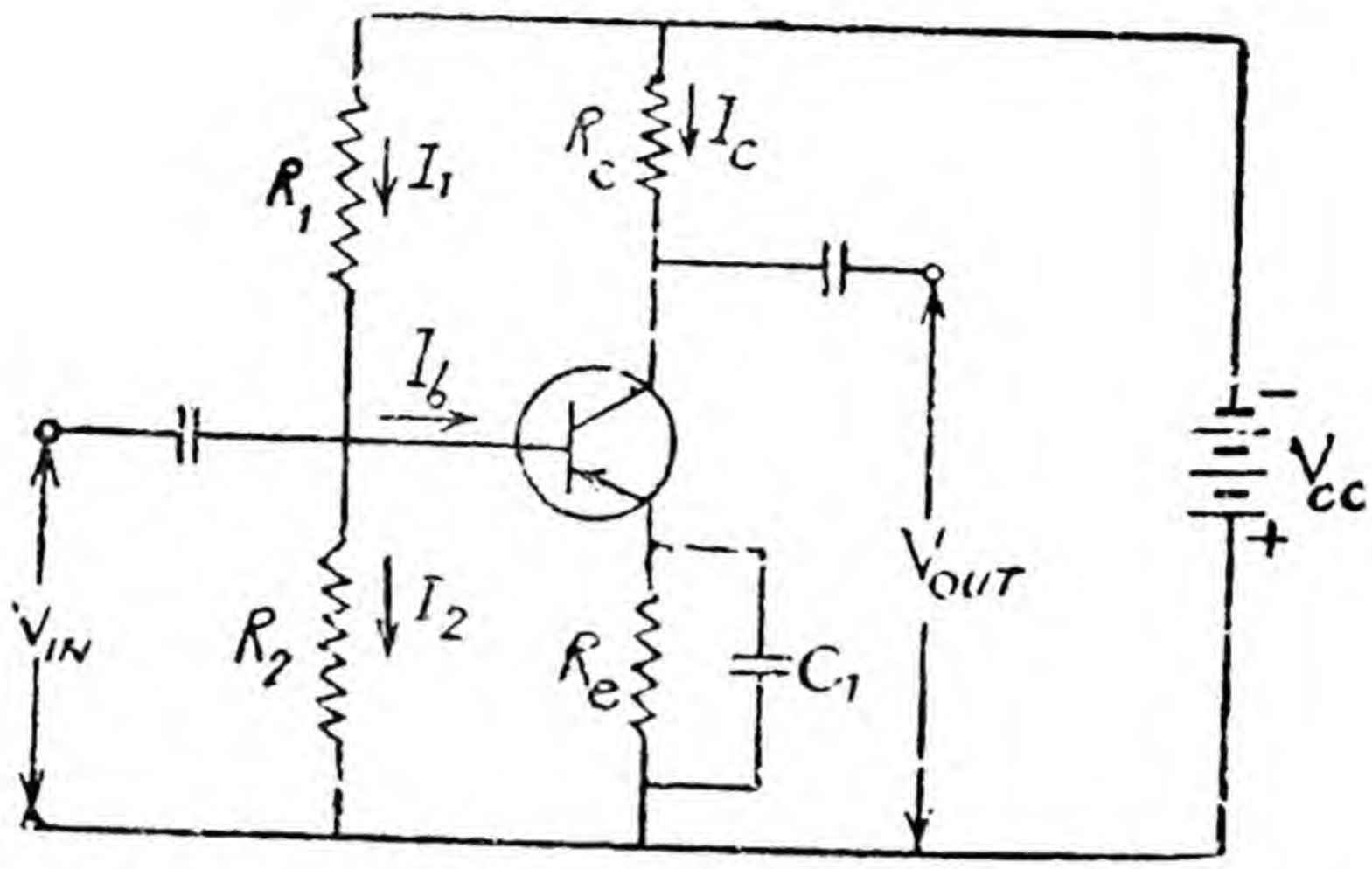


Fig 17. *R-C* coupling *C-E* amplifier with biassing arrangement.

Mathematically the condition can be described as follows :

$$I_c=\beta I_b+(1+\beta)\ I_{Co} \qquad ...(1)$$

From figure $I_1=I_b+I_2$

or $I_b=I_1-I_2.$ ...(2)

Therefore $I_c=\beta\ (I_1-I_2)+(1+\beta)\ I_{Co}$ ...(3)

and $V_{cc}=R_1I_1+R_2I_2.$ ...(4)

Multiplying equation (3) by R_1,

$$I_oR_1=\beta R_1\ (I_1-I_2)+R_1\ (1+\beta)\ I_{Co} \qquad ...(5)$$

and equation (4) by β,

$$\beta V_{oo}=\beta I_1R_1+\beta R_2I_2. \qquad ...(6)$$

Subtracting (5) from (6),

$$\beta V_{co}-I_cR_1=\beta I_2\ (R_1+R_2)-(1+\beta)\ I_{Co}$$

or $\beta V_{co}-R_1\ [I_o-(1+\beta)\ I_{Co}]=\beta I_2\ (R_1+R_2).$ ...(7)

Since voltage drop between base and emitter is very small,

$$I_2R_2=I_3R_e$$

or $$I_2=\frac{I_eR_e}{R_2}$$

But $I_e=I_b+I_c$. Putting this value of I_e in the expression for I_2, we get

$$I_2=\frac{(I_b+I_c)\ R_e}{R_2}.$$

Substituting this value of I_2 in equation (7), we get

$$\beta V_{cc} - R_1 [I_c - (1+\beta)\, I_{Co}] = \frac{\beta(I_b + I_c)\, R_e}{R_2}(R_1 + R_2)$$

$$= \frac{\beta R_c I_b\, (R_1 + R_2)}{R_2} + \frac{\beta R_e I_c (R_1 + R_2)}{R_2} \quad \ldots(8)$$

From equation (1), we get

$$\beta I_b = I_c - (1+\beta)\, I_{Co}.$$

Putting this value of βI_b in equation (8), we get

$$\beta V_{Co} - R_1 [I_c - (1+\beta)\, I_{Co}] = \frac{R_e\, [I_c - (1+\beta)\, I_{Co}]\, (R_1 + R_2)}{R_2}$$

$$+ \frac{\beta R_e I_c\, (R_1 + R_2)}{R_2},$$

or $$\frac{\beta V_{cc}}{R_1} - [I_c - (1+\beta)\, I_{Co}] = \frac{R_e (R_1 + R_2)\, [I_c - (1+\beta)\, I_{Co}]}{R_1 R_2}$$

$$+ \frac{\beta R_e I_c\, (R_1 + R_2)}{R_1 R_2}$$

$$\therefore \quad I_c \left[1 + \frac{R_e\, (R_1 + R_2)}{R_1 R_2} + \frac{\beta R_e\, (R_1 + R_2)}{R_1 R_2}\right]$$

$$= \frac{\beta V_{cc}}{R_1} + (1+\beta) I_{Co} \left[1 + \frac{R_e\, (R_1 + R_2)}{R_1 R_2}\right]$$

Let $\dfrac{R_1 R_2}{R_1 + R_2} = R_b$,

$$I_c \left[1 + \frac{R_e}{R_b}(1+\beta)\right] = \frac{\beta V_{cc}}{R_1} + (1+\beta)\, I_{Co} \left[1 + \frac{R_e}{R_b}\right].$$

Differentiating with respect to I_{Co}

$$\frac{\partial I_c}{\partial I_{Co}} \left[1 + \frac{R_e}{R_b}(1+\beta)\right] = \left[1 + \frac{R_e}{R_b}\right](1+\beta)$$

$$\therefore \quad \text{Stability factor } S = \frac{\left(1 + \dfrac{R_e}{R_b}\right)(1+\beta)}{1 + \dfrac{R_e}{R_b}(1+\beta)}$$

$$= \frac{(1+\beta)}{1 + \dfrac{\beta R_e}{R_e + R_b}}. \quad \ldots(9)$$

Equation (9) shows S to be less than one—the value demanded by stabilisation. To secure better stability, β and R_e should be large but R_b should be small.

The important feature of this method of stabilisation, in contrast to previous one, is that R_b (hence R_1 and R_2) and R_e can be valued within large limits to obtain desired stability, S is improved if R_c is increased but too high a value will also limit the collector current.

Example 1. *In a self bias circuit of fig.* (17) *a silicon NPN transistor is used with* $\beta=50$, $V_{bb}=0{\cdot}6V$, $V_{cc}=18V$ *and* $R_e=4{\cdot}3K\Omega$. *It is desired to establish a quiescent point at* $I_c=1{\cdot}5\,mA$, $V_{ce}=10V$ *and a stability factor* $S \leqslant 4$. *Find* R_e, R_1 *and* R_2.

The current through R_e is $I_e+I_b \approx I_c \quad \because \quad I_b \ll I_c$

$$R_c+R_e=\frac{V_{cc}-V_{ce}}{I_c}=\frac{18-10}{15\times10^{-3}}=5{\cdot}33\ K\Omega$$

or $$4{\cdot}3+R_e=5{\cdot}33$$

or $$R_e=5{\cdot}33-4{\cdot}3=1{\cdot}03K\Omega \approx 1\ K\Omega.$$

Now $$S=\frac{(1+\beta)}{1+\beta.R_e/(R_e+R_b)}$$

Solving of R_b/R_e, we have

$$\frac{R_b}{R_e}=\frac{(1+\beta)\ (S-1)}{(1+\beta-S)}=\frac{(1+50)\ (4-1)}{(1+50-4)}=3{\cdot}25$$

.. $$R_b=R_e\times3{\cdot}25=1\times3{\cdot}25=3{\cdot}25K\Omega$$

$$I_b \approx \frac{I_c}{\beta} \approx \frac{1{\cdot}5mA}{50} \approx 30\mu A$$

For a self bias arrangement,

$$V_{Th}=V_b=\frac{R_2}{R_1+R_2}.V_{cc} \quad \text{and} \quad R_{Th}=R_b=\frac{R_1R_2}{R_1+R_2}$$

Solving we get,

$$R_1=R_b\ \frac{V_{cc}}{V_b} \quad \text{and} \quad R_2=\frac{R_1V_b}{V_{cc}-V_b}$$

Also $$V_b=I_b\ R_b+V_{be}-(I_b+I_e)\ R_e$$

Substituting the given values, we have

$$V_b=(30\times10^{-6})\ 3{\cdot}25\times10^3+0{\cdot}6$$
$$+(30\times10^{-3}+1{\cdot}5\times10^{-3})\ (1\times10^3)$$
$$=2{\cdot}23 \text{ volts.}$$

Hence $$R_1=\frac{3{\cdot}25\times18}{2{\cdot}23}\ k\Omega=26{\cdot}23\ K\Omega$$

and $$R_2=\frac{26{\cdot}23\times2{\cdot}23}{18-2{\cdot}23}\ k\Omega=3{\cdot}7K\Omega.$$

Example 2. *A common emitter power amplifier uses an NPN transistor having* $\alpha=0{\cdot}95$. *It is operated with* $I_c=500\ mA$ *and* $V_{cc}=15$ *volts. Collector reverse saturation current is likely to vary from* 5 *to* 15 *mA over the working range of temperat ure. If the variation in the collector current* I_c *is to be kept within* 50*mA, design a stabilizing current using self bias such that the voltage drop in emitter circuit is not appreciably more than* 5 *volts. Assume* $V_{be}=0$.

$\Delta I_{C_0}=(15-5)=10mA$ and ΔI_c should not exceed $50mA$.

$$\therefore \qquad S=\frac{\Delta I_c}{\Delta_{C_0}}=\frac{50\text{mA}}{10mA}=5.$$

This is the maximum value of stablity factor.

The maximum value of R_e is given by

$$R_e=\frac{5 \text{ volt}}{I_e}=\frac{5 \text{ volt}}{I_c/\alpha}=\frac{5\times 0{\cdot}95}{500\times 10^{-3}}=9{\cdot}5\Omega.$$

For better stabilization, R_e may be used as 10Ω.

Now $$\beta=\frac{\alpha}{1-\alpha}=\frac{0{\cdot}95}{1-0{\cdot}95}=19$$

We know that

$$S=\frac{1+\beta}{1+\beta.R_e/(R_e+R_b)}$$

Solving we get

$$\frac{R_b}{R_e}=\frac{(1+\beta)\,(S-1)}{(1+\beta-S)}=\frac{(1+19)\,(5-1)}{(1+19-5)}=\frac{16}{\cdot}$$

$$\therefore \qquad R_b=R_e\times(16/3)=10\times(16/3)=53{\cdot}3\Omega$$

Assuming $V_{be}=0$, $V_b=$ voltage drop across $R_e=5$ volt

But $$V_b=\frac{R_2\,V_{cc}}{R_1+R_2}=\frac{15\,R_2}{R_1+R_2}$$

$$\therefore \qquad 5=\frac{15\,R_2}{R_1+R_2} \qquad \ldots(1)$$

Again $$R_b=\frac{R_1R_2}{R_1+R_2}=53{\cdot}3 \qquad \ldots(2)$$

Solving eqs. (1) and (2), we get

$$R_1=159\Omega \text{ and } R_2=79{\cdot}5\Omega.$$

16·11. TWO BATTERY BIAS STABILISATION :

In fig. 18, emitter bias is supplied by one battery of potential V_1 through resistor R_1 and battery V_2 supplies potential to the collector through resistor R_c.

If R_1 is sufficiently large compared to r_e, the internal emitter resistance, and the total resistance in base circuit is not excessive, the emitter current will be relatively constant. Since in junction transistors collector current is essentially independent of collector potential, the collector potential will remain constant. To achieve such an arrangement, large power is dissipated in stabilising resistors.

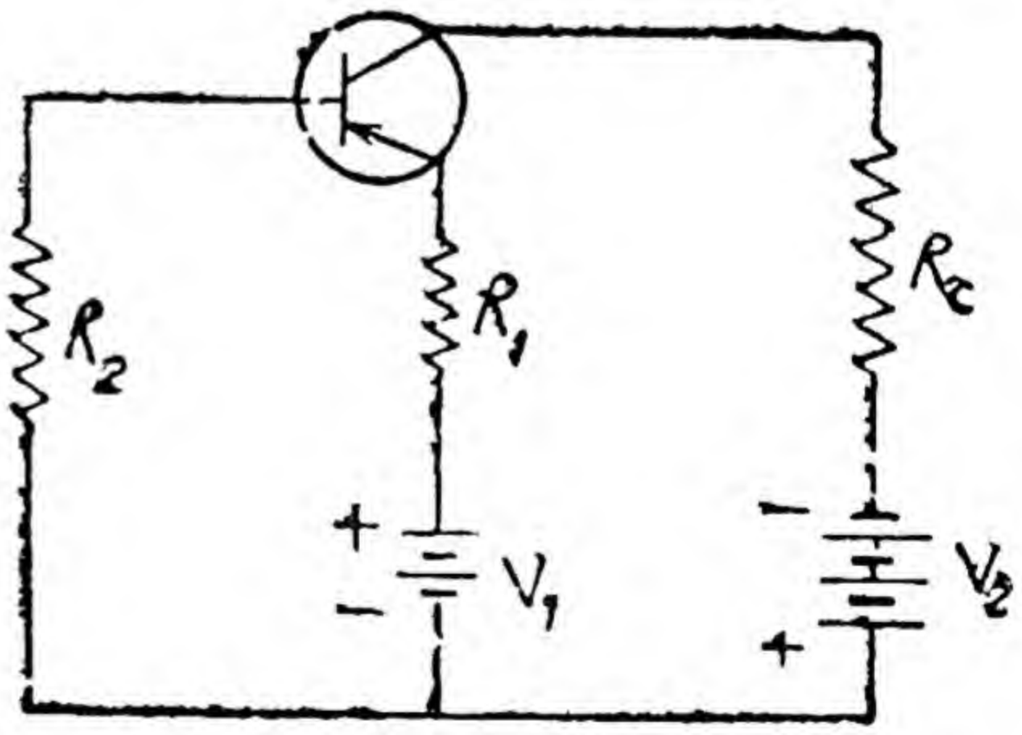


Fig 18. Two battery bias stabilisation.

16·12. BIAS COMPENSATION

We have seen that a fixed bias circuit has a poor operating point stability while self bias circuit has better operating point stability. In self bias circuit, the current I_b varies in such a way that I_c

remains constant inspite of variation in I_{CO}, V_{be} and β. In some applications, it is found useful to employ compensation techniques to reduce the drift of operating point. Very often both stabilization and compensation techniques are used to provide maximum bias and thermal stabilization. The bias compensation methods are discussed below :

(1) Diode compensation for I_{Co}—In case of a germanium transistor, the change of I_{Co} with temperature contributes more towards change in I_c. Fig. (19) shows the circuit of a germanium transistor amplifier with diode D used to compensate the variation in I_{Co}. A diode of the same material as the transistor is reverse biased by the base-emitter junction voltage V_{be}, allowing leakage current I_0 to flow through diode D. The principle of this method is that when any change in temperature causes the increase in I_{Co}, the leakage current I_0 in diode D also increases in such a way as to keep I_b constant. From fig. (19),

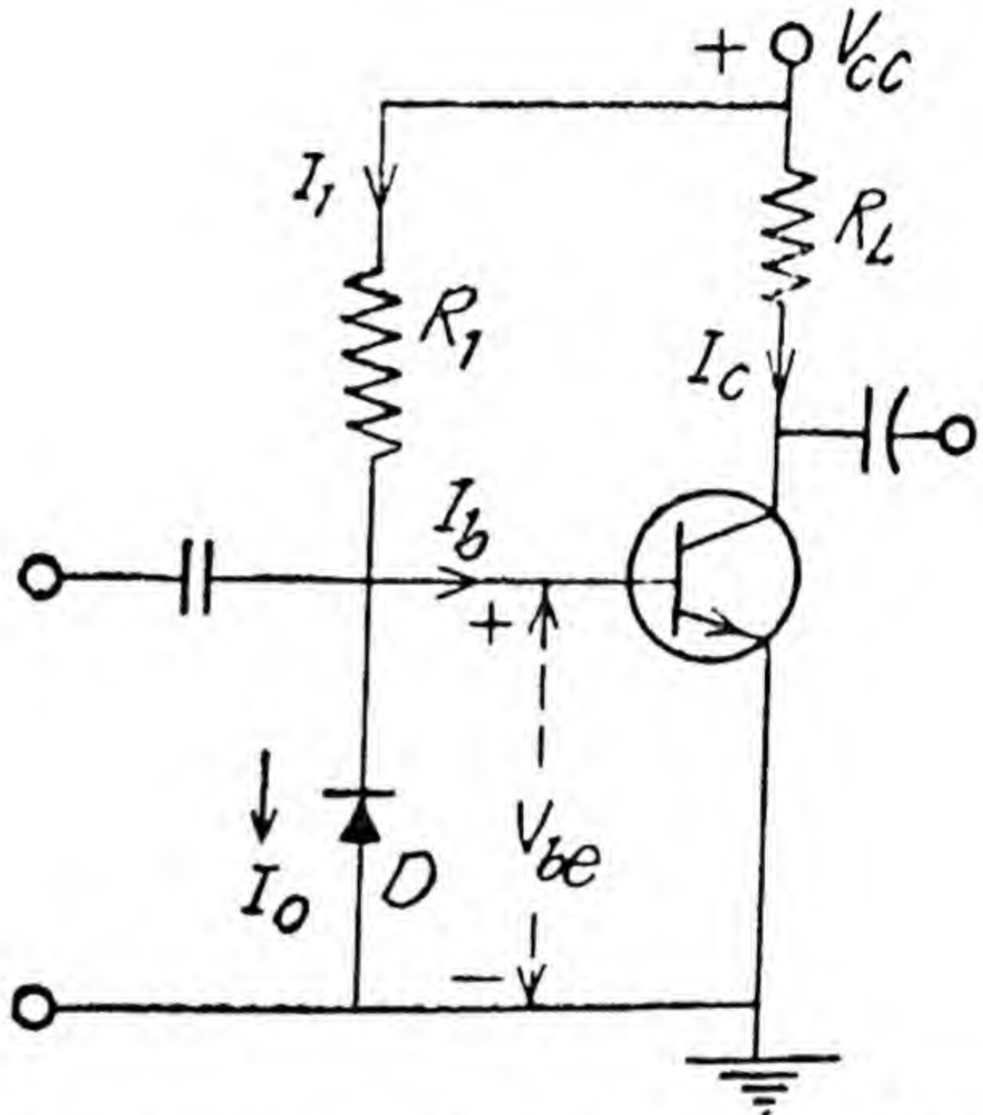


Fig. 19. Diode compensation for germanium transistor.

$$I_1 = \frac{V_{co} - V_{be}}{R_1} \approx \frac{V_{cc}}{R_1} = \text{constant}$$

and $$I_b = I_1 - I_0$$

We know that $I_c = (1 + \beta;\ I_{Co} + \beta I_b$

or $$I_c = (1+\beta)\ I_{Co} + \beta\ (I_1 - I_0)$$

or $$I_c = (1+\beta)\ I_{Co} + \beta I_1 - \beta I_0$$

If $\beta \gg 1$, then $I_c = \beta\ I_{Co} + \beta I_1 - \beta I_0$

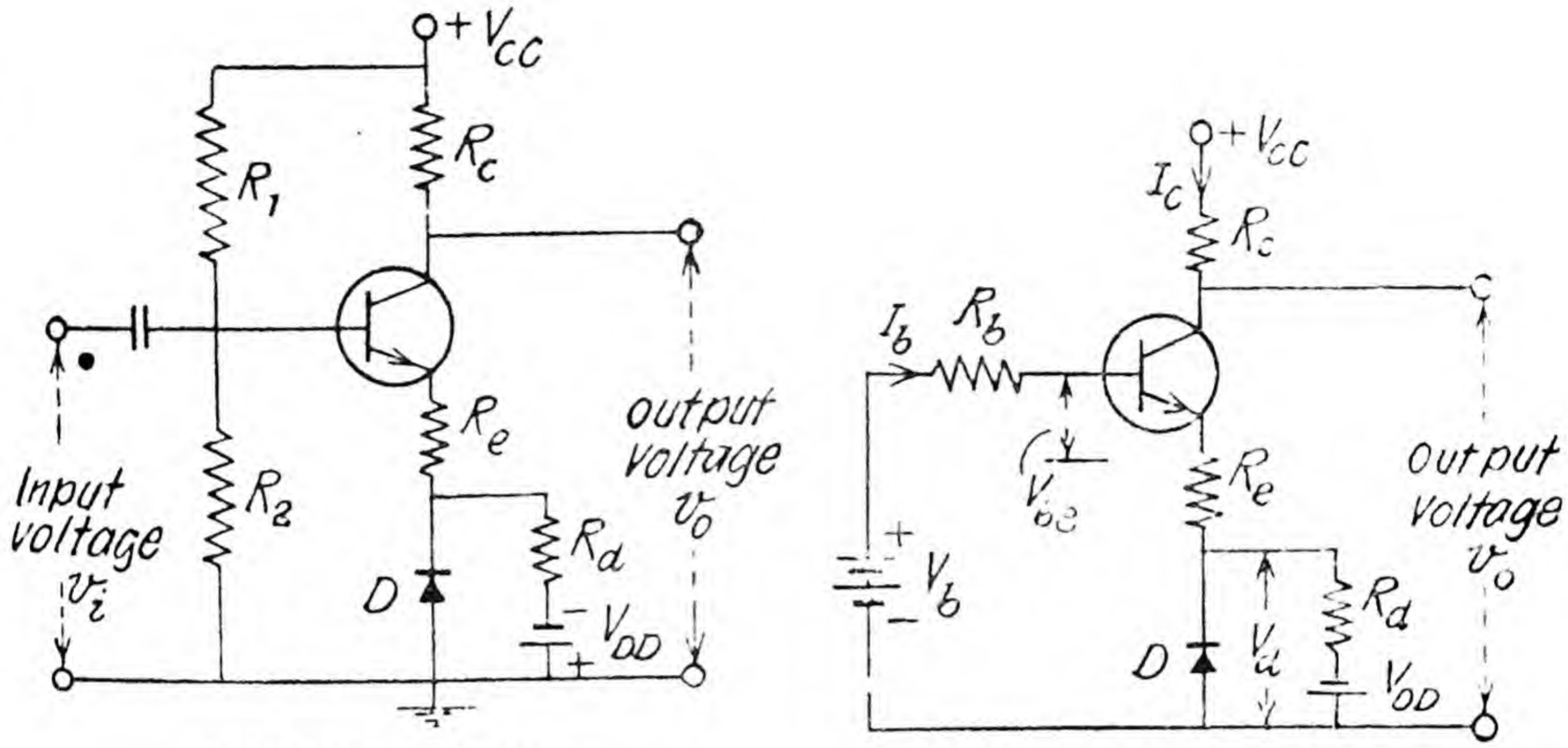
So if I_0 of diode and I_{Co} of transistor track each other over the desired range of temperature, then I_c remains essentially constant as I_1 is constant.

(2) Diode Compensation for V_{be}—In case of a silicon transistor, the change of V_{be} with temperature contributes significantly to the change in collector current. Fig. 20 (*a*) and (*b*) show the circuit diagram and equivalent circuit for self bias circuit with diode compensation for V_{be}.

The diode D is forward biased by the source V_{DD} and resistance R_d. Diode D is of the same material as the transistor so that the voltage V_d across the diode has the same temperature coefficient as V_{be}. From base circuit, we have

$$V_b - V_{be} + V_d = I_b R_b + (I_c + I_b)\ R_e \qquad \ldots(1)$$

But $$I_c = (1 - \beta)\ I_{Co} + \beta I_b \qquad \ldots(2)$$



(a) Circuit diagram (b) Equivalent circuit

Fig. 20. Self bias circuit with diode compensation for V_{be}.

Eliminating I_b from eqs. (1) and (2), we get

$$V_b - V_{be} + V_d = \frac{R_b + R_e(1-\beta)}{\beta} I_c - (R_b + R_e)\frac{1+\beta}{\beta} I_{Co}$$

$$\therefore \quad I_c = \frac{\beta[V_b - (V_{be} - V_d)] + (R_b + R_e)(1+\beta) I_{Co}}{R_b + R_e(1+\beta)} \quad \ldots (3)$$

Since V_{be} tracks V_d with respect to variations in temperature, it is clear from eq. (3) that I_c will become insensitive to variation in V_{be}.

16·13. CURRENT AND VOLTAGE AMPLIFIERS :

Voltage amplifiers : The prime requirements for voltage amplifier are :

(*a*) when a signal generator is connected at its input, generator voltage should not change. For this condition to be satisfied input resistance (r_i) of the amplifier should be high ;

(*b*) any change in load resistance should not affect the output voltage of the amplifier. This condition demands that output resistance of the amplifier (r_0) should be low.

For class *A* amplifiers, $r_i \simeq \infty$ and $r_0 < 1 \ k\Omega$ and therefore this stage can be regarded as voltage amplifier.

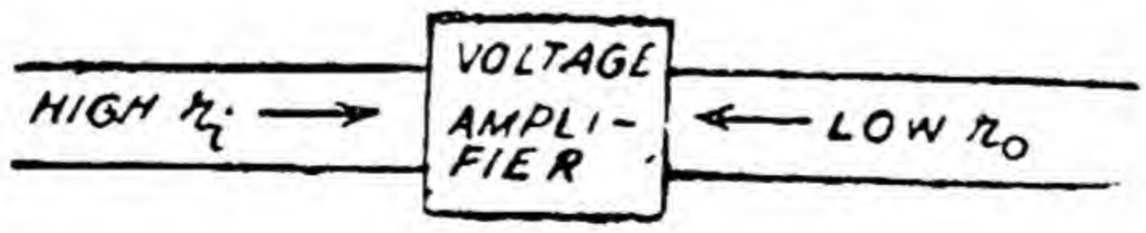


Fig 21. Pictorial representation of the voltage amplifier conditions.

Out of three configurations, only common collector amplifier has a high input resistance and a low output resistance and is therefore, operated as voltage amplifier.

Current amplifier : The essential features of a current amplifier are :

(*a*) when a signal generator is connected as its input, generator current should not change. This condition demands low value of input resistance (r_i).

(*b*) any change in load resistance should not affect the output current, *i.e.*, output current of the amplifier should be independent of load resistance. The condition demands a high value of output resistance (r_0) of the amplifier.

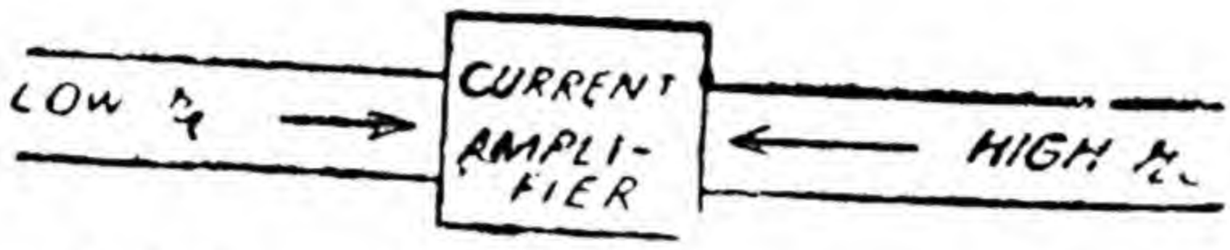


Fig. 22. Requirement of current amplifier.

A common base and common emitter amplifier stage have small input resistance and high output resistance. Thus both configurations serve as current amplifier.

6·14. THE COMMON EMITTER AMPLIFIER WITH EMITTER RESISTOR ;

We know that transistor parameters depend upon temperature, operating point, etc. Moreover, the parameters differ from device to device even for same type of transistor. The voltage gain, therefore, is not stabilize. To stablise the voltage gain a resistor is included in emitter circuit as shown in fig. (23*a*) The emitter resistor produces inverse feed back which causes the stabilization of voltage gain. Fig. (23*b*) shows the equivalent

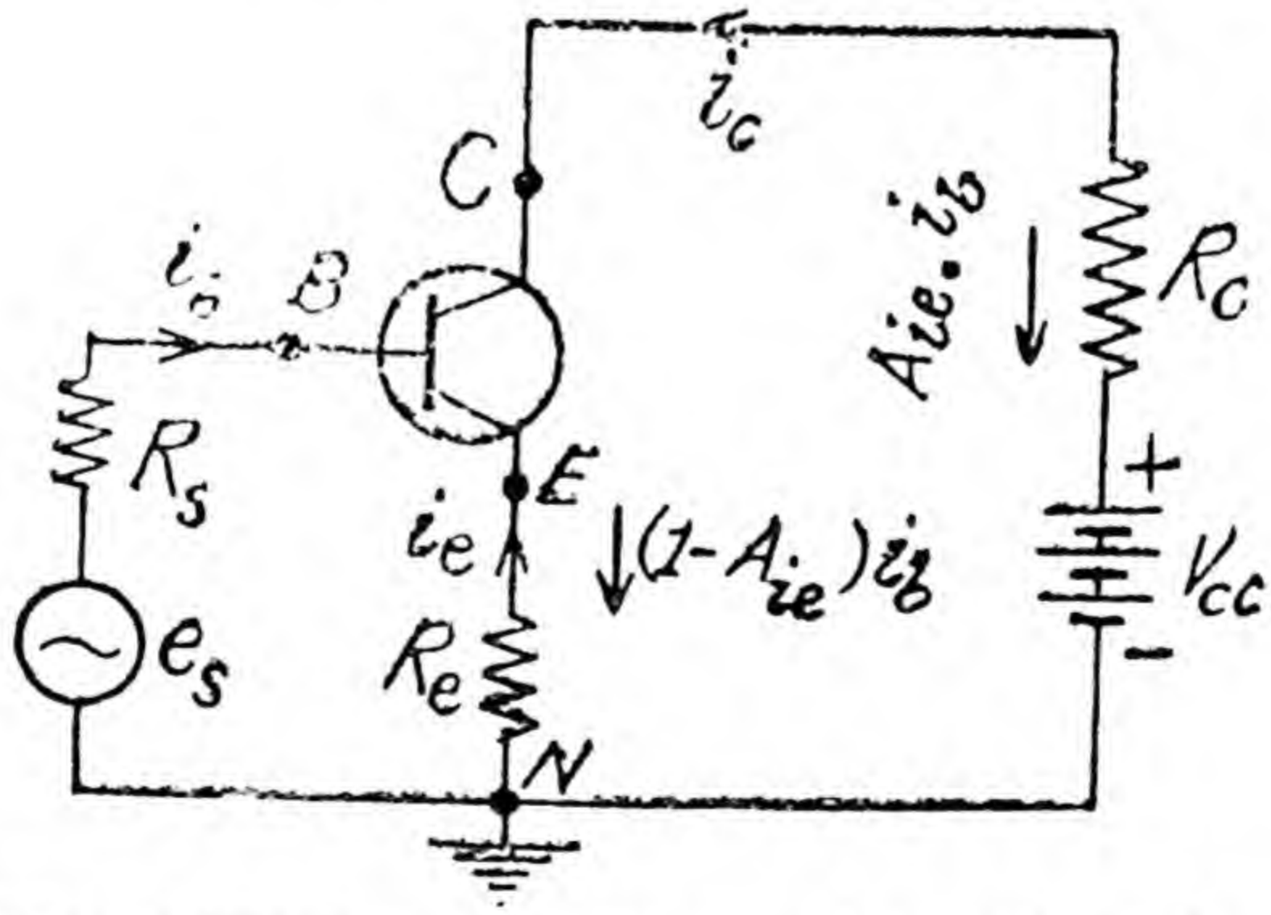
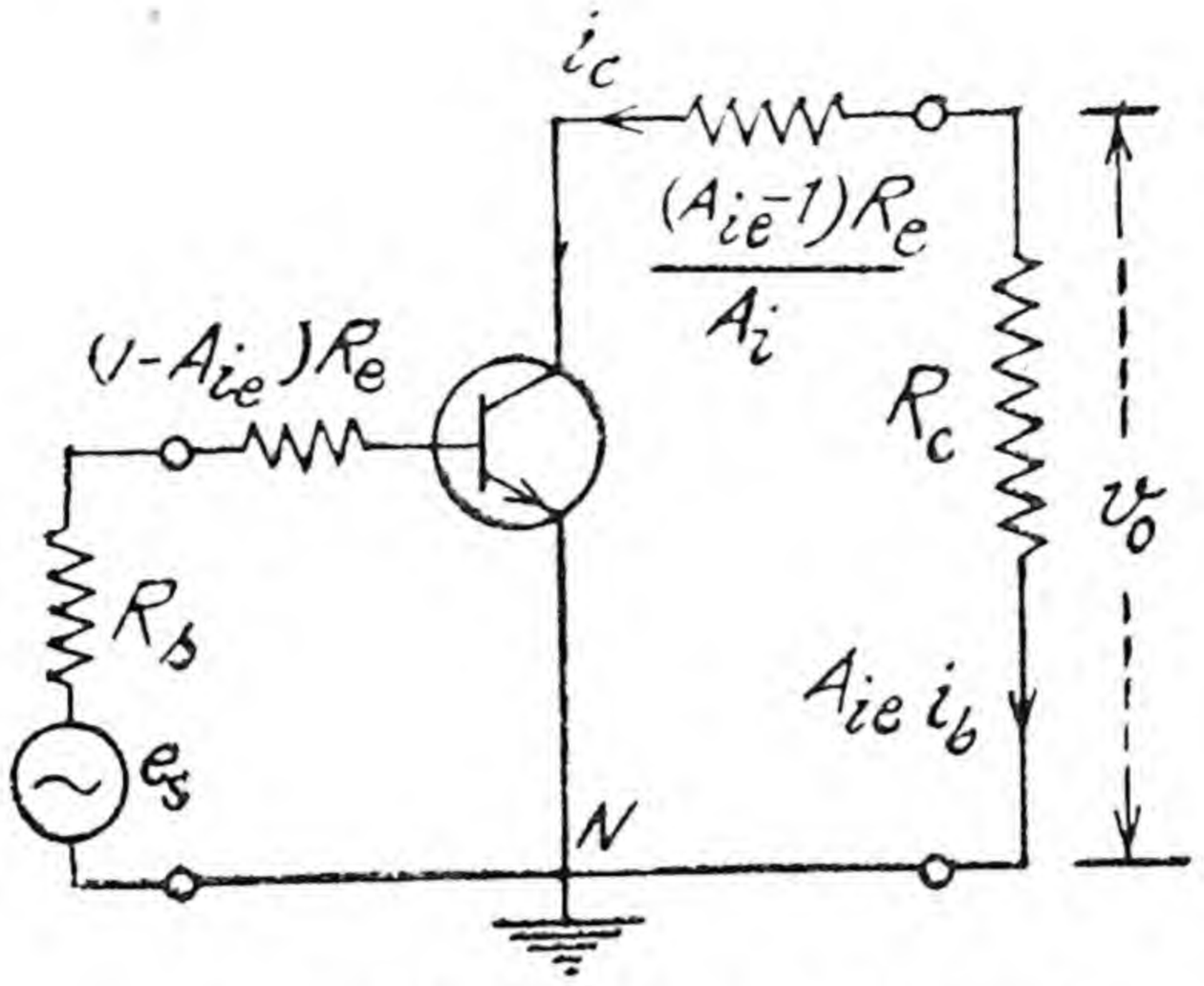


Fig. 23. (a) *CE* amplifier with emitter resistor R_e without base biasing network

circuit. The current in load resistor R_c [fig. (2 *a*)] has been shown as $A_{ie}\, i_b$ where A_{ie} is the ratio of collector current to base current *i.e.*, $A_{ie} = -i_c/i_b$. Now the current in R_e is $(1-A_{ie})\, i_b$. In fig. (23*b*) the emitter resistor R_e has been replaced by two resistors as $(1-A_{ie})R_e$



(b) Small signal equivalent circuit

Fig. 23. Common emitter amplifier with emitter resistor.

in base curcuit and $(A_{ie}-1)R_e/A_{ie}$ in the collector circuit. If we apply Kirchhoff's voltage law to collector circuits of fig. 23(a) and (b), the collector current is the same and hence the two circuits are equivalent.

(1) We know that the current gain of a common emitter amplifier is given by

$$A_{ic}=\frac{-h_{fe}}{1+h_{oe}\,R_L} \qquad ...(1)$$

where R_L is effective load impedance.

The effective load impedance in our present case is given by

$$R_L=R_c+\frac{(A_{ic}-1)\,R_e}{A_{ie}} \qquad ...(2)$$

Substituting the value of R_L from eq. (2) in eq. (1), we have

$$A_{ic}=\frac{-h_{fe}}{1+h_{oe}\,[R_c+\{(A_{ie}-1)R_e\}/A_{ie}]}$$

Solving this equation for A_{ie} we have

$$A_{ie}=\frac{h_{oe}\,R_e-h_{fe}}{1+h_{oe}\,(R_c+R_e)} \qquad ...(3)$$

(2) The input resistance R_{ie} is given by

$$R_{ie}=\frac{v_i}{i_b}=(1-A_{ie})\,R_e+h_{ie}+h_{re}A_{ie}R_L$$

or
$$R_{ie}=(1-A_{ie})\,R_e+h_{ie}+h_{re}A_{ie}\left[\,R_c+\frac{(A_{ie}-1)}{A_{ie}}R_e\right] \qquad ...(4)$$

(3) The voltage gain is given by the expression

$$A_{ve}=\frac{A_{ic}.R_c}{R_{ie}}, \qquad ...(5)$$

where A_{ie} and R_{ie} are values as given by eqs. (3) and (4).

(4) The output impedance R_{oc} of this amplifier is given by

$$R_{oc}=h_{oc}+(R_S+h_{ic})\left(1+\frac{1}{h_{oc}R_e}\right) \text{ where } (R_e \gg R_S+h_{ie}) \quad ...(6)$$

16·15. SIMPLIFIED COMMON EMITTER HYBRID MODEL

Fig. (24) shows an exact equivalent circuit of a common emitter hybrid model.

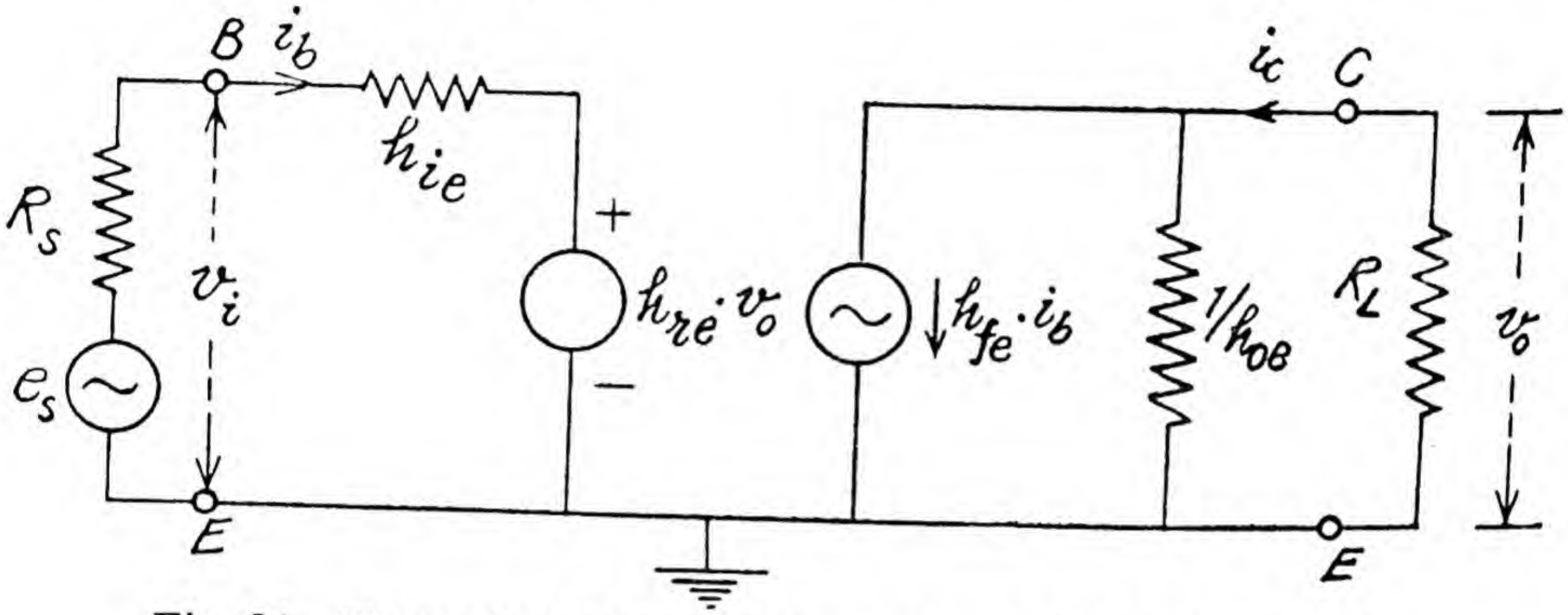
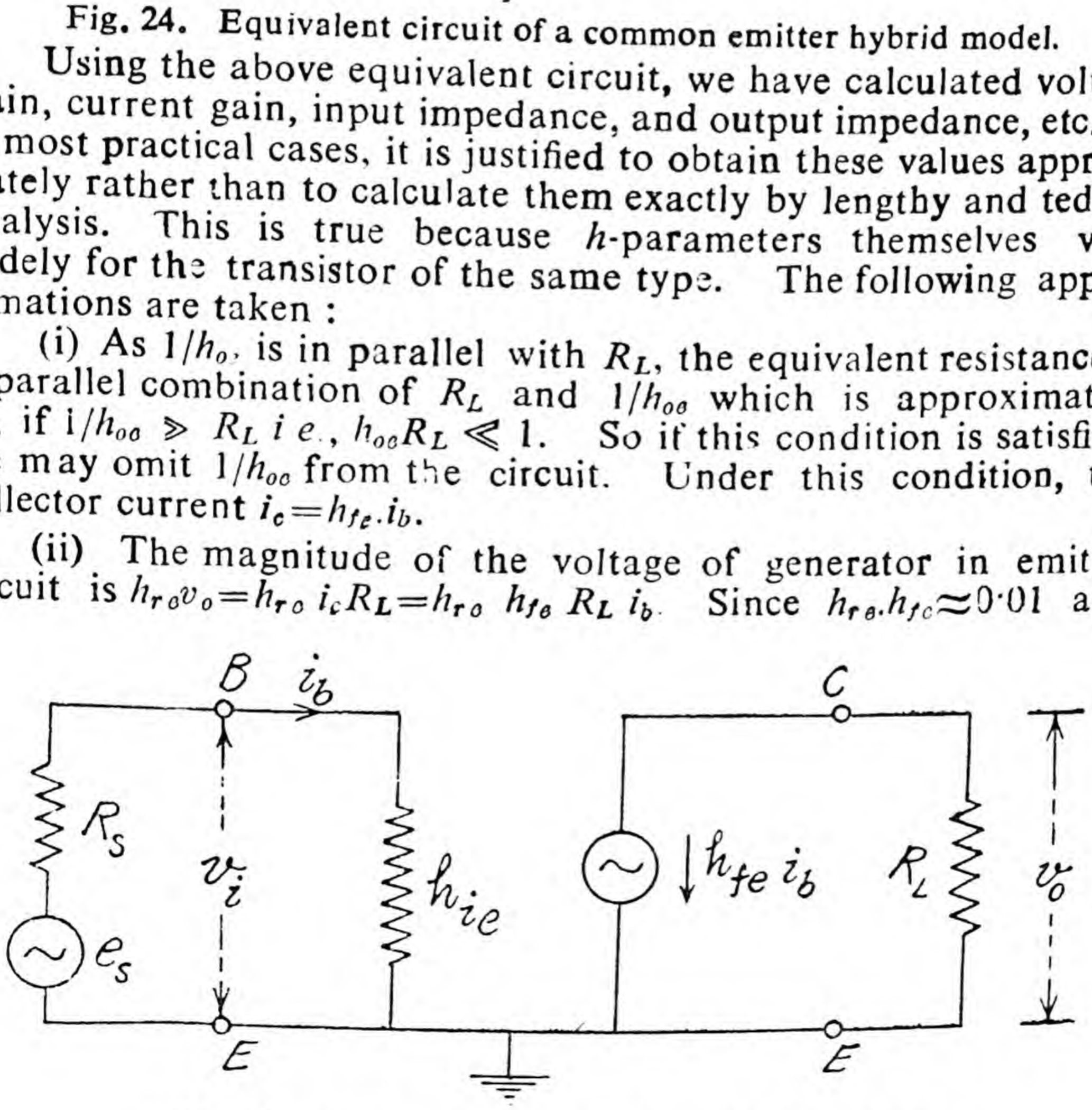


Fig. 24. Equivalent circuit of a common emitter hybrid model.

Using the above equivalent circuit, we have calculated voltage gain, current gain, input impedance, and output impedance, etc. But in most practical cases, it is justified to obtain these values approximately rather than to calculate them exactly by lengthy and tedious analysis. This is true because *h*-parameters themselves vary widely for the transistor of the same type. The following approximations are taken :

(i) As $1/h_o$, is in parallel with R_L, the equivalent resistance is a parallel combination of R_L and $1/h_{oc}$ which is approximately R_L if $1/h_{oc} \gg R_L$ *i e.*, $h_{oc}R_L \ll 1$. So if this condition is satisfied, we may omit $1/h_{oc}$ from the circuit. Under this condition, the collector current $i_c=h_{fe}.i_b$.

(ii) The magnitude of the voltage of generator in emitter circuit is $h_{re}v_o=h_{re}\, i_cR_L=h_{re}\ h_{fe}\, R_L\, i_b$. Since $h_{re}.h_{fe}\approx 0{\cdot}01$ and



Flg. 25. Simplified common emitter hybrid model.

hence this voltage may be neglected in comparison with $h_{ie}\, I_b$ drop across h_{ie} provided that R_L is not too large.

Thus we conclude that if R_L is small, then we can neglect h_{re} and h_{oe} from the exact equivalent circuit. Now the approximate equivalent hybrid model at low frequencies is shown in fig. (25).

16·16. EFFECT OF AN EMITTER BYPASS CAPACITOR IN LOW FREQUENCY RESPONSE

Fig. (26) shows the circuit of a common emitter with a resistor R_e and capacitor C_e in the emitter lead. The aim of this article is to consider the effect of C_e on the low frequency response. Here we assume that the parallel combination of R_1 and R_2 is much larger than input resistance and hence the voltage gain is unaffected by the presence of R_1 and R_2. They provide a self biasing arrangement

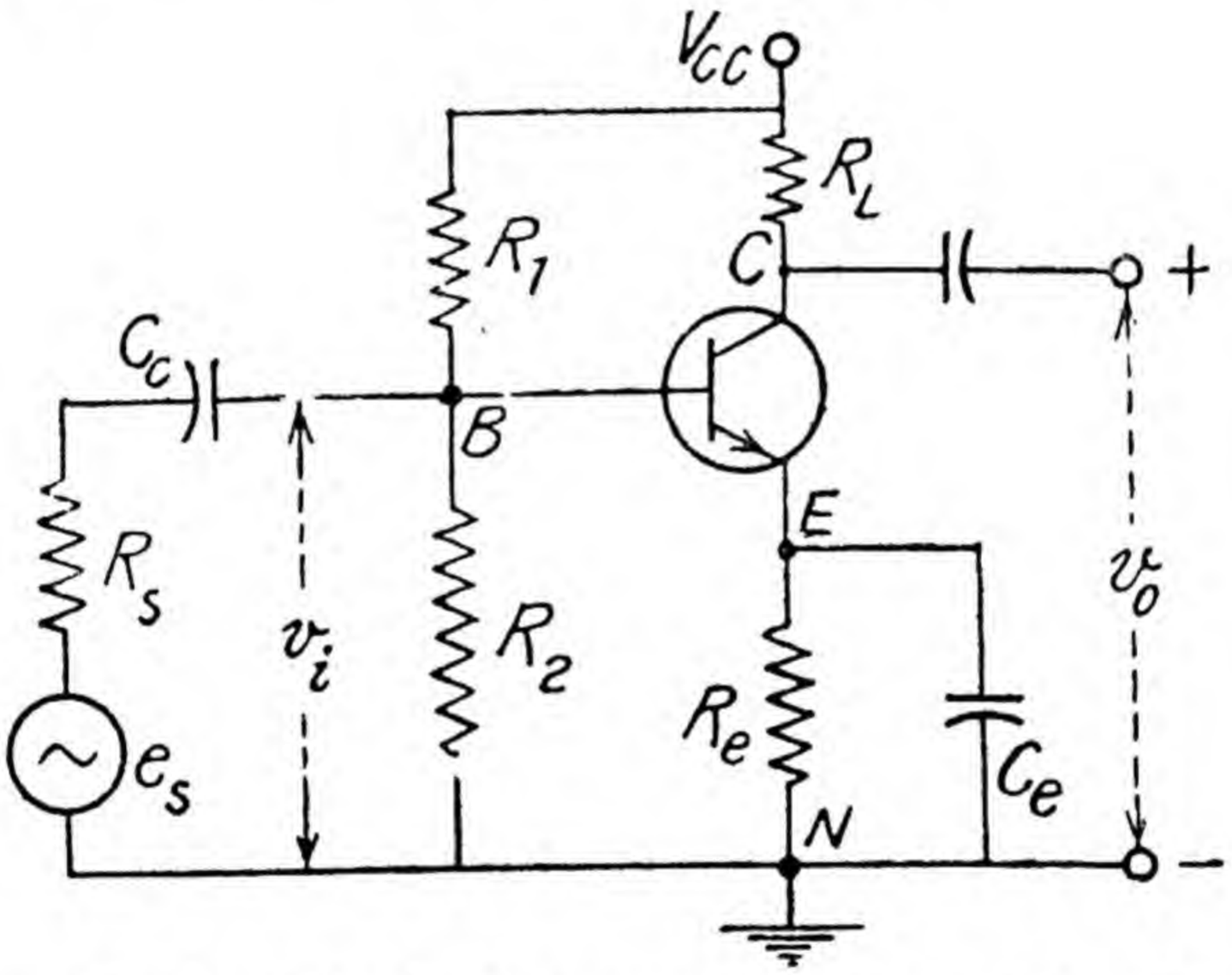


Fig. 26. Circuit of a common emitter with resistor R_e and capacitor C_e.

An approximate small signal hybrid equivalent circuit at low frequencies for the common emitter amplifier is shown in fig. (27).

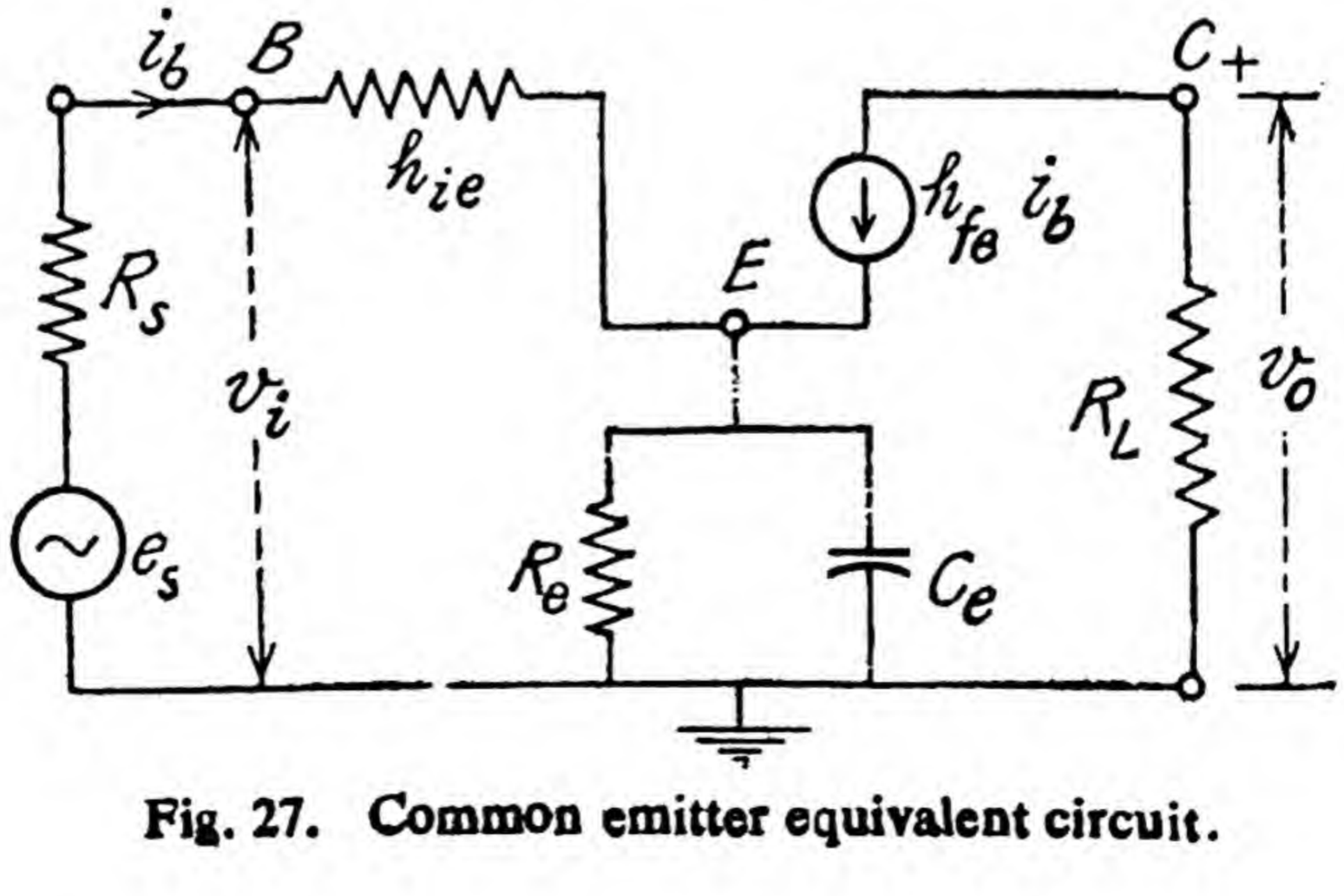


Fig. 27. Common emitter equivalent circuit.

From fig. (27), $v_0 = -h_{fe} i_b R_L$ ...(1)

The value of i_b is given by

$$i_b = \frac{e_S}{R_S + R_i} \quad ...(2)$$

where, for common emitter with an unbypassed R_e,

$$R_i = h_{ie} + (1 + h_{fe})\, Z_e. \quad ...(3)$$

Here Z_e is a parallel combination of R_e and C_e, *i.e.*,

$$Z_e = \frac{R_e}{1 + j\omega C_e R_e} \quad .. (4)$$

$$\therefore \quad v_0 = -h_{fe}\, R_L . \frac{e_S}{R_S + h_{ie} + \dfrac{(1 + h_{fe})\, R_e}{1 + j\omega C_e R_e}} \quad ...(5)$$

Now voltage gain at low frequency is given by

$$(A_{ve})_{L\cdot F.} = \frac{v_0}{e_S} = \frac{-h_{fe} R_L}{R_S + h_{ie} + \dfrac{(1 + h_{fe})\, R_e}{1 + j\omega C_e R_e}} \quad ...(6)$$

when ω is large, *i.e.*, at mid frequency range, the voltage gain is given by

$$(A_{ve})_{M\cdot F.} = \frac{-h_{fe} R_L}{R_S + h_{ie}}. \quad ...(7)$$

16·17. HIGH FREQUENCY EFFECTS AND HYBRID PI MODEL :

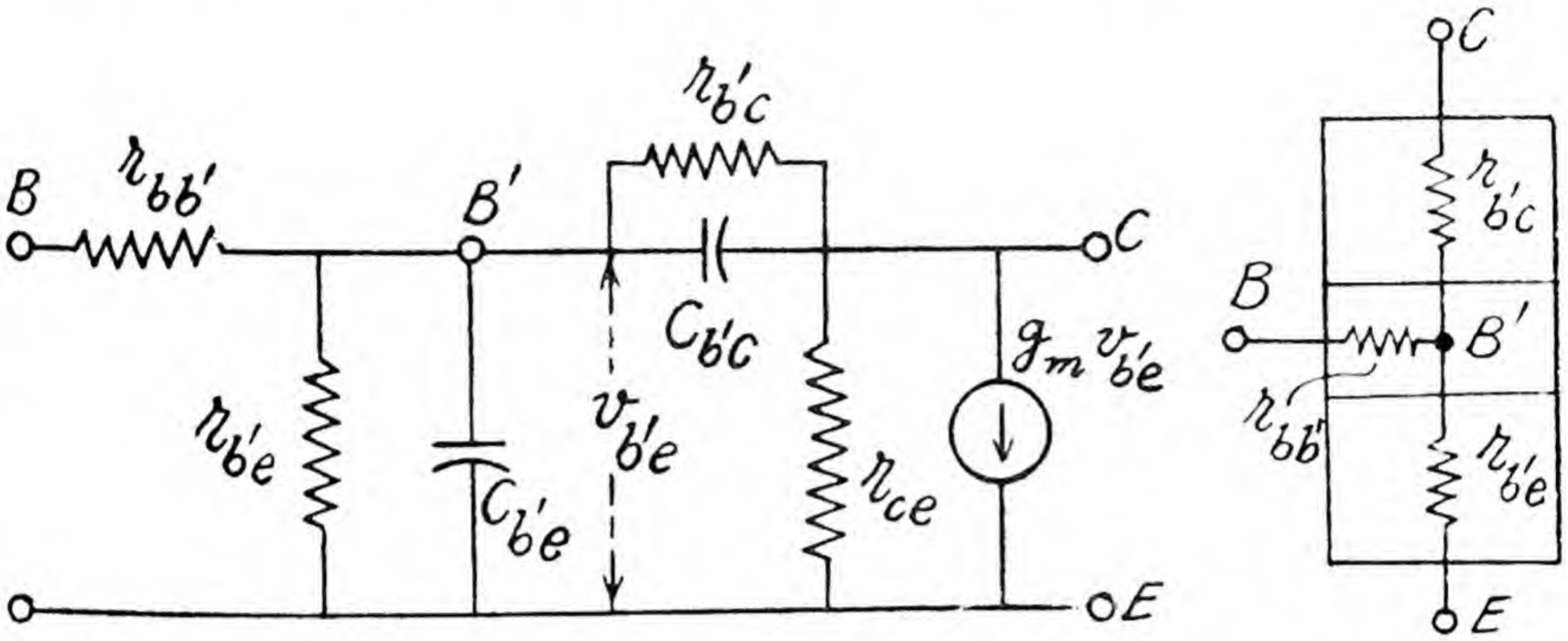
The aim of this article is to consider the equivalent circuit of common emitter configuration which is valid at high frequencies. At low frequencies a transistor responds instantaneously to changes of input voltage or current but at high frequencies. the mechanism of transport of charge carriers from emitter to collector is essentially one of diffusion. The diffusion involves time delay. So at high frequencies the diffusion mechanism should be taken into consideration. The diffusion analysis is very complicated. Here we shall describe the hybrid π model or Giacoletto model of the transistor in common emitter configuration at high frequencies. This model is simple, accurate and gives result which are in good agreement with experiments at all frequencies and hence popularly used. Following assumptions are made in this model :

(i) All the parameters (resistances and capacitances) are independent of frequency

(ii) The resistive component in this model may be obtained from the low frequency *h* parameters.

(iii) The parameters may vary with quiescent operating point but under given bias conditions they are reasonably constants for small signal variations.

Fig. 28 (*a*) shows the hybrid π model for a transistor in common emitter configuration and fig. 28 (*b*) shows the ohmic base spreading resistance $r_{bb'}$ and virtual base B'.



(a) Hybrid π model for a *CE* amplifier. (b) Showing virtual base.

Fig. 28.

Different parameters :

The Hybrid π Capacitances. Using a *PNP* transistor, when the input voltage increases, there is an increase in hole density in the base. To maintain the charge neutrality, electrons enters the base. When the input voltage falls, the stored charge returns to the circuit. Thus there is a charge flow or current which occurs with change in voltage. To represent this charge effect, we place a capacitor $C_{b'e}$ (diffusion capacitance) between B' and E.

The depletion layer of collector-base reversed biased junction also acts as a capacitance and is shown by $C_{b'c}$ between B' and C.

Base spreading resistance. The two capacitances discussed above have a common terminal at B' and not at the base terminal B. B' is known as internal node and is not physically accessible. The ohmic base-spreading resistance $r_{bb'}$ is represented as a lumped parameter between external base B and internal node B'. This represents the bulk resistance of the base and is the portion of the base through which the base current proper flows.

The increase in the minority carriers in the base results in increased recombination base current. This effect is taken into account by inserting a resistance $r_{b'e}$ between the internal node B' and E.

At low frequencies, $h_{ie} = r_{bb'} + r_{b'e}$.

The varying voltage across the collector to emitter junction results in base-width modulation. A change in the effective base width causes the emitter and collector current to change. This feed back effect between base and collector is taken into account

by connecting $r_{b'c}$ between internal node B' and C. This resistance is typically $4M$ ohms and can normally be omitted.

Resistance r_{ce} is the output resistance. When a load is connected between collector and emitter, this resistance may be omitted because $r_{ce} \gg R_L$.

Current generator. The term $v_{b'e}$ represents the voltage across the emitter junction. For small changes in $v_{b'e}$ across the emitter junction the extra minority carrier concentration injected into base region is proportional to $v_{b'e}$. Corresponding, small-signal collector current with the collector shorted to the emitter is, therefore, proportional to $v_{b'e}$. This effect accounts for the current generator $g_m v_{b'e}$ across the termina! C and E. Here g_m is the transconductance.

16·18. RELATION BETWEEN HYBRID π CIRCUIT COMPONENTS AND LOW FREQUENCY h-PARAMETERS

(1) **Transistor mutual conductance g_m**—Considering the case of a *PNP* transistor in common emitter configuration, the collector current is given by

$$i_c = i_{Co} - \alpha i_e$$

In this case, the emitter is forward biased while the collector is reversed biased. The transconductance is defined by

$$g_m \equiv \left.\frac{\partial i_c}{\partial v_{b'e}}\right|_{v_{cb}} = -\alpha \frac{\partial i_e}{\partial v_{b'e}} = \alpha \frac{\partial i_e}{\partial v_e} \qquad \ldots(1)$$

where v_e is the voltage of emitter with respect to B' and is equal to $-v_{b'e}$. It is assumed that α is independent of v_e.

The emitter diode resistance r_e is given by

$$r_e = \frac{\partial v_e}{\partial i_e} \qquad i.e., \qquad \frac{\partial i_e}{\partial v_e} = \frac{1}{r_e} \qquad \ldots(2)$$

From eqs. (1) and (2), we get

$$g_m = \alpha / r_e \qquad \ldots(3)$$

We know that

$$i \cong_e a_{11} \text{ exp. } (v_e/v_T) \qquad \text{Ebers-Moll Equation}$$

where v_T is the volt-equivalent of temperature and equals kT/q, k, being the Boltzmann constant in joules/K.

At cut off,

$$r_e = \frac{\partial v_e}{\partial i_e} \cong \frac{v_T}{a_{11} \text{ exp. } (v_e/v_T)} = \frac{v_T}{i_e}$$

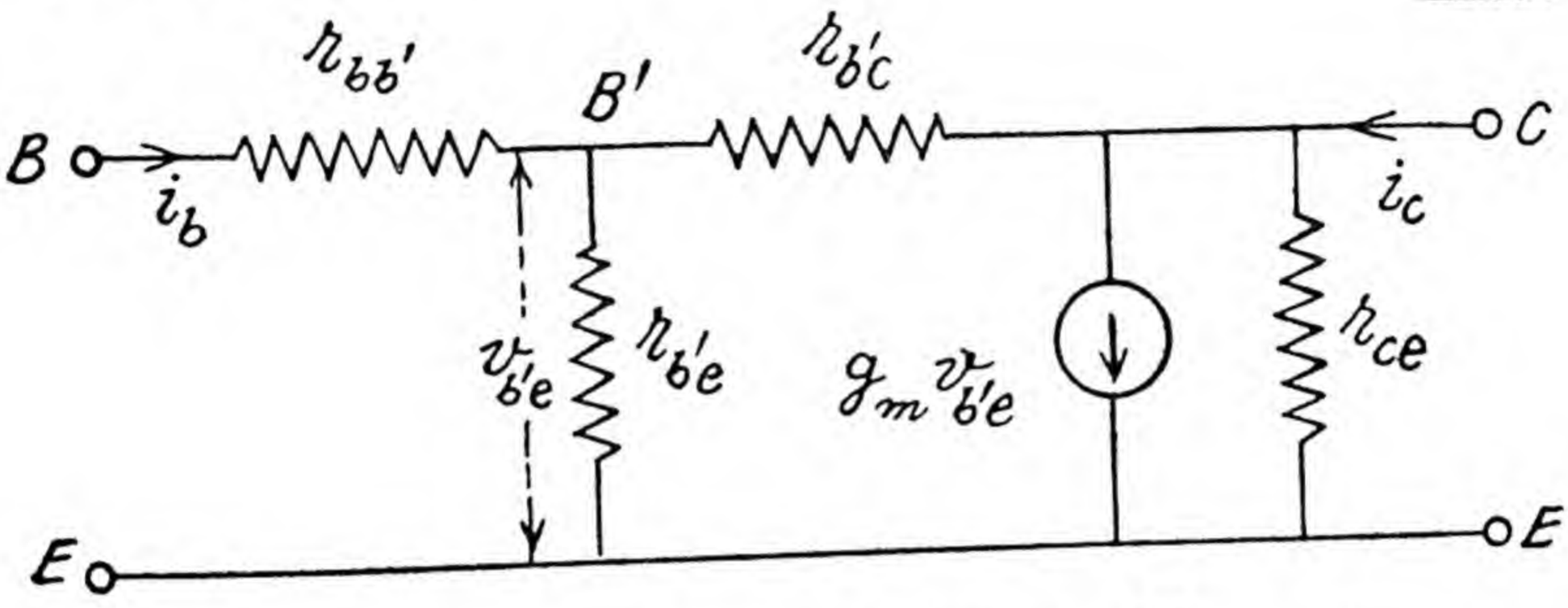
$$\therefore \qquad g_m = \frac{\alpha i_e}{v_T} = \frac{i_{C_o} - i_c}{v_T} \qquad \ldots(4)$$

For a *PNP* transistor i_c is negative and since $|i_{Co}| \ll |i_c|$, we have

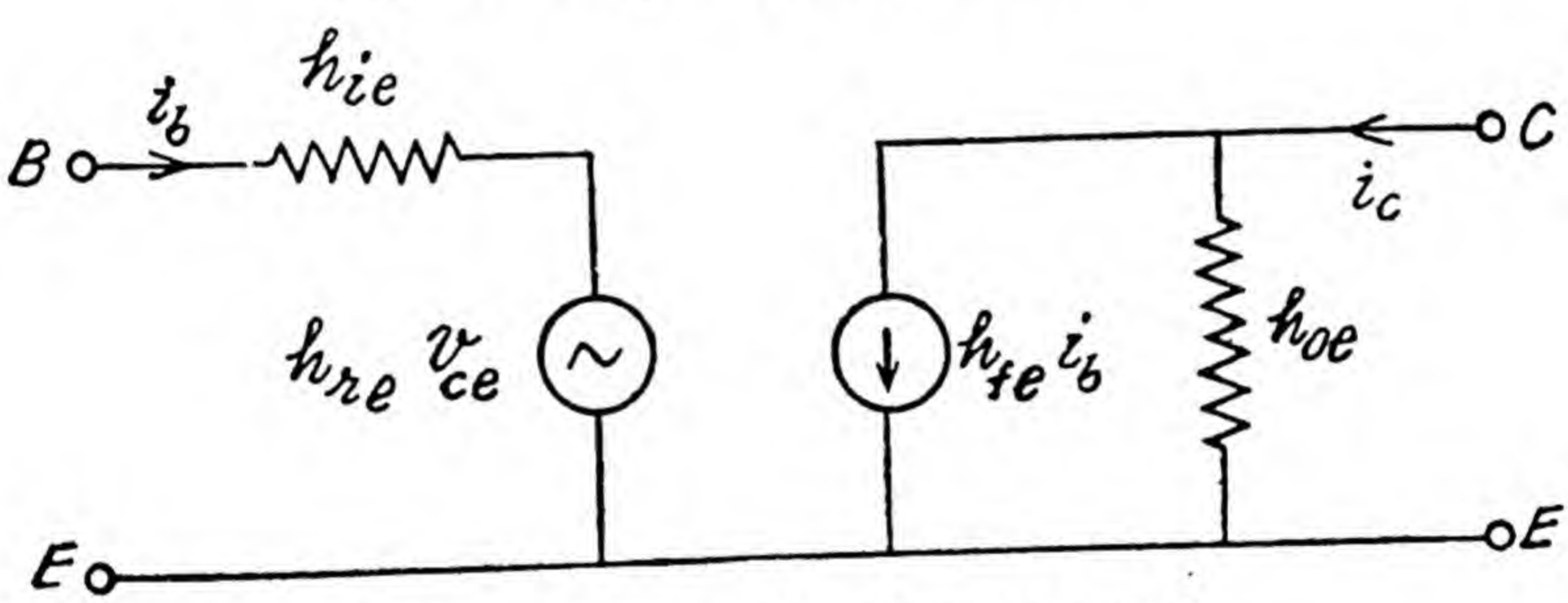
$$g_m \approx \frac{|\, i_o \,|}{v_T} \approx \frac{|\, i_c \,|}{T/11600} \quad (\because\ v_T = T/11600) \qquad \ldots(A)$$

For $i_c = 1\ mA$, $g_m \approx 40\ m\Omega$.

(2) Resistance $r_{b'c}$—Fig. (29 *a*) shows the hybrid π model at low frequencies where all capacitances are neglected and fig. (29 *b*) gives the equivalent circuit of same transistor at low frequencies using *h* parameters.



(a) Low frequency hybrid π model.



(b) Low frequency *h* parameter model.

Fig. 29. Low frequency hybrid π and *h* parameter models.

The reverse voltage gain is given by

$$h_{re} = \frac{v_{b'e}}{v_{ce}} \text{ with } i_b = 0 \qquad \ldots(5)$$

or $$h_{re} = \frac{r_{b'e}}{r_{b'e} + r_{b'c}}$$

$$\therefore \quad h_{re}\, r_{b'c} = r_{b'e}\,(1 - h_{re}) = r_{b'e} \qquad (\because\ 1 \gg h_{re})$$

or $$r_{b'c} = \frac{r_{b'e}}{h_{re}} \qquad \ldots(B)$$

(3) Resistance $r_{b'e}$—In general $r_{b'c} \gg r_{b'c}$. Hence almost the complete base current i_b flows into $r_{b'e}$ so that $v_{b'e} \approx i_b . r_{b'e}$.

The short circuit collector current i_c is given by

$$i_c = g_m v_{b'e} \approx g_m . i_b . r_{b'e} \qquad \ldots(6)$$

But the short circuit current gain h_{fe} is defined as

$$h_{fe}=\left.\frac{i_c}{i_b}\right|_{v_{ce}}=\frac{g_m.i_b.r_{b'e}}{i_b}=g_m r_{b'e}$$

$$\therefore \quad r_{b'e}=h_{fe}/g_m \qquad \ldots(C)$$

(4) Resistance $r_{bb'}$—The parameter h_{ie} forms the input resistance of the transistor with output short circuited. When the output is short circuited [Fig. (29*a*)], $r_{b'c}$ comes in parallel with $r_{b'e}$ But $r_{b'c} \gg r_{b'e}$, hence $r_{b'c} \parallel r_{b'e} \approx r_{b'e}$. Now the input resistance is approximately $r_{bb'}+r_{b'e}$

$$\therefore \quad h_{ie}=r_{bb'}+r_{b'e}$$

or $$r_{bb'}=h_{ie}-r_{b'e} \qquad \ldots(D)$$

(5) Resistance r_{ce}—For this input shorted, *i.e.*, $i_b=0$, we have

$$i_c=\frac{v_{ce}}{r_{ce}}+\frac{v_{ce}}{r_{b'c}+r_{b'e}}+g_m v_{b'e} \qquad \ldots(7)$$

Again $$v_{b'e}=h_{re}.v_{ce} \qquad \ldots(8)$$

From eqs. (7) and (8)

$$h_{oe}=\frac{i_c}{v_{ce}}=\frac{1}{r_{ce}}+\frac{1}{r_{b'c}+r_{b'e}}+g_m h_{re}$$

or $$h_{oe}=\frac{1}{r_{ce}}+\frac{1}{r_{b'c}}+g_m\, h_{re} \qquad \because r_{b'c} \gg r_{b'e}$$

or $$h_{oe}=\frac{1}{r_{ce}}+\frac{1}{r_{b'e}}+g_m.\frac{r_{b'e}}{r_{b'c}} \qquad \text{[using eq. (B)]}$$

or $$h_{oe}=\frac{1}{r_{ce}}+\frac{1}{r_{b'c}}+\frac{h_{fe}}{r_{b'c}} \qquad \text{[using eq. (C)]}$$

$$\therefore \quad r_{ce}=[h_{oe}-\{(1+h_{fe})/r_{b'c}\}]^{-1} \qquad \ldots(E)$$

(6) Capacitance $C_{b'e}$. This is the diffusion capacitance across the emitter junction and is equal to $C_{b'e}$. This is given by

$$C_{b'e} \approx \frac{g_m}{2\pi f_T} \qquad \ldots(F)$$

where f_T is the frequency at which common emitter short circuit current gain drops to unity.

(7) Capacitance $C_{b'c}$. The collector junction capacitance $C_{b'c}$ is nothing but the common base output capacitance with the input open circuited *i.e.*, $i_e=0$ and is usually specified by the manufacturers as C_{ob}. Since the collector junction is reversed biased in the active region, then $C_{b'c}$ is the transition capacitance and, therefore varies as v_{cb}^{-n} where $n=\frac{1}{2}$ or $\frac{1}{3}$ for an abrupt or graded junction respectively. Hence $C_{b'c}=kv_{cb}^{-n}$ where k is a constant.

Different parameters in the hybrid π model at high frequencies :

Parameter	Expression
g_m	$\mid i_c \mid / v_T = \mid i_c \mid / (T/11600)$
$r_{b'e}$	h_{fe}/g_m
$r_{bb'}$	$h_{ie} - r_{b'e}$
$r_{b'c}$	$r_{b'e}/h_{re}$
r_{ce}	$\left[h_{re} - \frac{(1+h_{fe})}{r_{b'c}} \right]^{-1}$
$C_{b'e}$	$g_m/2\pi f_T$
$C_{b'c}$	$k v_{ce}^{-n}$

Example 1. *The following low frequency parameters are known for a given transistor at* $i_e = 10\ mA$, $v_{cb} = 10V$ *and at*

$$h_{ie} = 500\,\Omega, \quad h_{oe} = 10^{-5} \text{ mhos}$$
$$h_{fe} = 100, \quad h_{re} = 10^{-4}$$

At the same operating point, $f_T = 50MHz$ *and* $C_{ob} = 3PF$. *Compute the values of all the hybrid parameters.*

$$g_m = \frac{\mid i_e \mid}{0{\cdot}026} = \frac{10 \times 10^{-3}}{0{\cdot}026} = \textbf{385m. mhos.}$$

$$r_{b'e} = h_{fe}/g_m = 100/385 \times 10^{-3} = \textbf{260 ohm}$$

$$r_{bb'} = h_{ie} - r_{b'e} = 500 - 260 = \textbf{240 ohm .}$$

$$r_{b'e} = r_{b'e}/h_{re} = 260/10^{-4} = \textbf{2·6 M ohm}$$

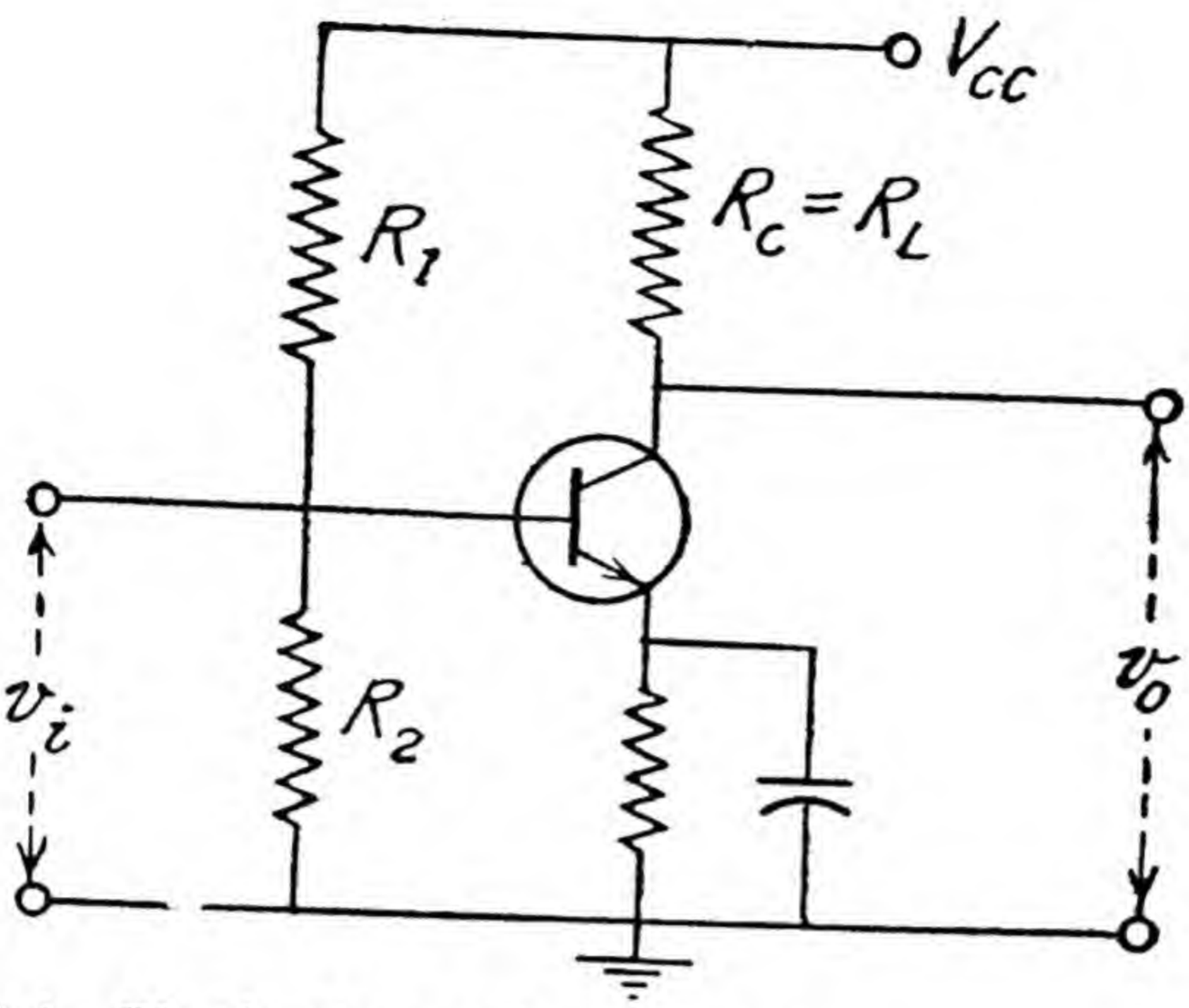
$$r_{ce} = \left[h_{oe} - \frac{(1+h_{fe})}{r_{b'e}} \right]^{-1} = \left[10^{-5} - \frac{(1+100)}{2{\cdot}6 \times 10^{-6}} \right]^{-1}$$
$$= \textbf{833K ohm}$$

$$c_{b'e} = g_m/2\pi f_T = \frac{385 \times 10^{-3}}{2 \times 3{\cdot}14 \times 50 \times 19^6}$$
$$= \textbf{1224 PF}$$

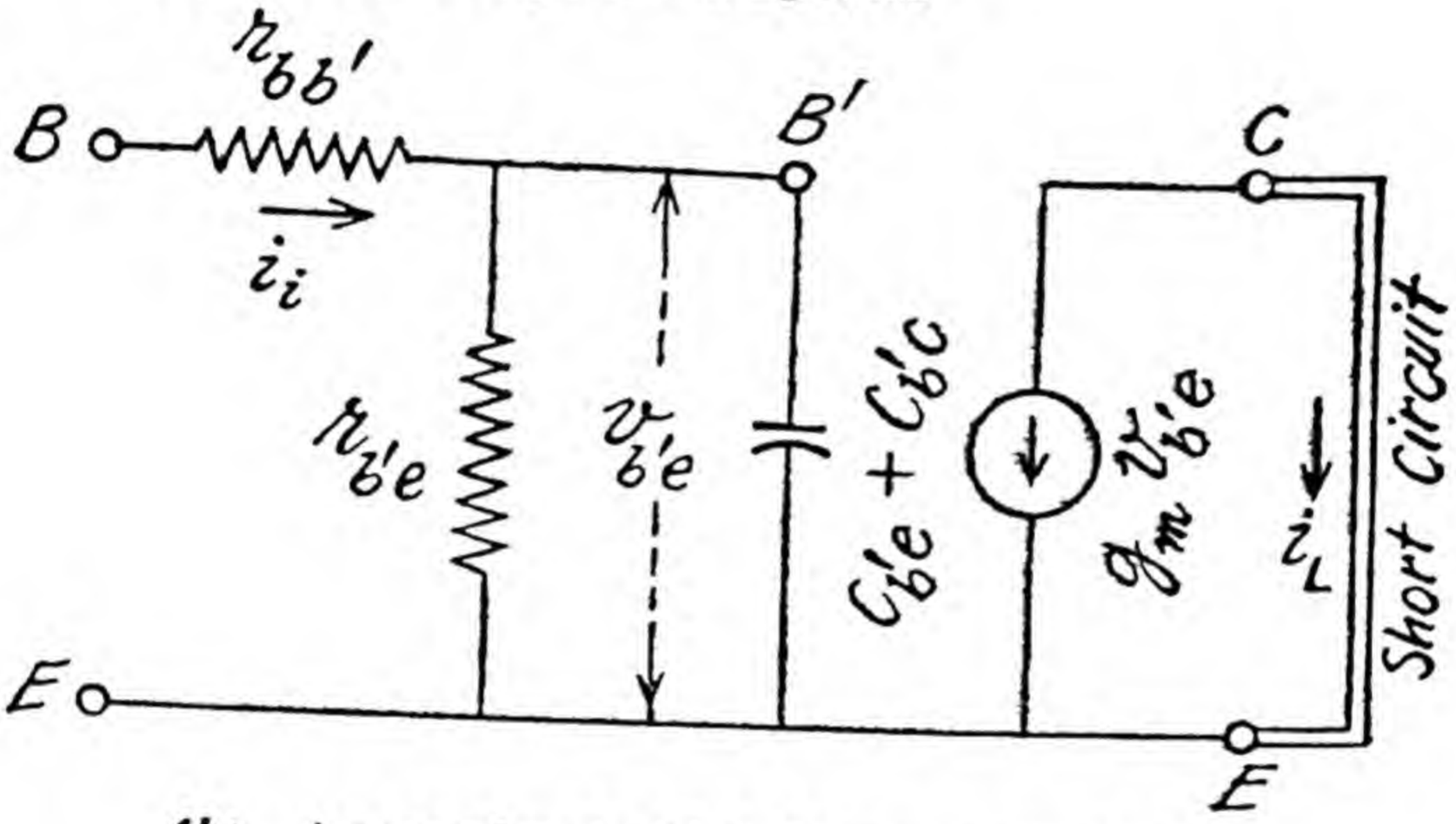
$c_{b'e}$ = Specified by manufacturer as $C_{ob} = 3$**PF.**

16·19. COMMON EMITTER SHORT CIRCUIT CURRENT GAIN USING HYBRID PI MODEL

Consider a single stage of *CE* transistor amplifier as shown in fig. (30*a*). The load R_L on this stage is the collector circuit resistor, so that $R_C = R_L$. In this section we assume that $R_L = 0$ *i.e.*, with the output short circuited. In next article we shall take the analysis when $R_L \neq 0$. The approximate equivalent circuit from which short circuited current gain is calculated is shown in fig. (30*b*).



(*a*) Single stage *CE* amplifier



(*b*) Approximate high frequency circuit.

Fig. 30. *CE* amplifier

The following assumptions are made :

1. In the equivalent circuit, the resistance $r_{b'e}$ is neglected as $r_{b'c} \gg r_{b'e}$.

2. The resistance r_{ce} disappears as it is in shunt with a short circuit.

3. $R_C = R_L = 0$.

From the equivalent circuit, $i_L = -g_m v_{b'e}$ ...(1)

The value of $v_{b'e}$ is given by

$$v_{b'e} = \frac{i_i}{(1/r_{b'e}) + j\omega (C_{b'e} + C_{b'c})} \quad \text{.. (2)}$$

From equation (1) and (2), we have

$$i_L = \frac{-g_m i_i}{(1/r_{b'e}) + j\omega (C_{b'e} + C_{b'c})} \quad \text{...(3)}$$

The short circuit current gain is given by

$$A_i = \frac{i_L}{i_i} = \frac{-g_m}{(1/r_{b'e}) + j\omega (C_{b'e} + C_{b'c})} \quad \text{...(4)}$$

We know that $r_{b'e}=h_{fe}/g_m$

$$\therefore \quad A_i=\frac{-g_m\, h_{fe}}{g_m+j\omega\,(C_{b'e}+C_{b'c})\,h_{fe}}$$

$$=\frac{-h_{fe}}{1+\{\,j\,2\pi f\,(C_{b'e}+C_{b'c})\,h_{fe}/g_m\}}$$

$$=\frac{-h_{fe}}{1+j\,(f/f_\beta)} \qquad \ldots(5)$$

where $$f_\beta=\frac{g_m}{h_{fe}.2\pi\,(C_{b'e}+C_{b\,c})}$$

Here we consider the following two cases :

(i) when $f=0$, $A_i=-h_{fe}$, which is in agreement with the definition of h_{fe} at the low frequency short circuit *CE* current gain.

(ii) when $f=f_\beta$, $A_i=-h_{fe}/\sqrt{2}$ *i.e.*, f_β forms the frequency at which *CE* short circuit current gain falls $1/\sqrt{2}$ of its low frequency value *i.e.*, falls by $3dB$. The frequency range upto f_β forms the *bandwidth* of the amplifier.

Frequency paramerer $\mathbf{f_T}$ —The frequency f_T is defined as the frequency at which the magnitude of *CE* short circuit current gain falls to unity.

Putting $|A_i|=1$ in eq. (5), we get

$$h_{fe}=1+j\,(f_T/f_\beta) \qquad (\because\ f=f_T)$$

or $$h_{fe}=f_T/f_\beta \qquad \because\ h_{fe}\gg 1$$

$$\therefore \quad f_T=h_{fe}.f_\beta \qquad \ldots(6$$

Substituting the value of f_β in eq. (6), we get

$$f_T=h_{fe}\,\frac{g_m}{h_{fe}.2\pi\,(C_{b'e}+C_{b'c})}\approx\frac{g_m}{2\pi C_{b'e}} \quad (\because\ C_{b'e}\gg C_{b'c}$$

Substituting eq. (6) in eq. (5), we have

$$A_i=\frac{-h_{fe}}{1+j\,h_{fe}\,(f/f_T)} \qquad \ldots(7$$

f_T forms an important parameter at high frequencies. The valu of f_T depends on the operating conditions of the device. The versus i_C (collector current) curve is shown in fig. (31).

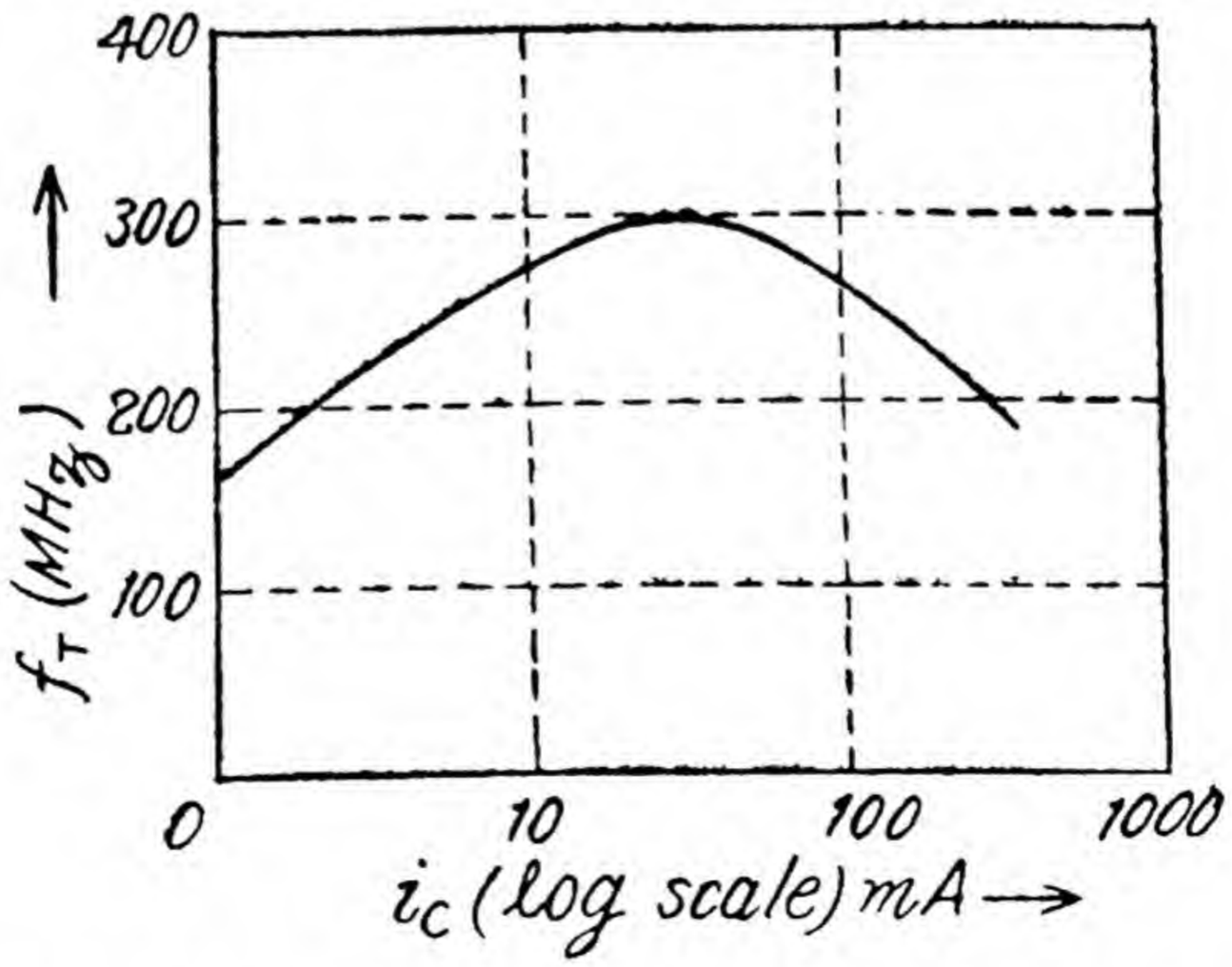


Fig. 31 (*a*) The f_T versus i_C curve at high frequencies.

In fig. (31*b*), A_i expressed in decibels (*i.e.*, 20 log $|A_i|$) is plotted against frequency on a logarithmic frequency scale.

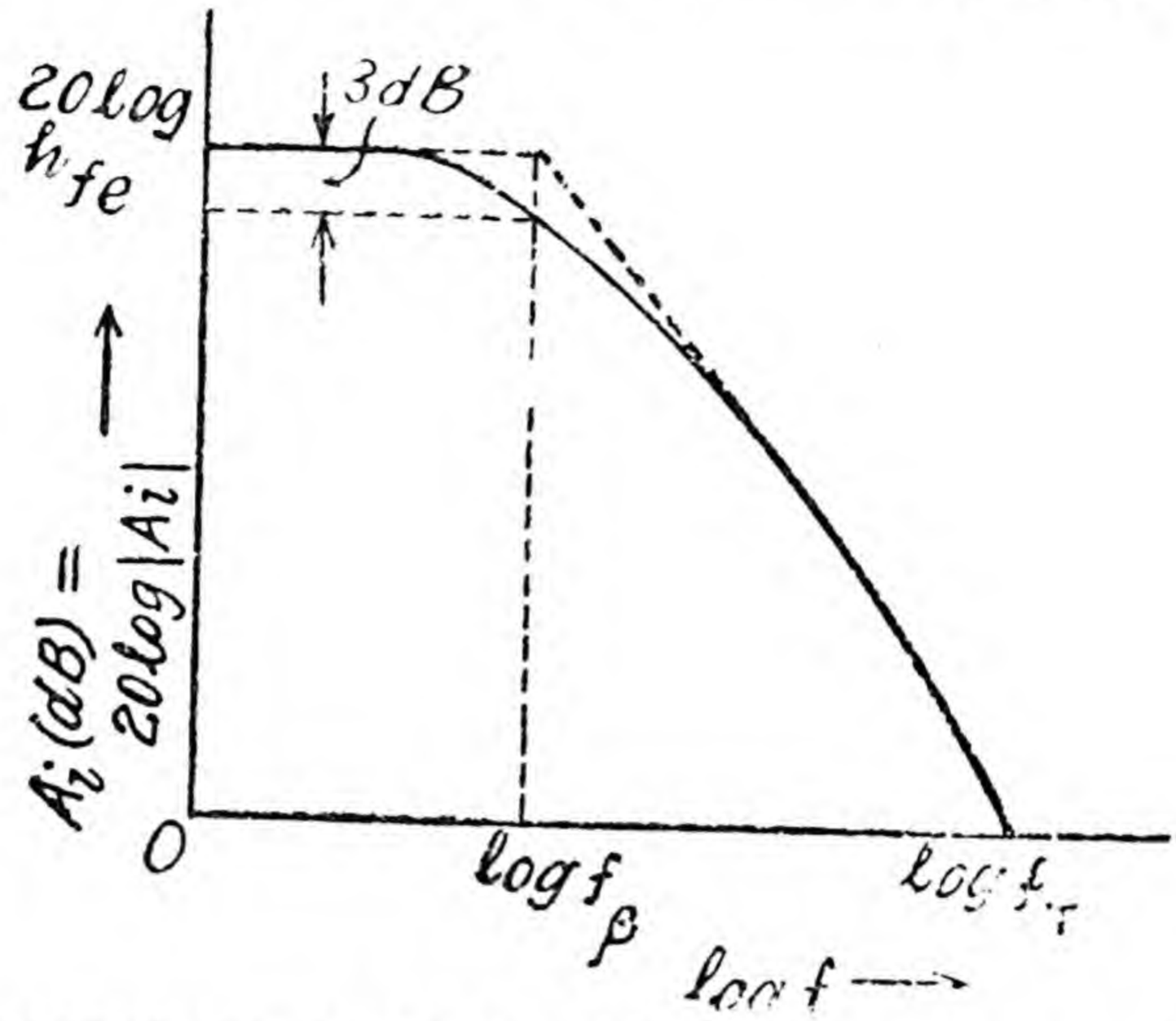


Fig. 31. (*b*) Variation of A_i (*dB*) with frequency on log scale.

We have the the following cases :

(i) when $f \ll f_\beta$, $A_i \approx -h_{fe}$ and A_i (*dB*) approaches asymptotically the horizontal line A_i (*dB*) $=20 \log h_{fe}$.

(ii) when $f \geqslant f_\beta$, $|A_i| \approx f_T/f$ so that

$$A_i \text{ (dB)} = 20 \log f_T - 20 \log f.$$

(iii) when $f = f_T$, A_i (*dB*) $=0$.

(iv) when $f \gg f_\beta$, the plot approaches as an asymptote to a straight line passing through the point (f_T, 0).

The intersection of the two asymptote occurs at a corner (frequency $f=f_\beta$) where A_i is down by 3*dB*. Hence f_β is called as the 3*dB* frequency.

6·20. HIGH FREQUENCY CE CURRENT GAIN WITH RESISTIVE LOAD.

Consider a single stage of *CE* transistor amplifier with resistive load R_L as shown in fig. (30*a*). When the transistor works into a resistive load, the equivalent circuit is shown in fig. (32).

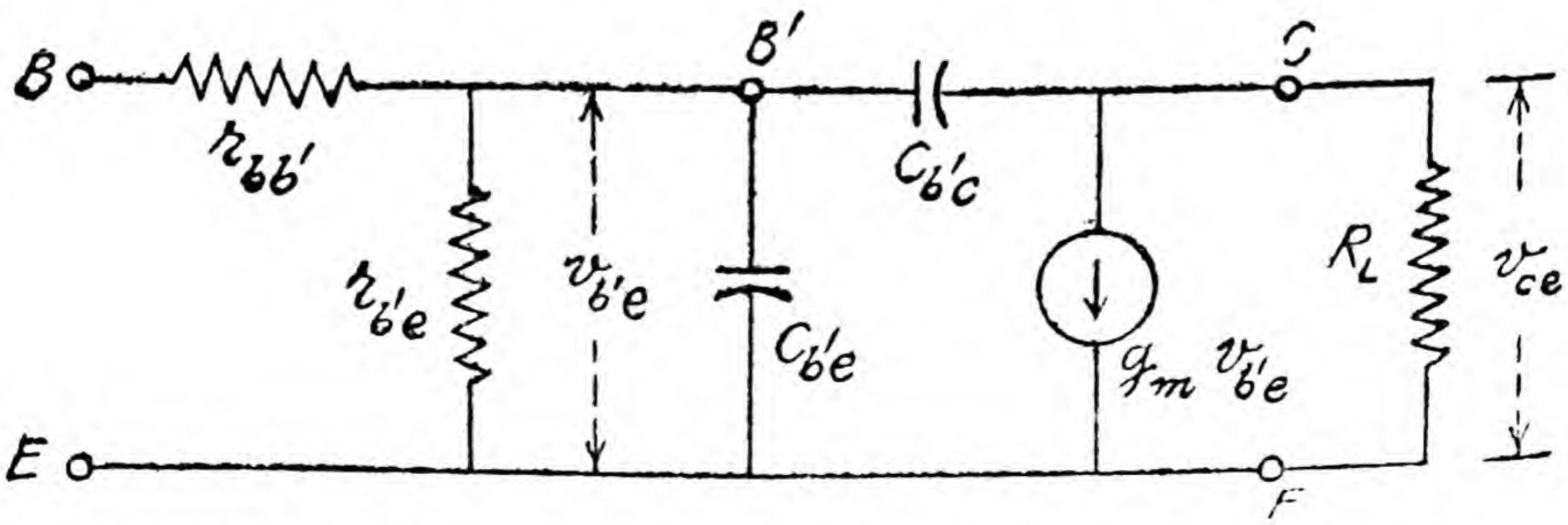


Fig. 32. Equivalent circuit of *CE* amplifier with resistive load.

The capacitance $C_{b'c}$ can be considered at the input side by applying the Miller's theorem. Miller's theorem states that an impedance Z connected between input and output can be replaced by an impedance $Z/(1-A)$ at the input side and an impedance $ZA/(A-1)$ at the output side [see fig. (33)], where A is the voltage gain. Refering to fig. (32), the voltage gain $A=v_{ce}/v_{b'e}=-g_m R_L$. At the input, the impedance is smaller by a factor $(1-A)$ and hence the input capacitance is larger by a factor $(1-A)$. So the capacitance $C_{b'c}$ can be replaced by $C_{b'c}\{1-(-g_m R_L)\}$ *i.e.*, $C_{b'c}(1+g_m R_L)$ at input side and a capacitance $C_{b'c}(+g_m R_L)(1+g_m R_L)$ at the output side. When $A \gg 1$, the value of capacitance at output is nearly $C_{b'c}$. This small value of capacitance in parallel with a low resistance R_L has a negligible time constant with the input time constant and hence can be neglected. Now the equivalent circuit of fig. (32) can be expressed as shown in fig.

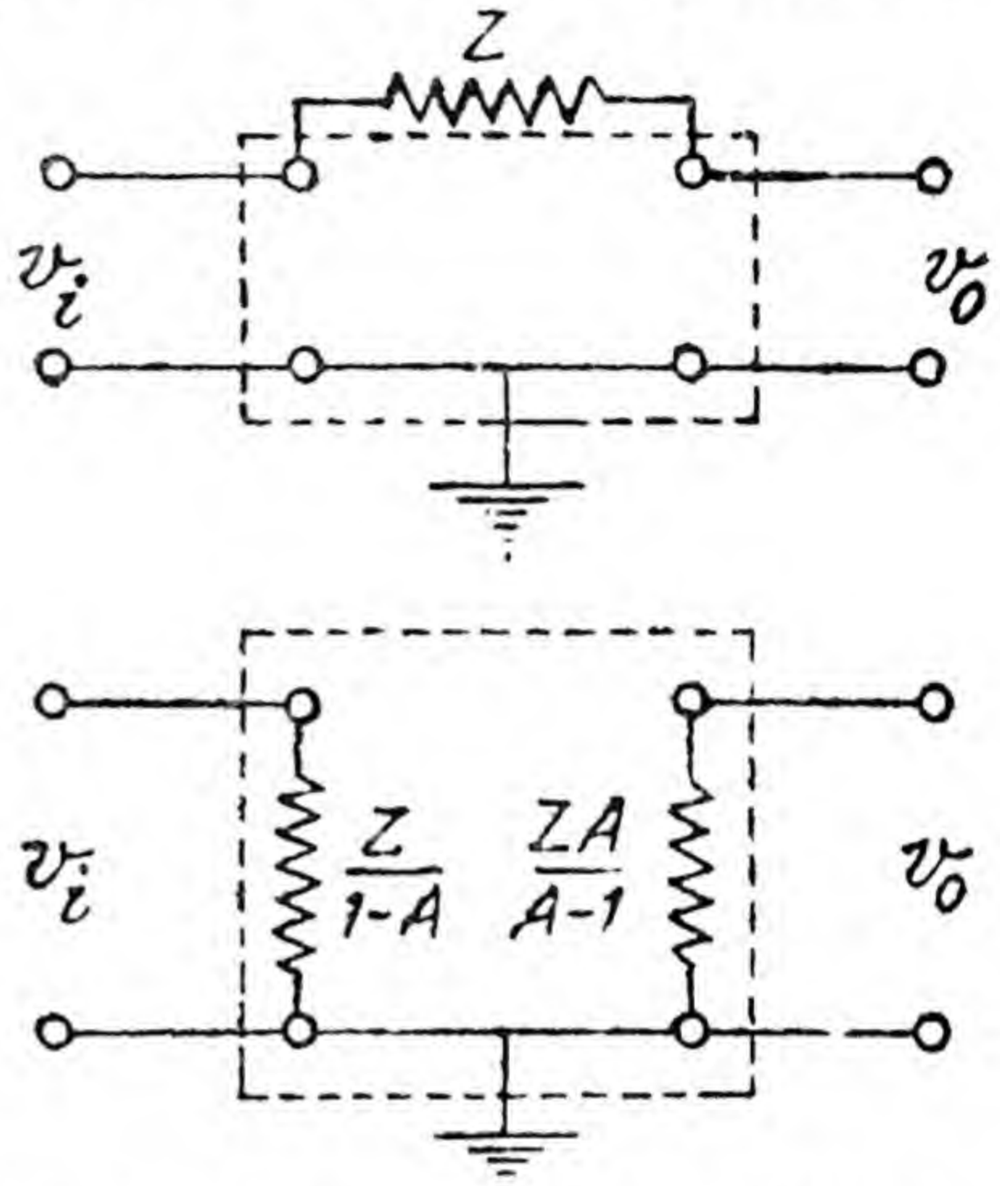


Fig. 33. Showiug Miller's theorem, $A=v_0/v_i$.

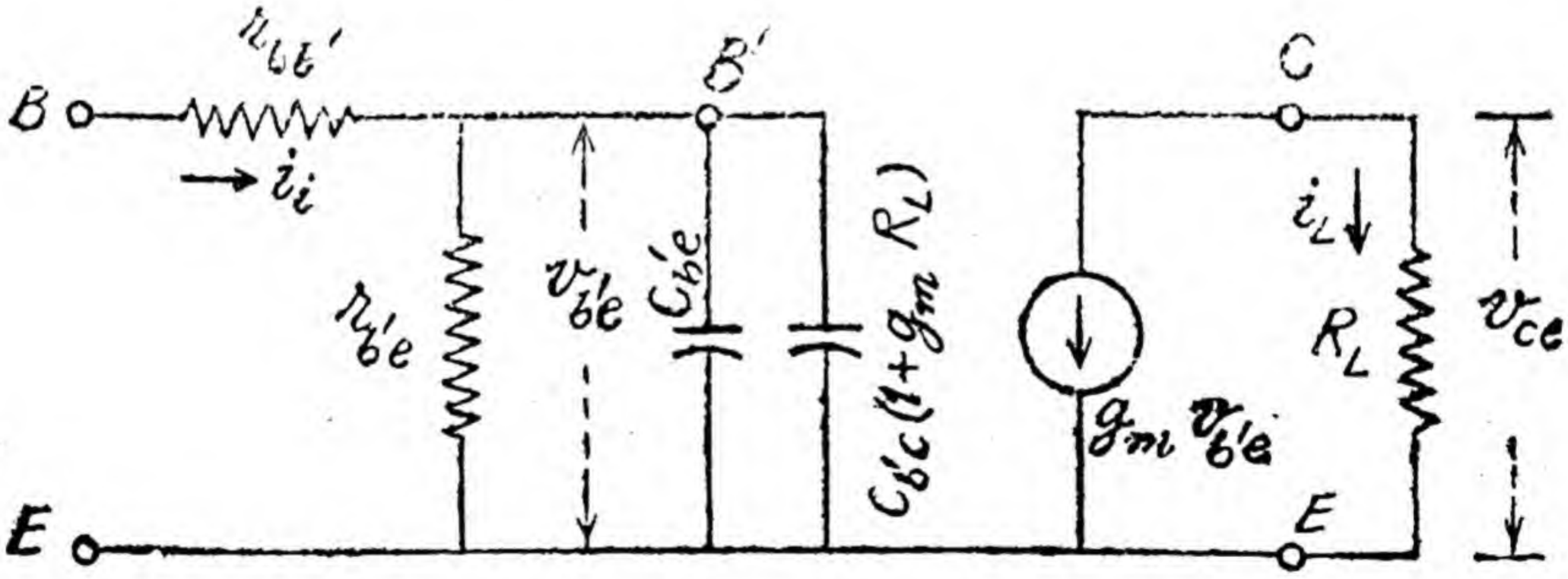


Fig. 33. Approximate model showing an additional capacitance $C_{b'c}(1+g_m R_L)$ at input due to Miller effect.

The capacitance C at the input side is

$$C=C_{b'e}+C_{b'c}(1+g_m R_L) \qquad \ldots(1)$$

This represents a considerable increase in capacitance at input side.

The value of load current i_L is given by

$$i_L=-g_m v_{b'e} \qquad \ldots(2)$$

Again $$v_{b'e}=\frac{i_i}{(1/r_{b'e})+j\omega[C_{b'e}+C_{b'c}(1+g_m R_L)]}$$

or $$v_{b'o}=\frac{i_i}{(1/r_{b'e})+j\omega C} \qquad ...(3)$$

From eqs. (2) and (3), we have

$$i_L=\frac{-g_m\ i_i}{(1/r_{b'e})+j\omega C} \qquad ...(4)$$

The current gain is given by

$$A_i=\frac{i_L}{i_i}=\frac{-g_m}{(1/r_{b'e})+j\omega C}$$

We know that $r_{b'e}=h_{fe}/g_m$

$$\therefore \quad A_i=\frac{-g_m\ h_{fe}}{g_m+j\omega C.h_{fe}}=\frac{-h_{fe}}{1+j\omega C.(h_{fe}/g_m)}$$

$$=\frac{-h_{fe}}{1+j\ 2\pi fC\ (h_{fe}/g_m)}=\frac{-h_{fe}}{1+j\ (f/f_H)} \qquad ...(5)$$

where $$f_H=\frac{1}{2\pi C}\cdot\frac{g_m}{h_{fe}}=\frac{1}{2\pi C\ r_{b'e}}$$

At $f=0$, $A_i=-h_{fe}$ (same as output short circuited)
At $f=f_H$, $A_i=-h_{fe}/\sqrt{2}$
Thus f_H forms the $3dB$ frequency.

16·21. EEFECT OF SOURCE RESISTANCE ON FREQUENCY RESPONSE OF CE AMPLIFIER AT HIGH FREQUENCIES :

In previous article, we have assumed that the current source is ideal *i.e.* the source resistance R_S to be infinity. In this article we consider that the source has a resistance R_S. The main aim here is to consider the effect of source resistance on frequency response of CE amplifier at high frequencies *i.e.*, how the $3dB$ frequency is modified due to the presence of R_S. The equivalent circuit is shown in fig. (34).

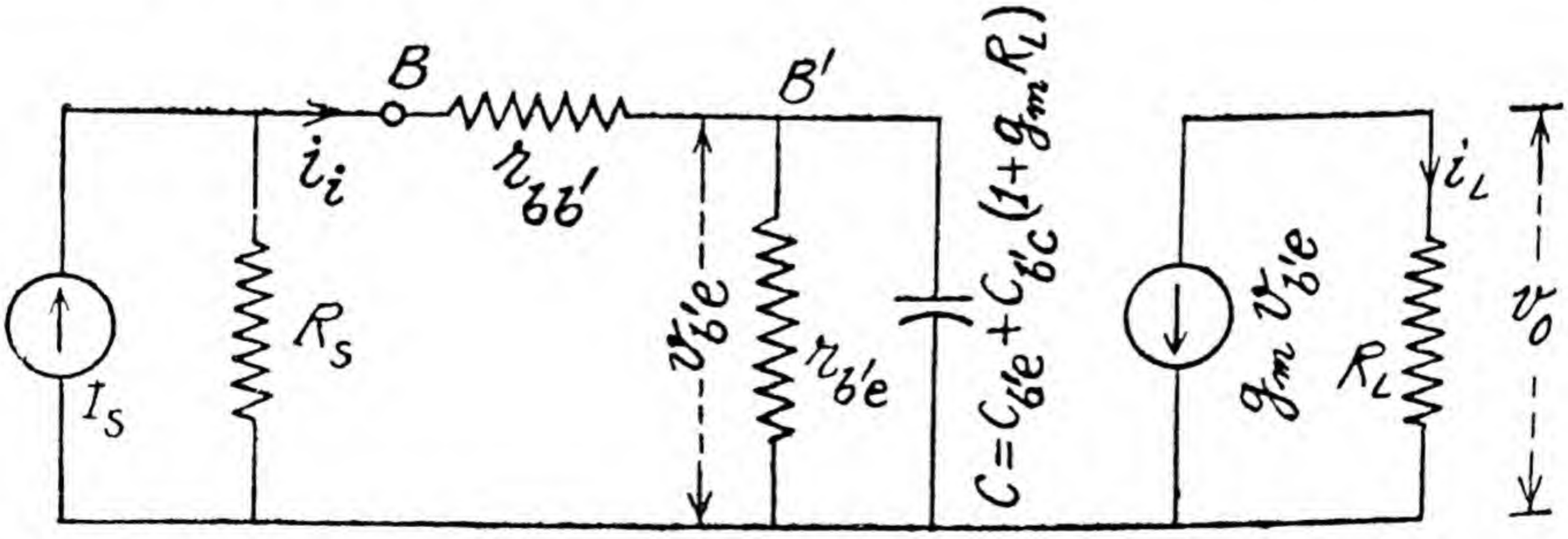


Fig. 34. Approximate equivalent circuit of CE amplifier using non-ideal current gain.

In this case the analysis remains the same as in previous article but in place of $r_{b'o}$ a resistance R is used which is a parallel combination of $(R_S+r_{bb'})$ and $r_{b'e}$, *i.e.*,

$$R=\frac{(R_S+r_{bb'}).r_{b'e}}{R_S+r_{bb'}+r_{b'e}}=\frac{(R_S+r_{bb'}).r_{b'e}}{R_S+h_{ie}} \quad ...(1)$$

The value of C is given by

$$C=C_{b'e}+C_{b'c}\,(1+g_mR_L) \quad ...(2)$$

The $3dB$ frequencies f_2 is now given by

$$f_2=\frac{1}{2\pi RC} \quad ...(3)$$

The following expressions can be derived :

(i) The current gain taken into account the source resistance is given by

$$(A_i)s=\frac{-h_{fe}\,R_S}{R_L+h_{ie}} \quad ...(4)$$

(ii) The voltage gain is given by

$$(A_v)s=(A_i)s\,\frac{R_L}{R_S}=\frac{-h_{fe}}{R_S+h_{ie}}\frac{R_L}{} \quad ...(5)$$

(iii) Gain bandwidth product is given by

$$|\,(A_v)s.f_2\,|=\frac{g_mR_L}{2\pi C\,(R_S+r_{bb'})}=\frac{f_T\,R_L}{(1+2\pi f_TC_{b'c}R_L)(R_S+r_{bb'})} \quad ...(6)$$

and current-gain bandwidth product is

$$|\,(A_i)s\,f_2\,|=\frac{f_TR_S}{(1+2\pi f_TC_{b'c}R_L)\,(R_S+r_{bb'})} \quad ...(7)$$

Example. *Calculate the voltage gain upper cut-off frequency of a transistor amplifier in CE configuration operating at high frequencies which has the following parameters.*

$f_T=6MHz$, $h_{fe}=50$, $g_m=0{\cdot}04$ mho

$r_{bb'}=100\Omega$, $R_S=500\Omega$, $C_{b'c}=10PF$

$R_L=1000\Omega$.

$r_{b'e}=h_{fe}/g_m=50/0{\cdot}04=1250\Omega$

$C_{b'e}=g_m/2\pi f_T=0{\cdot}04/2\pi\times6\times10^6=1060\ PF$

$\therefore\quad C=C_{b'e}+C_{b'c}\,(1+g_mR_L)$

$=1060+10\,(1+0{\cdot}04\times1000)$

$=1460PF.$

Now $h_{ie}=r_{bb'}+r_{b'e}=100+1250=1350\Omega$

$$R=\frac{600\times1250}{600+1250} \qquad (\because\quad R_S+r_{bb'}=500+100=600)$$

$=405{\cdot}4\Omega.$

$$\therefore\quad f_2=\frac{1}{2\pi RC}=\frac{1}{2\pi\times405{\cdot}4\times1460\times10^{-12}}$$

$=\mathbf{0{\cdot}27\ MHz.}$

16·22. R-C COUPLED AMPLIFIER :

This type of coupling is very commonly used for cascading two audio frequency amplifiers. A two stage *RC* coupled *CE* amplifier is shown in figure 35. Coupling condenser has been denoted by C_{c1} through which the output of the first stage is fed to the input of the second stage. The d.c. component of the output voltage of the first stage is blocked by the coupling condenser in reaching to the input of the second stage.

Value of Coupling Condenser : The capacitance C_{c1} is determined by the low frequency limit of the passband, the response being 3 *db* down at the frequency for which the capacitive reactance equals the resistance effectively in parallel with C_{c1}. If in this amplifier the resistance is say, 80 ohms from the input of the following transistor, 3·3*K*-ohms from the output circuit of the preceding transistor, and if the low frequency limit is 10 c/s, we have an approximate idea of the value of coupling condenser by putting (for exact value see example given in the end).

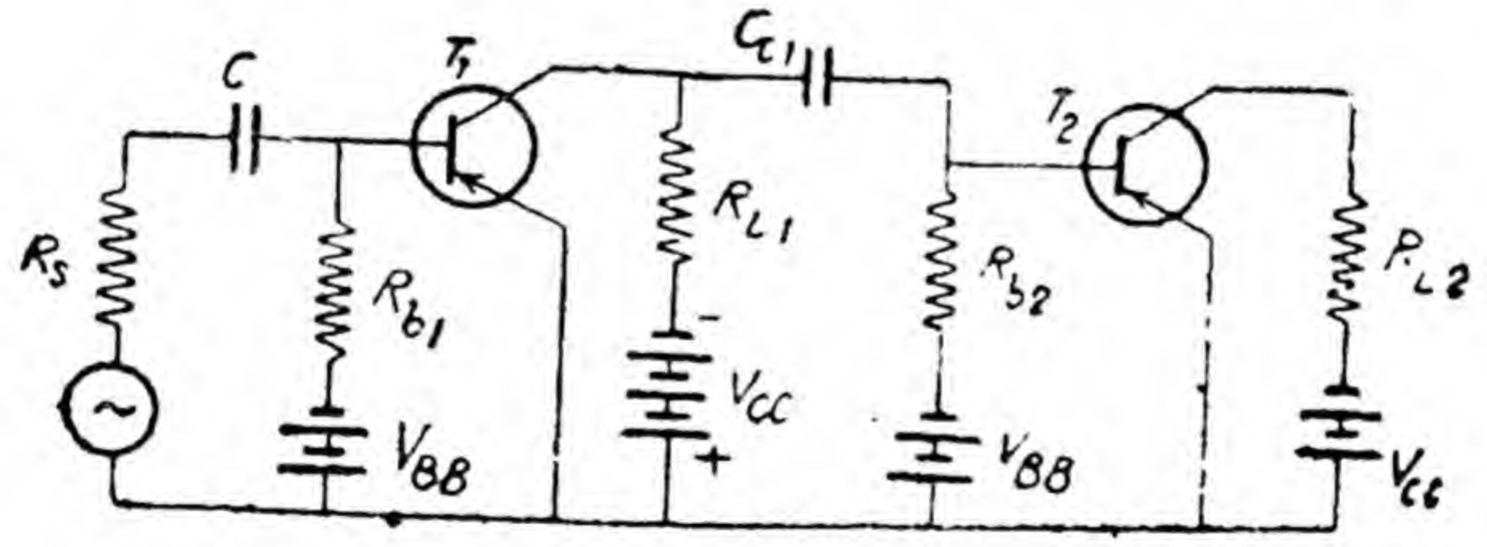
$$\frac{1}{2\pi f C_{c1}} = r$$

or

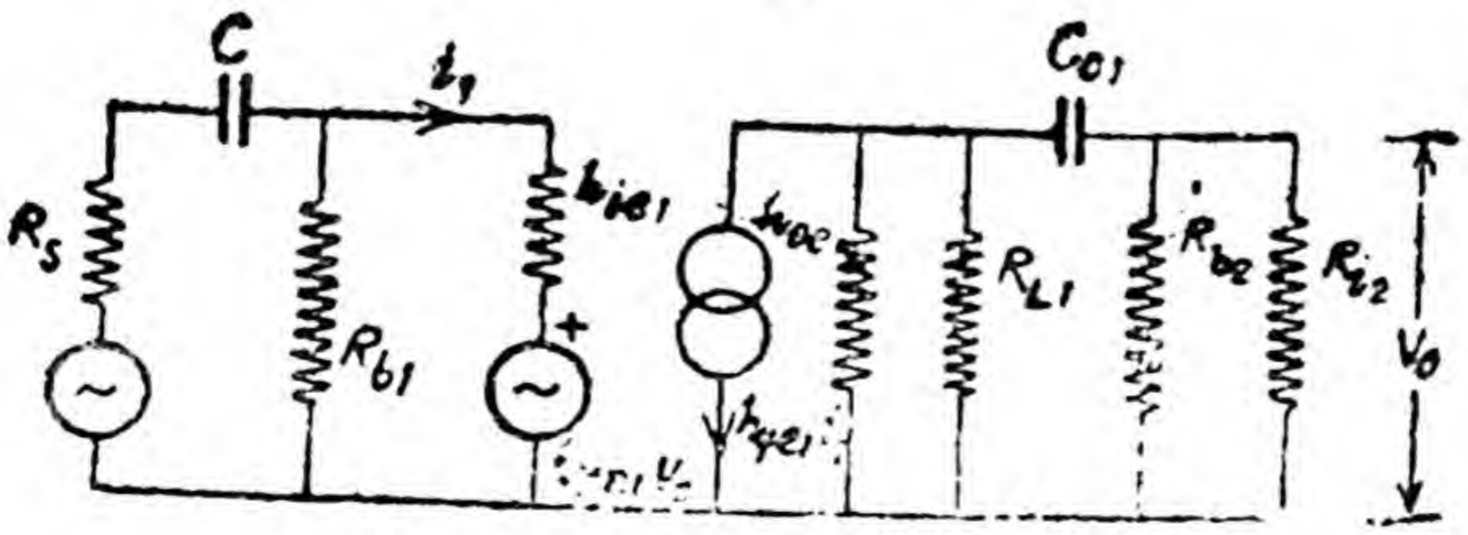
$$C_{c1} = \frac{1}{2\pi f r} = \frac{1}{6{\cdot}284 \times 10 \times 4100} \text{ Farads}$$

$$= 4{\cdot}0\mu F \text{ approximately.}$$

Note that this value depends primarily on the value of R_{L1} (3·3 $K\Omega$) of the first stage. This differs considerably from vacuum



(*a*)



(*b*)

Fig. 35. R-C coupled amplifier with equivalent circuit.

tube circuit practice, where the required capacitance depends primarily upon the input resistance of the following stage. The input impedance of the transistor circuit is far lower than that of the vacuum tube amplifier and as such, a considerable amount of current is to be supplied to the input of a transistor amplifier.

The frequency characteristic with its bandwith is shown in fig. 36. It is convenient to analyse the performance of the amplifier in different frequency ranges.

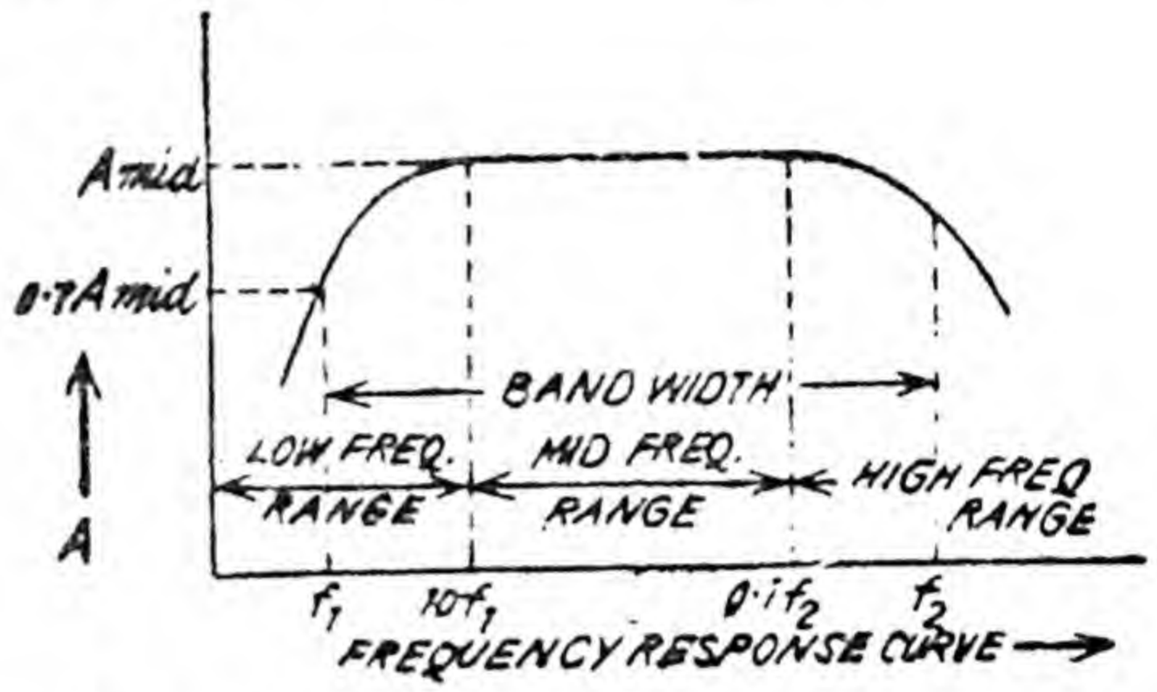
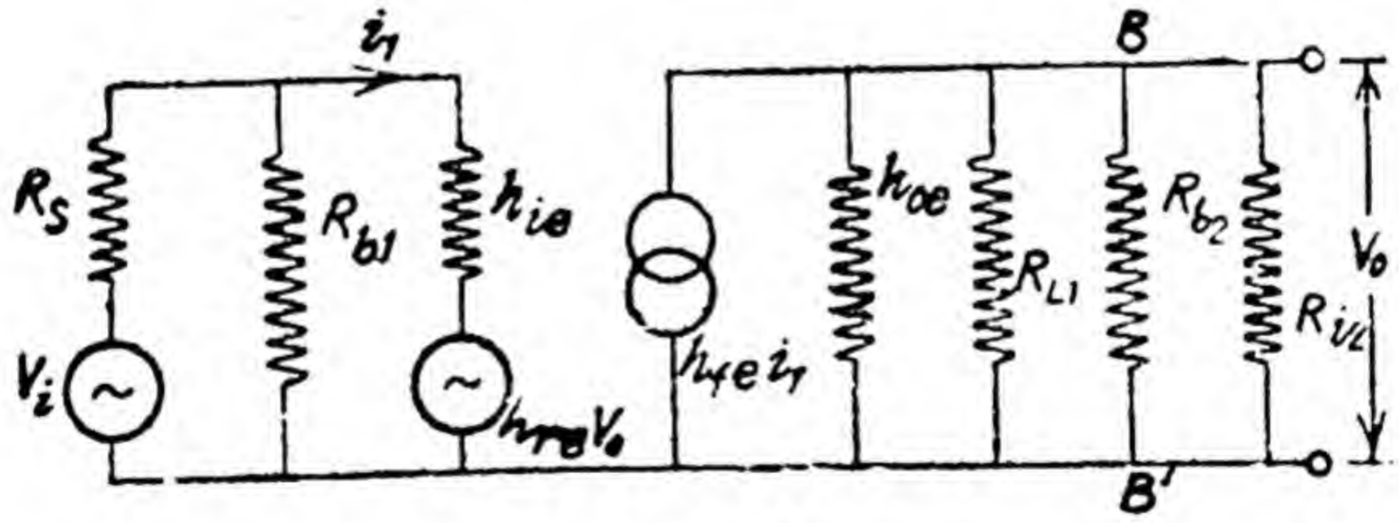


Fig. 36. Frequency response curve.
f_1 = lower cut-off frequency.
f_2 = upper cut-off frequency.

(i) Mid Frequency Range: At mid frequencies, the impedance offered by coupling condenser C_c is so small as to be an effective short circuit, while the impedance offered by the shunting capacity is comparatively so large as to be an effective open circuit. Thus both can be eliminated while drawing equivalent circuit of *RC* coupled amplifier at mid frequencies as shown in fig. 37 (*a*). Simplified equivalent circuit is shown in fig. 37 (*b*), which is obtained by applying, Thevenin's theorem across *bB* of fig. 37 (*a*). G_{L1} is the conductance corresponding to R_{L1}. R_{i2} is the input impedance of the second stage. Suppose R'_L is the effective load resistance, then

$$\frac{1}{R'_L} = \frac{1}{R_{L1}} + \frac{1}{R_{b2}} + \frac{1}{R_{i2}}.$$



Ftg. 37 (a) Equivalent circuit at mid frequencies.

From equations (14) and (17) of article 16·2, the input impedance of the first stage R_{i1} and the output impedance of the first stage R_{o1} may be written as

$$R_{i1} = \frac{R'_L \Delta + h_{ie1}}{h_{oe1} R'_L + 1}$$

and

$$R_{o1} = \frac{R_S + h_{ie1}}{R_S h_{oe1} + \Delta},$$

where $\Delta = h_{ie1}\ h_{oe1} - h_{re1}\ h_{fe1}$; and subscripts 1 and 2 indicate that the parameters belong to the transistors 1 and 2 respectively.

Refer to the equivalent circuit drawn in fig. 37 (*b*). Assuming that the loop current is i_0, we may write loop equation of the circuit of fig. 37 (*b*)

$$i_0 = \frac{h_{fe1}\ i_1}{h_{oe1} + G_{L1}}$$

$$= \left(R_{L0} + \frac{R_{b2} R_{i2}}{R_{b2} + R_{i2}} \right) i_0.$$

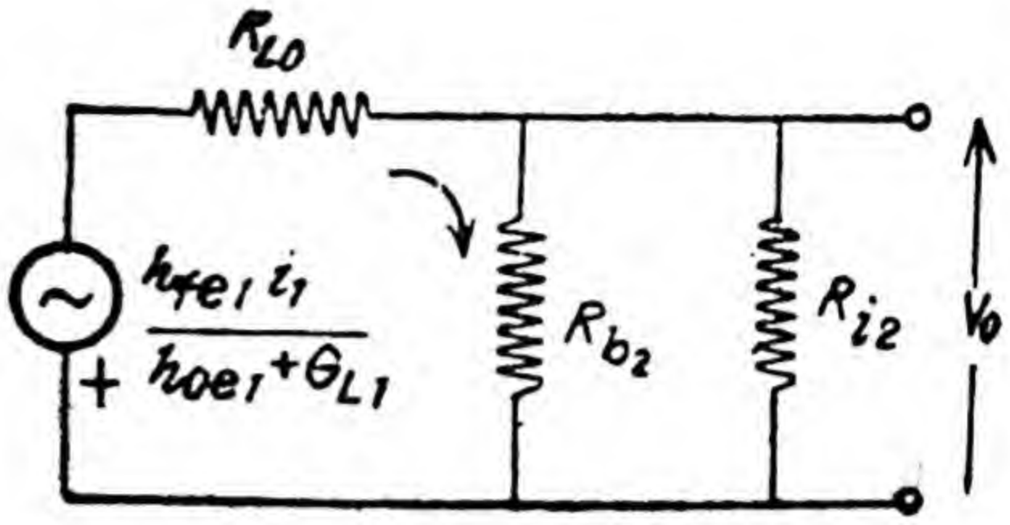


Fig. 37 (*b*) Simplified equivalent circuit drawn from fig. 37 (a) with open circuited voltage across $BB' = \frac{h_{fe1} i_1}{h_{oe1} + G_{L1}}$ and open circuited impedance across BB'

$$= R_{L0} = \frac{R_{L1} R_{01}}{R_{L1} + R_{01}}.$$

Therefore the current gain at mid frequencies will be

$$A_{mid} = \frac{i_0}{i_1} = \frac{-h_{fe1}}{h_{oe1} + G_{L1}} \cdot \frac{1}{R_{L0} + R'_{i2}}, \qquad \ldots(1)$$

where $$R'_{i2} = \frac{R_{L2} R_{i2}}{R_{b2} + R_{i2}} \simeq R_{i2},$$

when $$R_{b2} > > R_{i2}.$$

(2) Low Frequency Range : At this frequency range the impedance offered by coupling condenser is comparable to the load resitance so that coupling capacity largely affects the current amplication. Therefore we shall have to include it in the equivalent circuit. On the other hand, the impedance offered by the shunting capacity is so large as to be an affective open circuit and, therefore, shunting capacity C_{sh} can be neglected. With these considerations, we can draw the equivalent circuit as shown in fig. 38 (*a*).

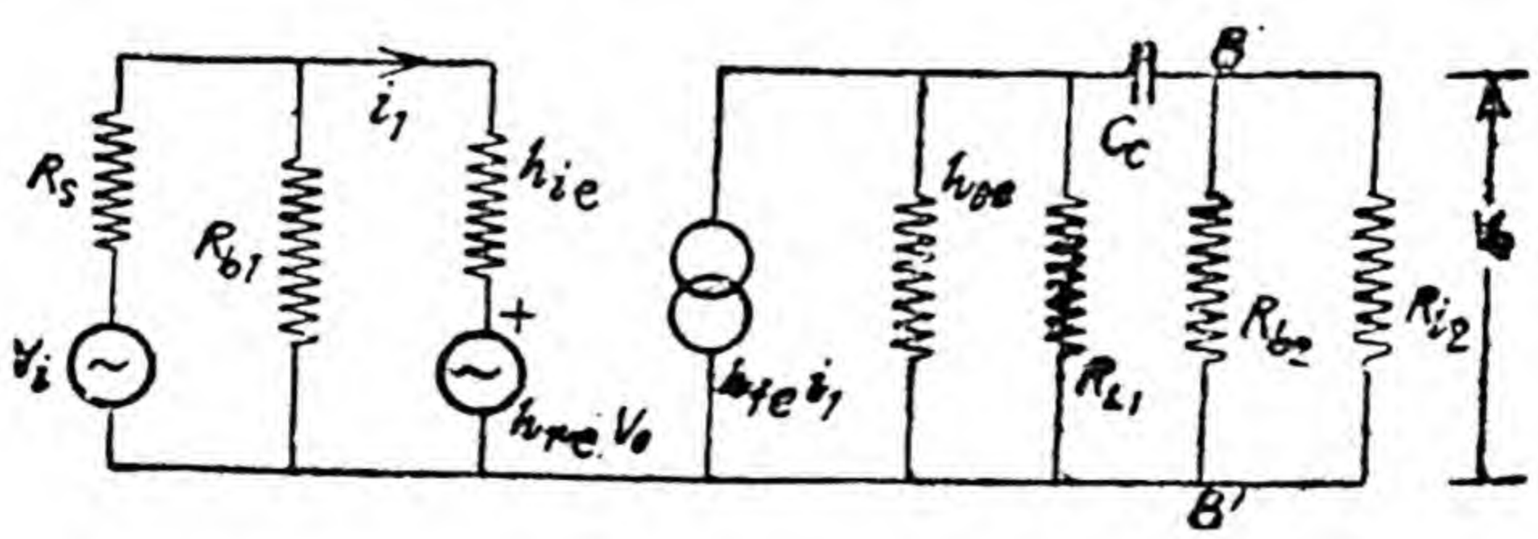


Fig. 38 (*a*) Equivalent circuit of RC coupled amplifier at low frequencies.

We now apply Thevenin's theorem to the equivalent circuit of fig. 38 (*a*) across BB' and draw the simplified equivalent circuit of this frequency range as shown in fig. 38 (*b*).

Assuming that i_0 is the loop current, then loop equation of circuit of fig. 38 (*b*) is

$$i_0=\frac{h_{fe1}\, i_1}{h_{oe1}+G_{L1}}$$

$$=i_1\left(R_{L0}+R'_{i2}+\frac{1}{j\omega C_o}\right)$$

Therefore the current gain at low frequencies is

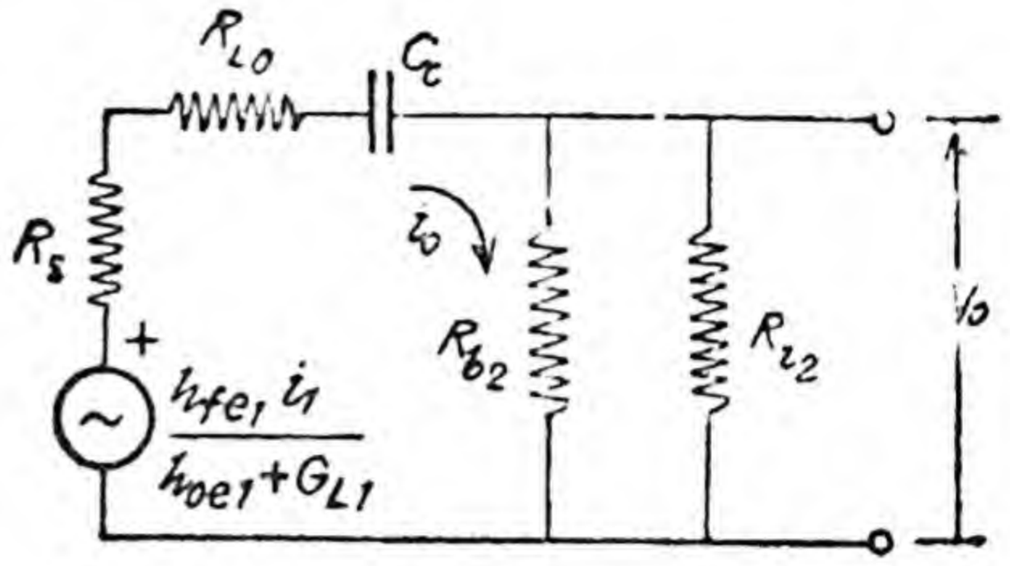


Fig. 38 (b) Simplified equivalent circuit of RC coupled amplifier at low frequencies.

$$A_{low}=\frac{i_0}{i_1}$$

$$=\frac{h_{fe1}}{h_{oe1}+G_{L1}}\cdot\frac{1}{R_{L0}+R'_{i2}+\frac{1}{j\omega C_o}} \qquad ...(2)$$

Dividing equation (2) by equation (1), we get

$$\frac{A_{low}}{A_{mid}}=\frac{R_{L0}+R'_{i2}}{R_{L0}+R'_{i2}+\frac{1}{j\omega C_o}}$$

$$=\frac{1}{1-\frac{j}{\omega C_o(R_L+R'_{i2})}} \qquad ...(3)$$

We know that lower cut off frequency f_1 refers to a frequency at which the gain of the amplifier is 70% of its mid frequency gain, and may be expressed as

$$f_1=\frac{1}{2\pi C_o\,(R_{L0}+R'_{i2})},$$

Putting the value in equation (3), we get

$$\frac{A_{low}}{A_{mid}}=\frac{1}{1-\frac{jf_1}{f}}=\frac{1}{[1+(f_1/f)^2]^{1/2}}\,\underline{|\tan^{-1}(f_1/f)}$$

$$=0{\cdot}707,$$

when $f=f_1$ and phase shift is equal to 42°. Thus at the lower cut-off frequency f_1,

$$A_{low}=70\%\ A_{mid}.$$

(3) High Frequency Range :** At this frequency range, the impedance offered by the coupling capacitor is negligible. In addition, the wiring and the shunting capacitances becomes effective and lowers the current and voltage gain of the amplifier as the frequency increases. If C_w and C_{oe} denote the wiring capacity and output capacity of the common emitter circuit then we write the

**The exact analysis at this frequency range is different on account of the variation of transistor parameters such as h_{fe} and h_{ie} with frequency and also due to the presence of shunting capacitances.

total shunt capacitance $C_{sh}=C_w=C_{ob}$ or shunt admittance as $Y_{sh}=j\omega C_{sh}$. The equivalent circuit is shown in fig. 24 (*a*) and a simplified equivalent circuit is shown in fig. 39 (*b*).

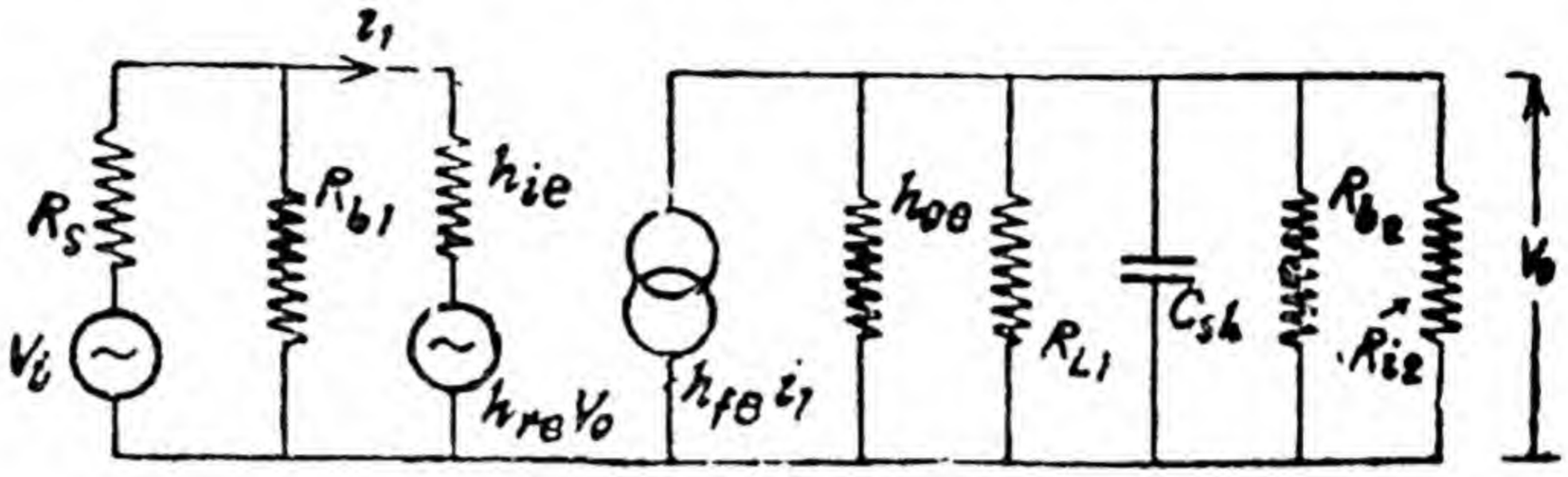


Fig. 39 (*a*) Equivalent circuit of *RC* coupled amplifier at high frequencies.

where
$$\frac{1}{R_{sh}}=\frac{1}{R_{L1}}+\frac{1}{R_{b2}}+\frac{1}{R_{i2}}$$
$$=G_{L1}+Y_{b2}+Y_{i2}.$$

The equivalent admittance Y_{L1} can then be expressed as

$$Y_{L1}=\frac{1}{R_{sh}}+\frac{1}{j\omega C_{sh}}.$$

Therefore current i_2 flowing across Y_{L1} will be

$$i_2=\frac{h_{fe1}\, i_1 \frac{1}{h_{oe1}}}{\frac{1}{Y_{L1}}+\frac{1}{h_{oe1}}}$$

or
$$i_2=\frac{h_{fe1}\, Y_{L1}\, i_1}{h_{oe1}+Y_L}$$

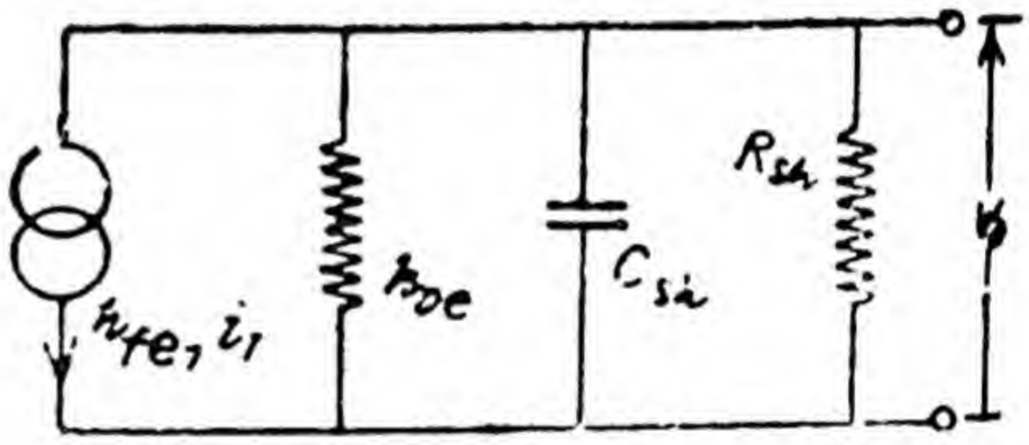


Fig. 39. (*b*) Simplified equivalent circuit of RC coupled amplifier at high frequencies.

The current i_0 flowing through R_{sh} will be

$$i_0=\frac{\frac{1}{Y_{sh}}\, i_2}{R_{sh}+\frac{1}{Y_{sh}}}$$

$$=\frac{i_2}{1+R_{sh}\, Y_{sh}}$$

$$=\frac{h_{fe1}\, Y_{L1}\, i_1}{(1+R_{sh}\, Y_{sh})\,(h_{oe1}+Y_{L1})}$$

Therefore current gain at high frequencies is

$$A_{hiah}=\frac{i_0}{i_1}=\frac{h_{fe}\, Y_{Li}}{(1+R_{sh}\, Y_{sh})\,(h_{oe1}+Y_{L1})} \qquad ...(4)$$

It may be mentioned there that h_{fe1} and Y_{L1} both are frequency dependent. Y_{L1} includes shunting capacitance and the input impedance of the following stage which also contains the frequency

dependent term h_{oe2} that increases with frequency. With the increase in frequency Y_{L1} increases so that according to eq. (4) current gain A_{high} also increases but due to the decrease in h_{fe1} with the increase in frequency, any increase in A_{high} is thereby counteracted, although the variation of h_{fe1} has very little effect on the response curve.

Example. In an RC coupled amplifier as shown in fig. 35(*a*), same type of transistors are used in both the stages and it is given that $h_{fe}=30$, $h_{ie}=1K\Omega$, $h_{oe}=10$ micromhos, $h_{re}=0{\cdot}5\times10^{-4}$, $R_{L1}=R_{L2}=6K$ ohms and $R_S=500$ ohms. If the lower cut off frequency of the amplifier is to be 10 *c/s*, find the values of the coupling capacity C_c.

We have shown that

$$C_c=\frac{1}{2\pi f\,(R_{Lo}+R'_{i2})}$$

we shall first find the values of R_{Lo} and R'_{i2}. We know that

$$R_{01}=\frac{h_{ie1}+R_S}{h_{oe1}\,R_S+h_{oe1}h_{ie1}-h_{fe1}h_{re1}}$$

$$=\frac{1000+500}{10\times10^{-6}\times500+10\times10^{-6}\times1000-0{\cdot}5\times10^{-4}\times30}$$

$$=\frac{1500}{14{\cdot}85\times10^{-3}}$$

$$\approx 101\ K \text{ ohms.}$$

$$R_{i2}=h_{ie2}-\frac{h_{ie2}h_{re2}\,R_{L2}}{1+h_{ce2}\,R_{L2}}$$

$$=1000-\frac{30\times0{\cdot}5\times10^{-4}\times6000}{1+10\times10^{-6}\times6000}$$

$$\approx 1\ K \text{ ohms.}$$

Therefore,

$$R_{Lo}=\frac{101\times6}{101+6}=\frac{606}{107}\ K \text{ ohms.}$$

$$=5{\cdot}6\ K \text{ ohms.}$$

$$R'_{i2}=\frac{30\times1}{30+1}\approx 1\ K \text{ ohms.}$$

Therefore

$$C_c=\frac{1}{2\times3{\cdot}14\times10\times(5{\cdot}6+1)\times10^3}$$

$$=2{\cdot}5 \text{ micro farads.}$$

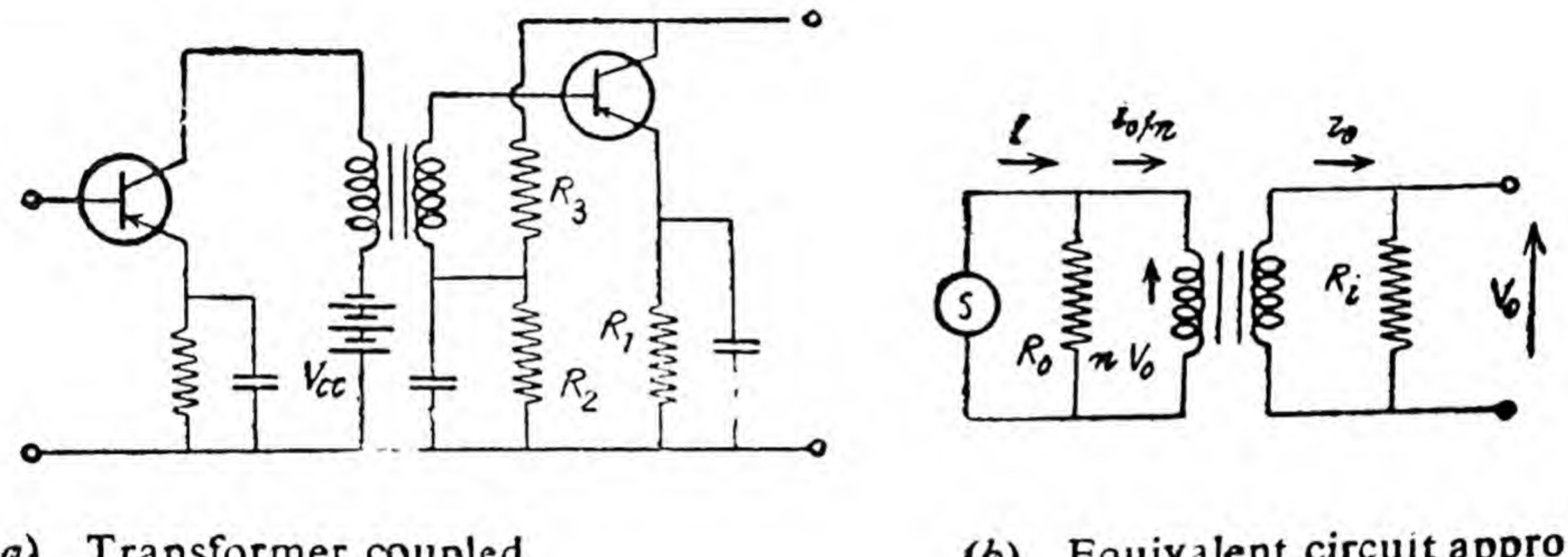
16·23. TRANSFORMER COUPLED AMPLIFIER :

The transformer as a coupling device has the following advantages :

(1) the ohmic resistance of the windings being small in comparison with a.c. impedance, the efficiency may be much higher than in *RC* coupled amplifier.

(2) the turns ratio may be chosen so that impedance matching may be utilised to obtain maximum power gain and hence maximum current gain. The input impedance in a transistor is usually lower than the output impedance. Consequently, a step down transformer can be used to increase the current gain of the amplifier.

The circuit diagram of a transformer coupled amplifier is shown in fig. 40 (*a*) with its equivalent circuit in fig. 40 (*b*). We take here the transformer as ideal that is, coefficient of coupling is unity, magnetising current negligible, and ohmic resistance of each winding negligible. Then the voltage and current ratio of the transformer are proportional to the turns ratio, *n*, (primary turns/secondary).



(*a*) Transformer coupled amplifier.

(*b*) Equivalent circuit approximation is used with $R_0 = 1/h_{oe}$ and $R_i = h_{ie}$

Fig. 40. Transformer coupled amplifier.

Then the source current

$$I = \frac{nV_0}{R_0} + \frac{i_0}{n} = \frac{nV_0}{R_0} + \frac{V_0}{nR_i}$$

or
$$V_0 \left(\frac{n}{R_0} + \frac{1}{nR_i} \right) = I$$

or
$$V_0 = I \frac{nR_0R_i}{n^2R_i + R_0}. \qquad \ldots(1)$$

Maximum power will be transferred to the load when

$$R_0 = n^2R_i$$

so that

$$V_0 = I \frac{n(n^2R_i)R_i}{n^2R_i + n^2R_i}$$

$$= \frac{InR_i}{2}. \qquad \ldots(2)$$

The secondary current

$$i_0 = \frac{V_0}{R_i} = \frac{nI}{2}.$$

When the amplifier is a common emitter transistor,

$$I = h_{fe}\, i_b.$$

Maximum power transfer is attained when

$$i_0 = \frac{nh_{fe}i_b}{2}$$

Then the current gain is given by

$$A_i = \frac{i_0}{i_b} = \frac{nh_{fe}}{2}.$$

$$= \frac{h_{fe}}{2}\left(\frac{R_0}{R_i}\right)^{1/2}.$$

In the case of a transistor, as already stated, $R_i \ll R_0$ and consequently $n > 1$ is preferred. The approximation $R_o = 1/h_{oe}$ and $R_i = h_{ie}$ is not valid under all conditions. Further, in transistor the input circuit can not be treated as isolated from output circuit.

The chief advantage of the transformer as a coupling device may be the large inductance required for satisfactory amplification of long pulses or low frequencies. But in that case, transformer would be bulky and costly. However, we employ miniature transformers for coupling transistor amplifiers, although the transformers designed for audio frequency amplification generally have poor low frequency response.

Another circuit for such an amplifier is shown in fig 41. In this two stage transformer coupled grounded emitter amplifier,

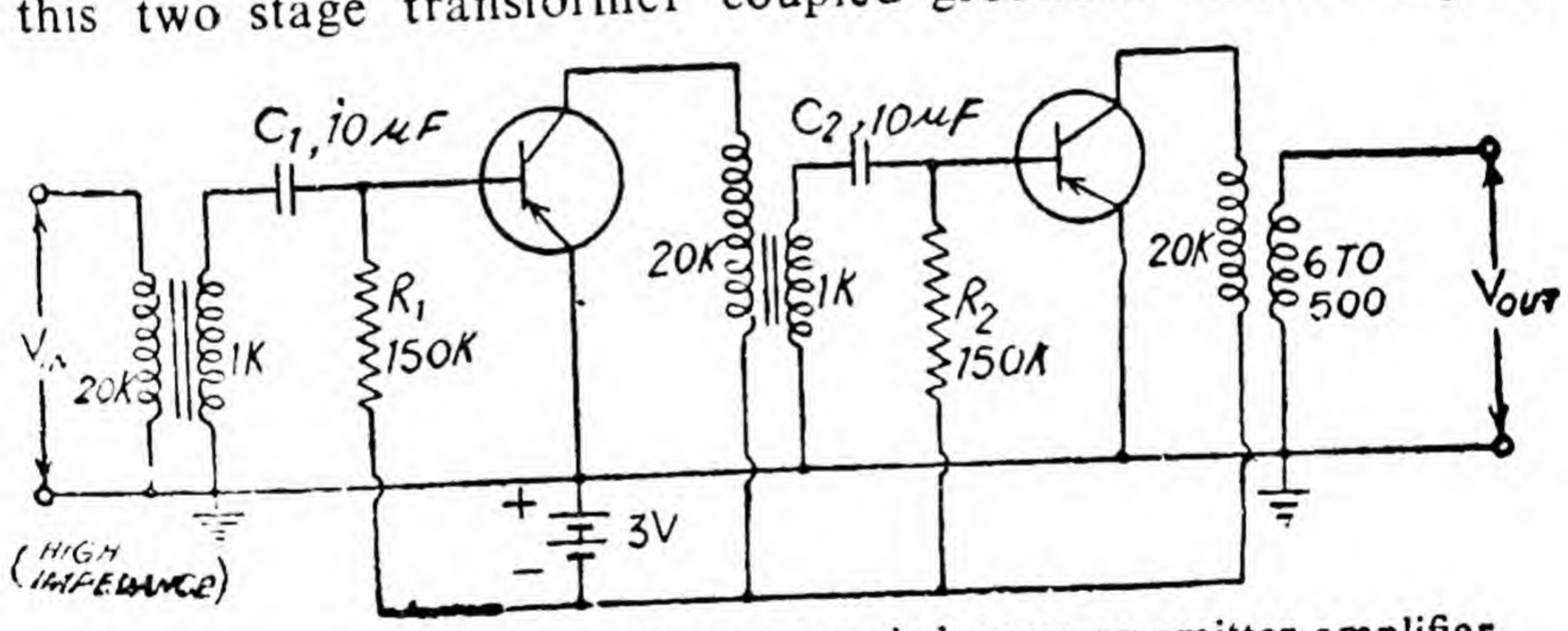


Fig. 41. Two stage transformer coupled common emitter amplifier.

interstage transformer have primary impedances of 20 *K*. ohms each, and secondary impedances of 1000 ohms each. To prevent earthing of base bias as through the low d.c. resistance of the transformer secondary winding, two capacitors each of $10\mu F$ in value are used. The two resistors R_1 and R_2 each of value 150 *K*-ohms are inserted to apply proper forward bias for the base emitter circuit.

For large gain turn ratio of the transformer should be of proper value so as to match the output resistance of the first transistor to the input resistance of the second. The primary, which is larger as compared to secondary is connected to the collector circuit to step down the voltage from the collector circuit of the first stage to the base of the next stage. The net result is to

step up the current which is required because transistors are current operated device.

Perfect matching results in distortion since variations in input resistance change the input current. A constant current will, however, be driven to the next stage if turns ratio is small resulting in lower distortion and lower gain.

The overall gain of transformer coupled amplifier is much greater than that of RC coupled amplifier. For this combination it is approximately 50 db (power gain).

16·24. NEGATIVE FEED BACK AMPLIFIERS :

Negative feed-back can be employed in transistor amplifiers for the same reason and in the same manner as in vacuum tube amplifiers. We have already discussed in chapter 11 that negative feed-back improves amplifier stability and reduces distortion. It has also been discussed that how negative feed back increases bandwith of an amplifier at high and low frequencies. The important feature of negative feed back in transistor amplifiers is the reduction in the variations in gain caused by different transistors.

Before proceeding to discuss feed back circuits, we shall show the effect of feed back on input and output impedances.

(*i*) *Effect of input impedance of a transistor amplifier* :

A is the normal gain of the amplifier without feed back. BV_0 is the fraction of the output voltage which is fed back to the input terminals. Without feed back the input impedance is

$$Z_i = \frac{e_1}{i_1} = \frac{V_i}{i_1} \text{ (since } V_i = e_1)$$

With feed back, the input impedance Z_{if}

$$Z_{if} = \frac{e_1 - BV_0}{i_1}.$$

But output voltage

$$V_0 = Ae_1,$$

so that

$$Z_{if} = \frac{e_1 - AB\,e_1}{i_1} = \frac{e_1}{i_1}(1 - AB)$$

$$= Z_i (1 - AB). \qquad \ldots(1)$$

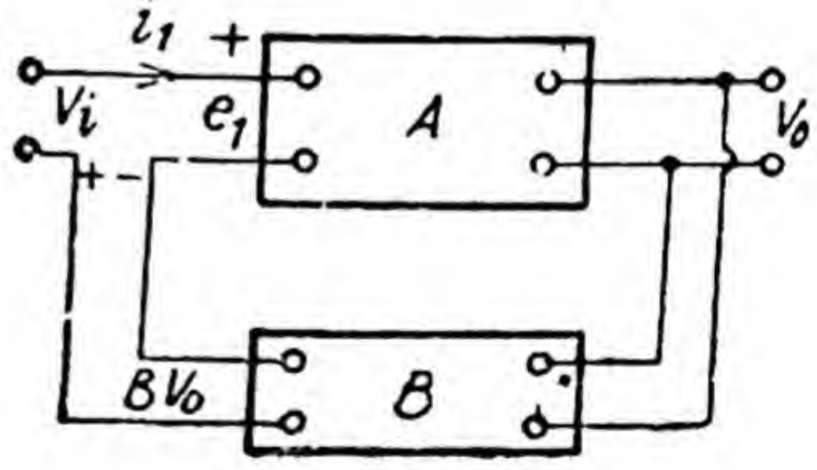


Fig. 42. Input impedance increases due to negative feed back B is feed back ratio.

In negative feed back, $(1 - AB)$ is greater than unity and consequently, Z_{if} is greater than Z_i. That is, due to negative feed back input impedance of a transistor amplifier increases.

(*ii*) *Effect on output impedance of a transistor amplifier :*

The output impedance without feed back is

$$Z_0 = \frac{V_0}{i_0}.$$

With BV_0 as the voltage fed back, the output loop current i_0' (after the feed back loop is closed) is given by

$$i_0' = \frac{V_0 - AB\,V_0}{Z_0}$$

$$= \frac{V_0}{Z_0}(1 - AB),$$

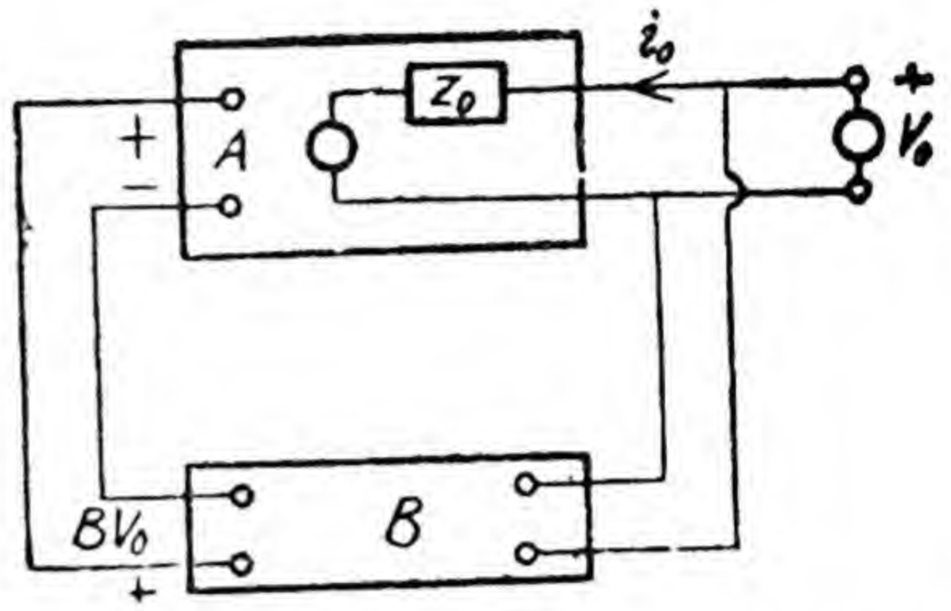


Fig. 43: Output impedance decreases due to negative feed back.

so that output impedance is

$$Z_{of} = \frac{V_0}{i_0'} = \frac{Z_0}{(1 - AB)} \qquad \ldots(2)$$

Since in negative feed back $(1 - AB) > 1$, Z_{of} is less than Z_0. That is, output impedance decreases due to negative feed back.

Feed back Circuits : Prior to discussing the feedback circuits, let us remind that voltage gain of an amplifier with feed back is

$$A_{vf} = \frac{A_v}{1 - B_v A_v} \qquad \ldots(3)$$

and similarly, current gain is

$$A_{if} = \frac{A_i}{1 - B_i A_i} \qquad \ldots(4)$$

where A_v and A_i denote the voltage and current gain without feed back; and B_v and B_i denote the voltage and current feed back ratio.

(1) Current Feed back amplifier : The circuit is shown in fig. 44. The emitter resistor R_e is not being by-passed by any condenser, and the negative feed back is provided by this resistor. This carries substantially the same current as the load resistor R_L when an input signal is applied. The a.c. voltage developing across R_e is also applied at the input circuit of the amplifier. Thus it provides a feed back voltage proportional to the current in the load. Since $R_L \gg R_e$, the presence of R_e in the emitter circuit does not substantially change the current gain of the amplifier. This is also known as '*series derived*' feed back and it tends to maintain constant current in the load. It effectively increases the output resistance of the transistor and is useful therefore, when the output is required in the form of a current.

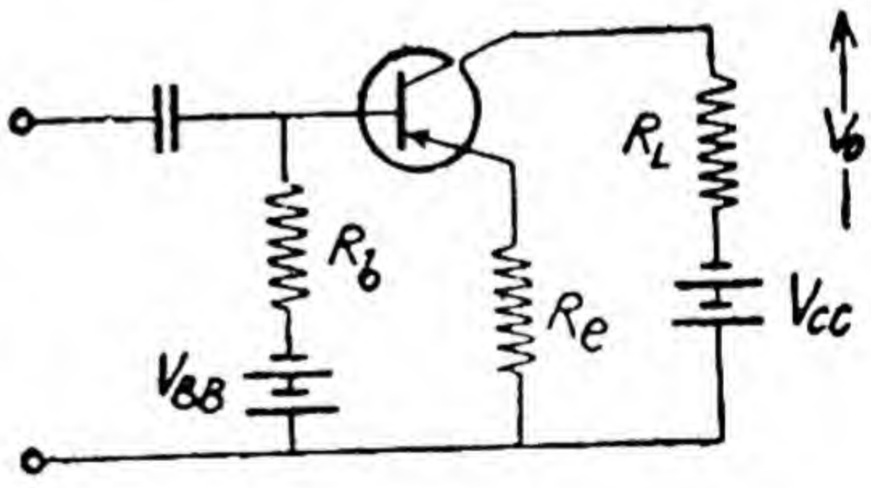


Fig. 44. Curren' feed back amplifier.

(*a*) *Voltage gain*—The voltage gain of such an amplifier without feed back is given by equation (15) art. 16·2,

$$A_v = \frac{-h_{fe}\, R_L}{R_L\,(h_{ie}\, h_{oe} - h_{fe}\, h_{re}) + h_{ie}}. \qquad \ldots(5)$$

The feedback ratio B_v based on the circuit voltage is

$$B_v = \frac{V_e}{V_0} = \frac{-i_e\, R_e}{i_c\, R_L} = \frac{(i_b + i_c)\, R_e}{i_c\, R_L},$$

where the symbols have the same meaning as mentioned in art. 16·2. Since $i_c \gg i_b$, we have

$$B_v = \frac{R_e}{R_L}. \qquad \ldots(6)$$

Therefore voltage gain of such an amplifier with feed back will be

$$A_{vf} = \frac{A_v}{1 - B_v A_v}.$$

Using equations (5) and (6), we get

$$A_{vf} = \frac{\dfrac{-h_{fe}\, R_L}{R_L\,(h_{ie}\, h_{oe} - h_{fe}\, h_{re}) + h_{ie}}}{1 + \dfrac{\dfrac{R_e}{R_L}.\, h_{fe}\, R_L}{R_L\,(h_{ie}\, h_{oe} - h_{re}\, h_{fe}) + h_{ie}}}$$

$$= \frac{-h_{fe}\, R_L}{h_{ie} + R_L\,(h_{ie}\, h_{oe} - h_{re}\, h_{fe}) + R_e h_{fe}}. \qquad \ldots(7)$$

Since the denominator of A_{vf} includes one more term $R_e\, h_{fe}$ compared to the denominator of A_v, the voltage gain with feedback is less than that without feed back.

Example. Suppose in amplifier circuit discussed above

$h_{fe} = 50$ $\qquad$ $R_L = 5$ k. ohm.

$h_{ie} = 2$ k. ohm. $\qquad$ $R_e = 100$ ohms.

$h_{oe} = 5 \times 10^{-6}$ ohms.

$h_{re} = 10^{-4}$,

then to find the voltage gain, we proceed as follows :

We have shown that

$$A_{vf} = \frac{-h_{fe}\, R_L}{h_{ie} + R_L\,(h_{ie}\, h_{oe} - h_{re}\, h_{fe}) + R_e\, h_{fe}}$$

$$= \frac{-50 \times 5 \times 10^3}{2 \times 10^3 + 5 \times 10^3\,(2 \times 10^3 \times 5 \times 10^{-6} - 50 \times 10^{-4}) + 50 \times 100}$$

$$= \frac{-250}{2 + 5\,(100 - 50) \times 10^{-4} + 5}$$

$$= \frac{-250}{7{\cdot}025} \approx 35{\cdot}7.$$

(*b*) *Input resistance*—The input resistance can be calculated with help of fig. (45).

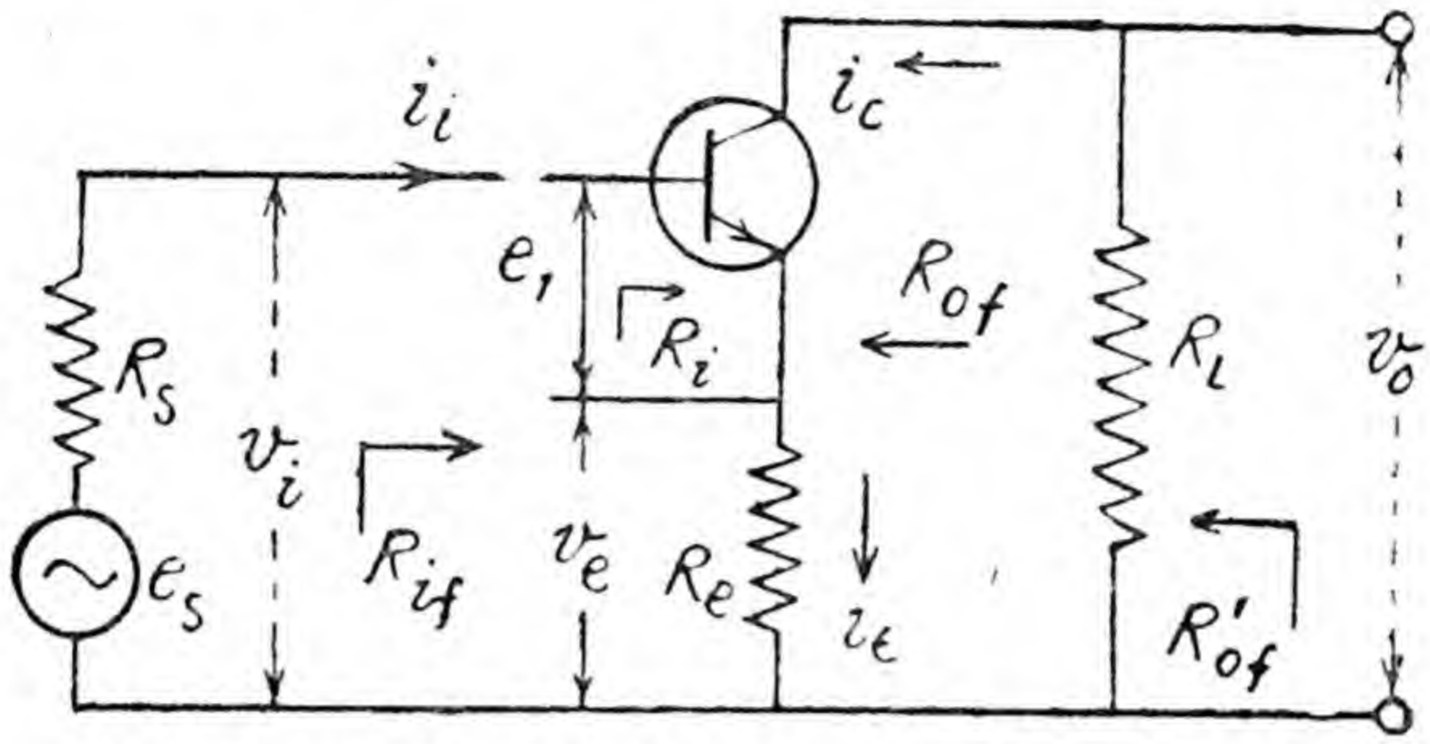


Fig. 45. Simplified diagram of a.c. quantities.

From figure (45),

$$R_{if}=\frac{v_i}{i_i}=\frac{e_1+v_e}{i_i} \quad (\because \quad v_i=e_1+v_e) \qquad ...(8)$$

Also $v_e=R_e.i_e$ and $i_i=i_e-i_c$

$$\therefore \qquad R_{if}=\frac{e_1}{i_i}+\frac{R_e.i_e}{i_e-i_c}=R_i+\frac{R_e}{1-i_c/i_e} \quad \left(\because \frac{e_1}{i_i}=R_i\right) \qquad ...(9)$$

We know that

$$i_c/i_e=h_{fb}=h_{fe}/(1+h_{fe})$$

Substituting this value in equ. (9), we get

$$R_{if}=R_i+\frac{R_e}{1-\{h_{fe}/(1+h_{fe})\}}$$

or
$$R_{if}=R_i+R_e\,(1+h_{fe}) \qquad ...(10)$$

Example : *A common emitter transistor has* $h_{fe}=50$ *and* $h_{ie}=1K$ *ohm. With* $R_L=1K$ *ohm,* $R_e=100$ *ohms, calculate the input resistance with feed back. Calculate the gain if* R_e *were adequately bypassed and input resistance with feed back.*

$$A_v=\frac{A_i\,R_L}{R_i}\approx\frac{-h_{fe}\,R_L}{h_{ie}}=\frac{-50\times 1k}{1k}=-50$$

$$R_i\approx h_{ie}=1\text{K ohm}$$

$$R_{if}=h_{ie}+R_e\,(1+h_{fe})$$

$$=1\text{K}+100\,(1+50)$$

$$=\mathbf{6{\cdot}1K\ ohm.}$$

(2) Voltage Feed back Amplifier : The circuit is shown in fig. 46.

The negative feed back is provided by a resistor R_f connected between collector and base. The current feedback through R_f is proportional to the voltage across R_L and this form of feedback is also termed as '*shunt derived*'. It maintains the voltage across R_L constant and thus effectively reduces the output resistance of the

transistor. This circuit is useful when output in the voltage form is desired.

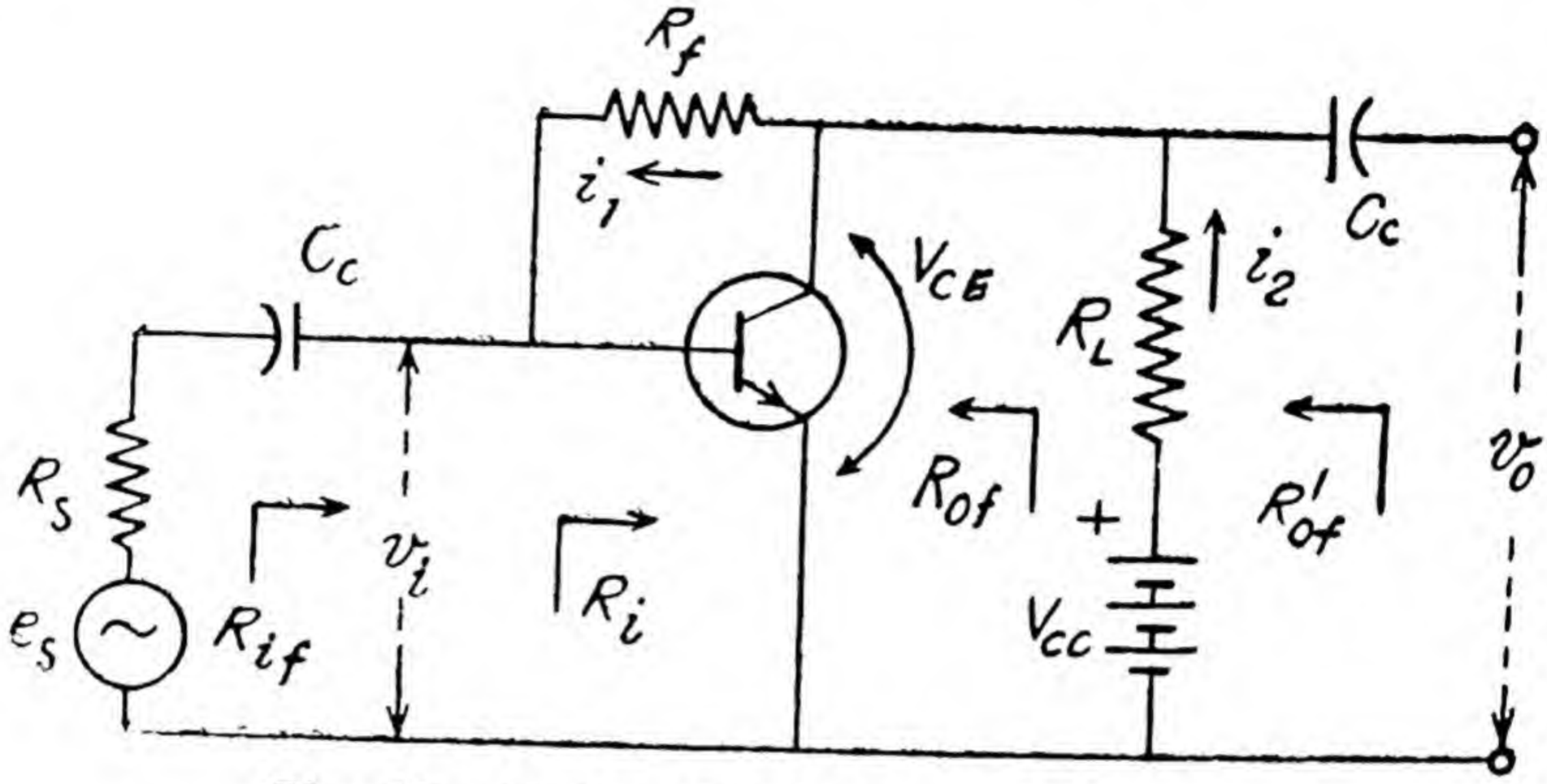


Fig. 46. Voltage feedback amplifier.

(a) *Voltage gain*—From equation (12) art. 16·2, the circuit gain of such an amplifier without feedback is given by

$$A_i = \frac{h_{fe}}{h_{oe}\, R_L + 1}. \qquad \ldots(11)$$

In order to find feedback ratio B_i, we find that major portion of current flows through R_L and the remainder flows back into the input circuit through R_f. If we denote the input resistance of the circuit by R_i, then

$$B_i = \frac{\text{current feedback into the input circuit}}{\text{current through the load}}.$$

$$= \frac{i_1}{-i_2} = \frac{\dfrac{V_{CE}}{R_i + R_f}}{-\dfrac{V_{CE}}{R_L}}$$

$$\simeq -\frac{R_L}{R_i + R_f}$$

$$= -\frac{R_L}{R_f} \text{ since } R_f \gg R_i. \qquad \ldots(12)$$

We have already mentioned that the current gain of the amplifier with beedback is

$$A_{if} = \frac{A_i}{1 - B_i A_i}$$

Using equations (11) and (12), we get

$$A_{if} = \frac{\dfrac{h_{fe}}{h_{oe}\, R_L + 1}}{1 + \dfrac{\dfrac{R_L}{R_f}\, h_{fe}}{h_{oe}\, R_L + 1}} \qquad \ldots(13)$$

$$= \frac{h_{fe}}{1 + h_{oe} R_L + \frac{R_L}{R_f} h_{fe}} \qquad \ldots(14)$$

From eqs. (11) and (14), by comparison, it becomes obvious that $A_{if} < A_i$. Thus the current is affected by the presence of R_f, which is far larger than R_L.

Example : Suppose in amplifier circuit discussed above

$h_{fe} = 50$
$h_{ie} = 2K$ ohm $\qquad R_L = 5K$ ohm.
$h_{oe} = 5 \times 10^{-6}$ ohms $\qquad R_f = 100K$ ohm.
$h_{re} = 10^{-4}$

then find the current gain.

From equation (14),

$$A_{if} = \frac{h_{fe}}{1 + h_{oe}\, R_L + \frac{R_L}{R_f}\, h_{fe}}$$

$$= \frac{50}{1 + 5 \times 10^{-6} \times 5 \times 10^3 + \frac{5 \times 10^3}{100 \times 10^3} \times 50}$$

$$= \frac{50}{1 + \cdot 025 + 2 \cdot 5} = \frac{50}{3 \cdot \diagdown 25}$$

$$\approx 14 \cdot 4.$$

(*b*) ***Input resistance***—The input resistance R_{if} is given by

$$R_{if} = \frac{h_{ie} . \{R_f/(1 - A_v)\}}{h_{ie} + \{R_f/(1 - A_v)\}} \qquad \ldots(15)$$

where $$A_v = \frac{v_o}{v_i} = \frac{-h_{fe}\, R_L'}{h_{ie}}$$

$$R_L' = \frac{R_f . R_L}{R_f + R_L}.$$

16·25. EFFECT OF FEED BACK ON BANDWIDTH :

We have seen that amplifier gain falls off at low and high frequencies. At high frequencies, the shunt capacitances can not be considered as open circuited as at mid frequencies and hence due to the reactance of shunt capacitances, the amplifier gain falls off. Let ω_1 be the frequency at which the gain falls to $1/\sqrt{2}$ times the mid-frequency value A_0 the gain at high frequency ω is given by the expression

$$A_v = \frac{A_0}{1 + j\, \omega/\omega_1} \qquad \ldots(1)$$

At low frequencies, the series capacitances can no longer be taken as short circuited and hence the gain falls off. The gain of the amplifier at low frequencies is given by

$$A_v = \frac{A_0}{1 - j\, \omega_2/\omega} \qquad \ldots(2)$$

where ω_2 is the low frequency at which the gain falls to $1/\sqrt{2}$ times the mid frequency value.

Due to feed back, the lower and higher frequency limits are changed. With feed back, the gain is expressed as

$$A_{vf} = \frac{A_v}{1 - A_v B} \qquad \ldots(3)$$

At high frequeucies, the gain of amplifier with feed back is given by

$$A_{vf} = \frac{A_0}{1 + j\frac{\omega}{\omega_1} - A_0 B} \qquad \ldots(4)$$

For low frequencies, the gain of amplifier with feed back is given by

$$A_{vf} = \frac{A_0}{1 - j\frac{\omega_2}{\omega} - A_0 B} \qquad \ldots(5)$$

It is clear from expressions (4) and (5) that the amplifier gain falls to $1/\sqrt{2}$ times the mid frequency value at frequencies ω' and ω'' in the presence of feed back.

$$\left.\begin{aligned} &\text{Here} \quad \frac{\omega'}{\omega_1} = 1 - A_0 B \text{ for high frequencies} \\ &\text{and} \quad \frac{\omega_2}{\omega''} = 1 - A_0 B \text{ for low frequencies.} \end{aligned}\right\} \qquad \ldots(6)$$

The new cut off frequencies ω' and ω'' in presence of feedback are well above and below the cut off frequencies of amplifier without feedback. Thus the bandwidth is extended as a result of feed back.

16·26. UNTUNED POWER AMPLIFIERS :

In large signal amplifiers, the prime requirement is to obtain maximum power output free from distortion. To avoid distortion, final stage is always operated as class *A* or class *B*.

16·26-1. CLASS A AMPLIFIERS :

Circuit for class *A* amplifier is shown in fig. 47. In class *A* condition, transistor operates during the whole of the input and

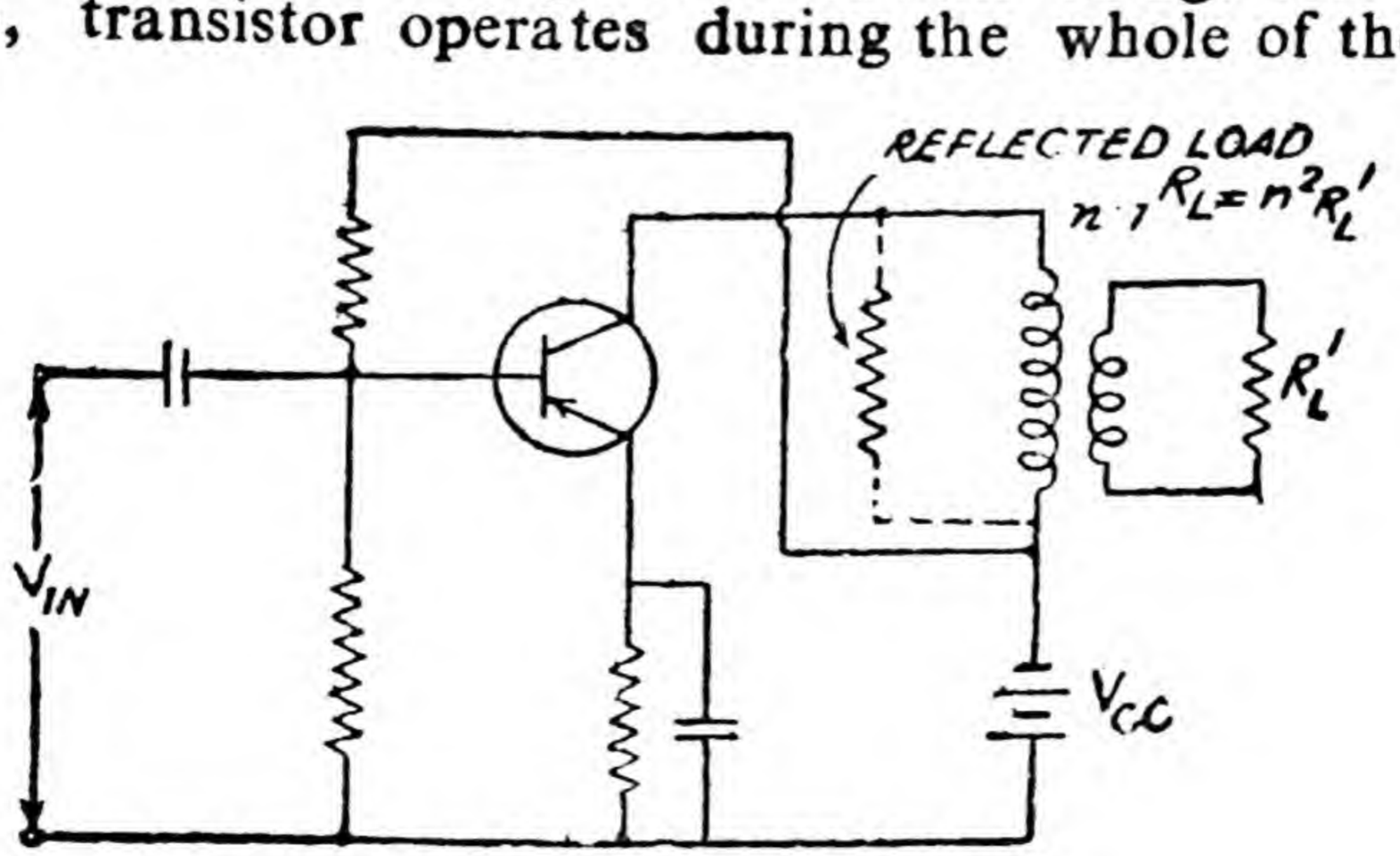


Fig. 47. Class *A* amplifier.

output cycle and results in distortion free amplification provided the device is linear.

For high efficiency it is essential that the collector current and collector voltage swings should be large.

An external load R_L (reflected load) is coupled by the transformer of turns ratio n to the power amplifier. The reflected load into the collector circuit is $n^2 R'_L$.

In fig. 48, idealised characteristics for common emitter configuration with load line of slope $-1/R_L$ are shown. Under a.c. conditions, reflected load appears across the transformer giving an a.c. load line. When a.c. signal is applied, collector current fluctuates due to which operating point moves up and down the load line and during a portion of the cycle instantaneous voltage across the transistor will exceed the supply voltage V_{cc}. For class A operation, operating point Q will be located half way up the load line. Corresponding to this point,

$$V_{C_1} = V_{cc} = \frac{V_{max}}{2},$$

$$I_{C_1} = \frac{I_{max}}{2},$$

where $V_{max.}$ and $I_{max.}$ are permitted voltage and current for the transistor in this configuration.

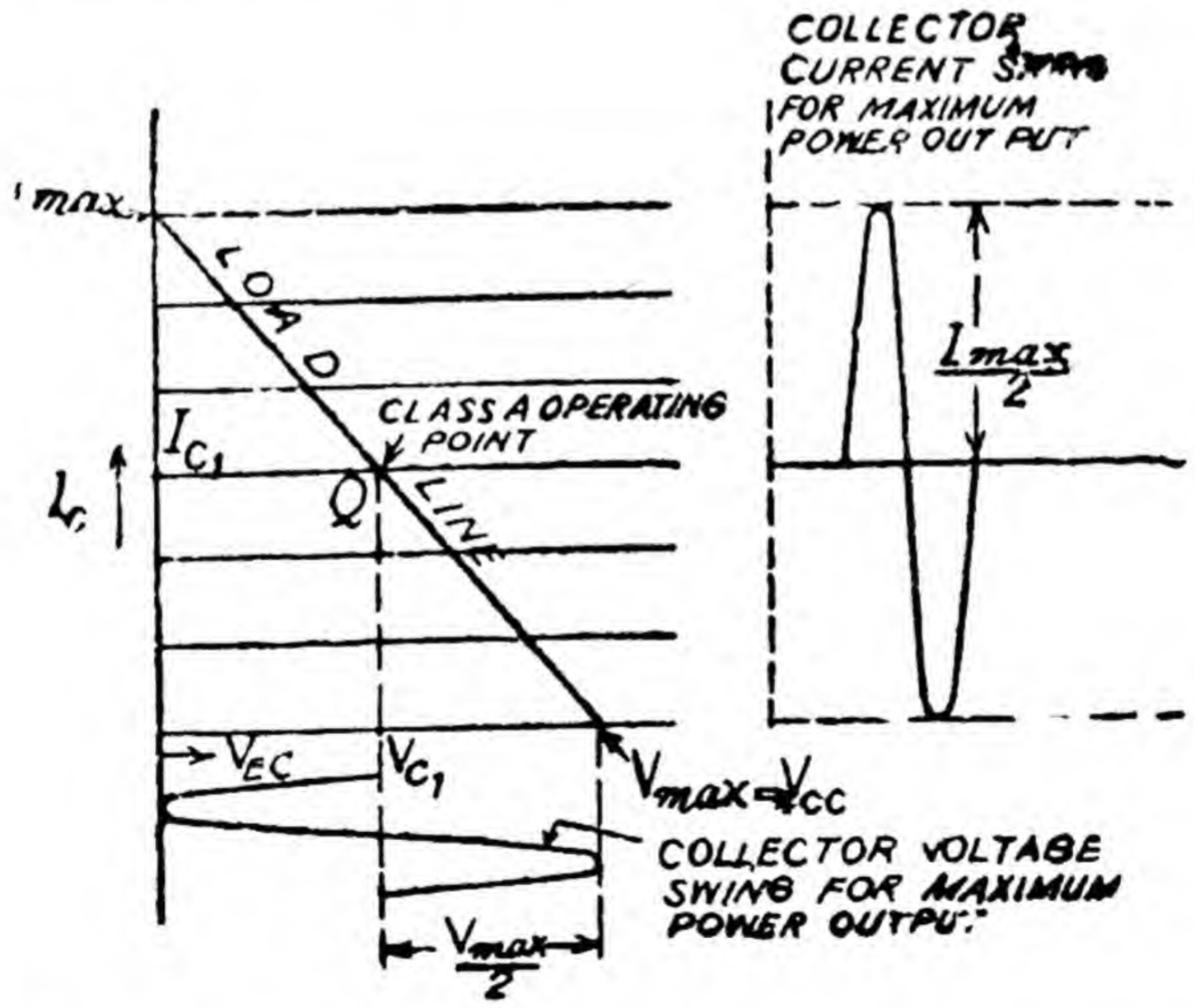


Fig. 48. Idealised CE characteristics with output voltage and current-wave forms.

Under maximum capacity of class A amplifier, voltage swings from $V_{max.}$ to zero and current from $I_{max.}$ to zero. Hence

$$V_{r.m.s.} = \frac{V_{max.}}{2\sqrt{2}}, \qquad I_{r.m.s.} = \frac{I_{max.}}{2\sqrt{2}}$$

$$\text{Useful a.c. power} = V_{r.m.s.}\ I_{r.m.s.} = \frac{V_{max.}\ I_{max.}}{4}$$

$$\therefore \quad P = \frac{V_{max.}\ I_{max.}}{8}.$$

Total power supplied to the circuit

$$P_s = V_{cc} I_{C_1} = \frac{V_{max.}\ I_{max.}}{4}.$$

Hence efficiency of power amplifier, $\eta = P / P_S$.

$$\text{percentage efficiency } \eta = \frac{P}{P_s} \times 100 = 50\%,$$

If, however, instead of transformer, series load resistor is used then because of the d.c. power dissipation in it, efficieney reduces to half of this value.

16·26-2. CLASS B AMPLIFIER :

Circuit for such an amplifier is shown in fig. 49. In class *B* conditions, the operating point is chosen at the collector current cut off (fig. 50) so that in the absence of the signal, collector current is almost negligible. The transistor will thus conduct only during one half of the input wave and will be cut off during the rest half cycle. Collector current increasing with signal amplitude. In contrast to class *A* amplifier, in this case current taken from the battery is proportional to signal amplitude.

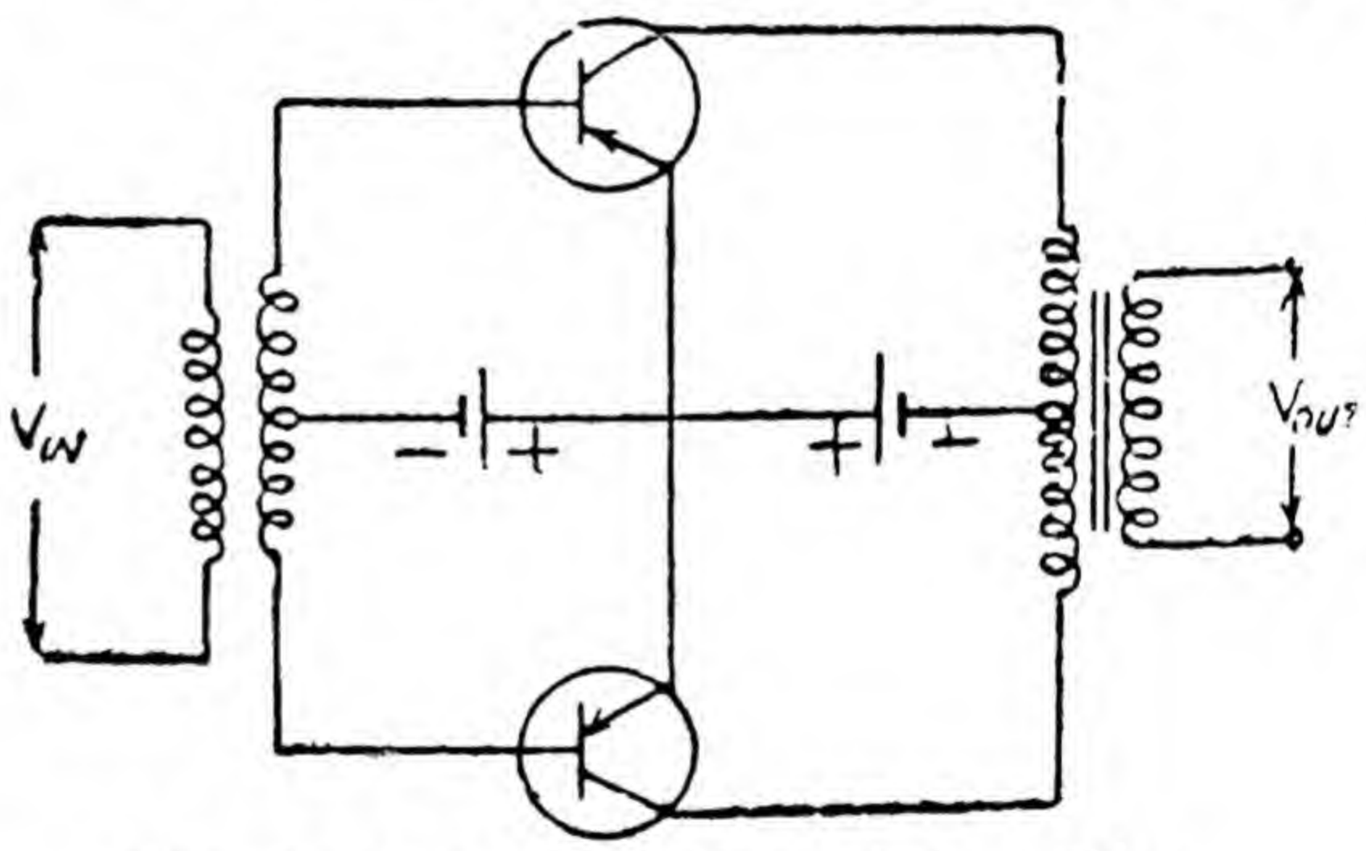


Fig. 49. Push pull Class B amplifier.

To produce original input signal (to get response for both halves of the input cycle) it is necessary to use two transistors operating out of phase (push pull arrangement).

The efficiency of class B amplifier :

For push pull arrangement since both halves are useful inputs, the calculation for a.c. power output will, however, remain the same, *i.e.*,

$$V_{C_1} = V_{max.}$$

$$I_{C_1} = I_{max.}$$

where $V_{max.}$ and $I_{max.}$ signify the maximum permitted voltage and current in class *B* operation. So maximum a.c. power output will be

$$P = V_{r.m.s.}\ I_{r.m.s.} = \frac{V_{max.}}{\sqrt{2}} \cdot \frac{I_{max.}}{\sqrt{2}}$$

$$P = \frac{V_{max.}\ I_{max.}}{2}$$

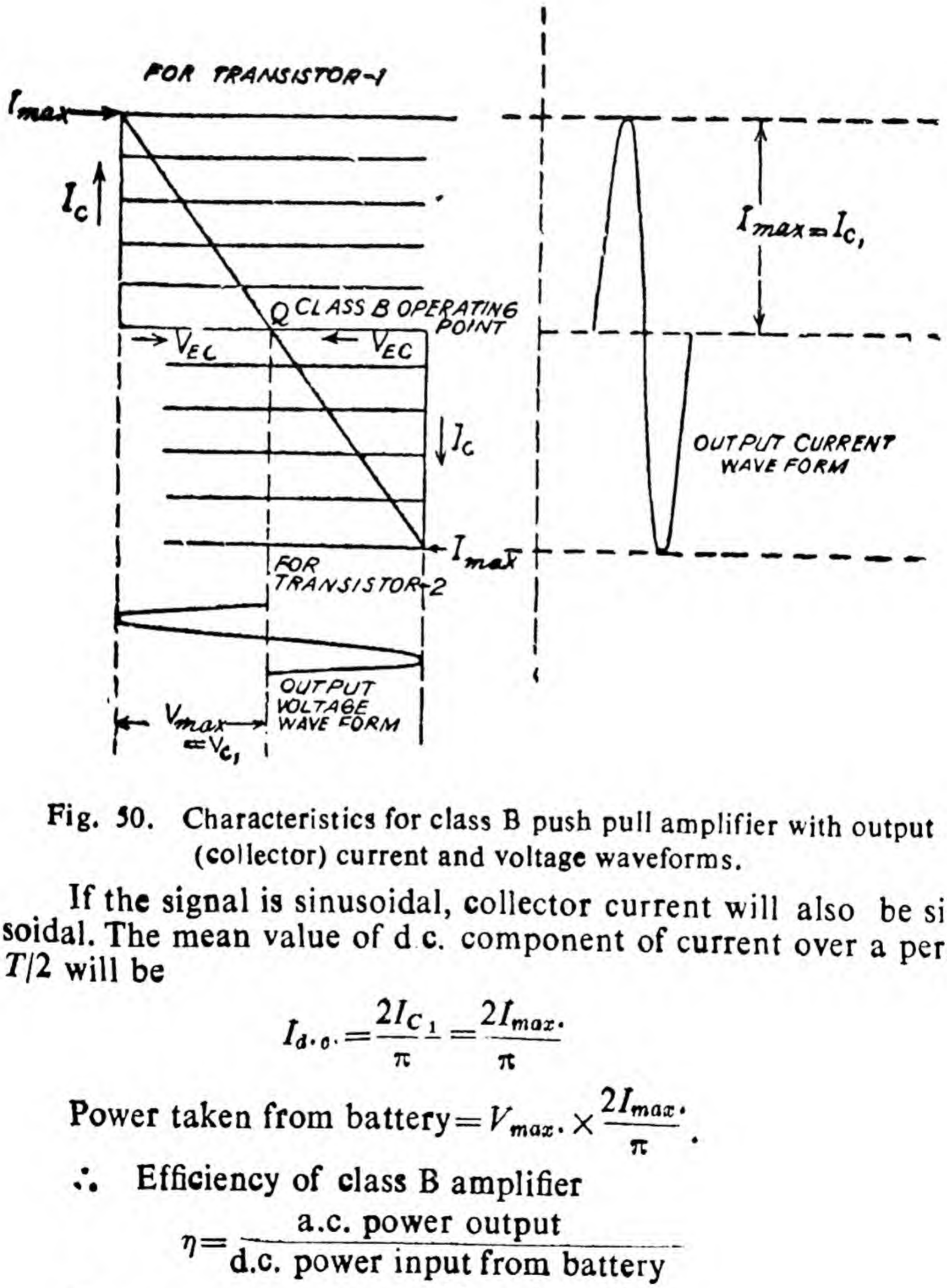


Fig. 50. Characteristics for class B push pull amplifier with output (collector) current and voltage waveforms.

If the signal is sinusoidal, collector current will also be sinusoidal. The mean value of d.c. component of current over a period $T/2$ will be

$$I_{d.c.} = \frac{2I_{C_1}}{\pi} = \frac{2I_{max.}}{\pi}$$

Power taken from battery $= V_{max.} \times \frac{2I_{max.}}{\pi}$.

∴ Efficiency of class B amplifier

$$\eta = \frac{\text{a.c. power output}}{\text{d.c. power input from battery}}$$

Percentage efficiency $\eta = \dfrac{V_{max.}\ I_{max.}/2}{2V_{max.}\ I_{max.}/\pi}$

$$= \frac{\pi}{4} \times 100 = 78{\cdot}50\%.$$

Now we shall proceed to calculate the maximum collector to collector load. Let the load resistance between primary terminals

be R_L ; then voltage drop across R_L will be $\frac{I_{max.}\ R_L}{2}$. For a single transistor, this drop will be halved, *i.e.*, voltage drop between collector and emitter of a single transistor

$$=\frac{1}{2}.\frac{I_{max.}\ R_L}{4}=\frac{I_{max.}\ R_L}{4}$$

Collector voltage $V_C=V_{max.}-\frac{I_{max.}\ R_L}{4}$

or $$V_{max.}-V_C=\frac{I_{max.}\ R_L}{4}.$$

$$\therefore \quad R_L=\frac{4}{I_{max.}}(V_{max.}-V_{C_1})$$

$$=4.\frac{V_{max.}}{I_{max.}}\left(1-\frac{V_{C_1}}{V_{max.}}\right)$$

If $\frac{V_{C_1}}{V_{max.}}$ is negligibly small, it may be neglected ; then

$$R_L=4.\frac{V_{max.}}{I_{max.}},$$

which gives collector to collector load in a class *B* push-pull amplifier.

Effect of signal amplitude on the efficiency : Suppose $|V_S|$ is the signal amplitude for which output is maximum. Now suppose it is reduced *a* time, so that new signal amplitude is $a|V_S|$. Corresponding to this value, maximum collector current and maximum collector voltage will be reduced by the same factor if load and battery voltages are kept constant. Therefore power output for reduced signal will be

=r.m.s. collector current × r.m.s. collector voltage

$$=\frac{aI_{max.}}{\sqrt{2}}\times\frac{a\ V_{max.}}{\sqrt{2}}=\frac{a^2\ V_{max.}\ I_{max.}}{2}.$$

Now to calculate the efficiency for the reduced signal, the further requirement is the calculation of power taken from battery. In class *B* amplifier, battery voltage equals V_{max}, the mean value of current being $\frac{2\ I_{max.}}{\pi}$ which for new signal amplitude will be $\frac{2a\ I_{max.}}{\pi}$. Therefore power taken from battery

$$=\frac{2a\ I_{max.}\ V_{max.}}{\pi}.$$

The efficiency for reduced input signal

$$=\frac{\text{a.c. power output}}{\text{d.c. power input}}$$

$$= \frac{a^2 I_{max.} \, V_{max.}/2}{2a I_{max.} \, V_{max.}/\pi} = a \frac{\pi}{4}$$

The expression for the efficiency shows that efficiency is reduced as amplitude decreases.

Power dissipation. Power dissipated,

P_d = power supplied from battery — power output

$$= \frac{2a I_{max.} \, V_{max.}}{\pi} - \frac{a^2 I_{max.} \, V_{max.}}{2}$$

To find the value of a for which power dissipated is maximum, we must put $\frac{\partial P_d}{\partial a} = 0$, *i.e.*,

$$\frac{2 I_{max.} \, V_{max.}}{\pi} = a I_{max.} \, V_{max.}$$

$$a = \frac{2}{\pi}.$$

Therefore power dissipation is maximum when signal amplitude is $\frac{2}{\pi} | V_S |$, where $| V_S |$ is the signal amplitude for maximum output.

$$\therefore \quad (P_d)_{max.} = \frac{2 . \frac{2}{\pi} I_{max.} \, V_{max.}}{\pi} - \frac{\frac{4}{\pi^2} . I_{max.} \, V_{max.}}{2}$$

$$= \frac{2}{\pi^2} I_{max.} \, V_{max.},$$

which is just half the power supplied from d.c. source and hence the efficiency of the amplifier is 50%.

In terms of collector to collector load, maximum dissipation can be expressed as

$$(P_d)_{max.} = \frac{I^2_{max.}}{2\pi^2} .4 \frac{V_{max.}}{I_{max.}}$$

$$= \frac{I_{max.}}{2\pi^2} R_L.$$

16·27. TUNED AMPLIFIERS :

A tuned amplifier is one which uses a parallel tuned circuit (anti resonant circuit) as its load impedance. These amplifiers are used for amplification of a narrow band of frequencies. This is due to the fact that a parallel tuned circuit has high impedance at its frequency of resonance and the impedance falls off sharply as the frequency departs from the frequency of resonance. In case of tuned amplifiers, the gain versus frequency curve is very similar to the impedance versus frequency curve of the tuned circuit. For

the adjustment of the resonant frequency at the centre of the band of frequencies to be amplified, a resonant circuit generally uses either a variable capacitor or a variable inductor. Thus the impedance of the tuned circuit or the gain of the tuned amplifier remains more or less constant over this narrow band of frequency. Transistor tuned amplifiers may be divided into two categories : (*i*) small signal tuned amplifier and (*ii*) large signal tuned amplifiers. Small signal transistor tuned amplifiers are meant for amplifying small signals at radio frequencies. The power involved is small. They operate under class *A* operation and hence the distortion is negligibly small. Large signal transistor tuned amplifiers, on the other hand, are meant for amplifying large signals at radio frequencies. In this case, power involved is large. They operate under class *AB*, *B* or *C* conditions providing large collector circuit efficiency The main drawback is that distortion gets increased. Of course, the tuned circuit itself eliminates most of the harmonic distortion. For further reducing the harmonic distortion, they use push-pull operation. Here we shall consider small signal tuned amplifiers. To obtain a large gain, a number of tuned amplifier stages in cascade are used. The cascade tuned amplifiers can be put in the following three categories :

(1) *Single tuned amplifiers.*

(2) *Double tuned amplifiers.*

(3) *Stagger tuned amplifiers.*

(1) Single tuned amplifiers. They use one parallel tuned circuit as the load impedance in each stage. The tuned circuits in different stages are tuned to the same frequency.

(2) Double tuned amplifiers. They use two inductively coupled tuned circuits per stage. Both the tuned circuits are tuned to the same frequency.

(3) Stagger tuned amplifiers. They use a number of single tuned stages in cascade. The successive tuned circuits being tuned to slightly different frequencies.

16·28. SINGLE TUNED INDUCTIVELY COUPLED TRANSISTOR AMPLIFIER :

Fig. (51) shows the circuit of a *CE* single tuned inductively coupled transistor amplifier. Here the voltage developed across the tuned circuit is inductively coupled to the next stage. In this circuit $R_1 - R_2$ along with $R_e - C_e$ combination in the emitter circuit provide the emitter bias. The coil L of the tuned circuit and the inductively coupled coil L_2 feeding the next stage form a transformer. The parallel tuned circuit formed by capacitance C and inductor L resonate at the frequency of operation.

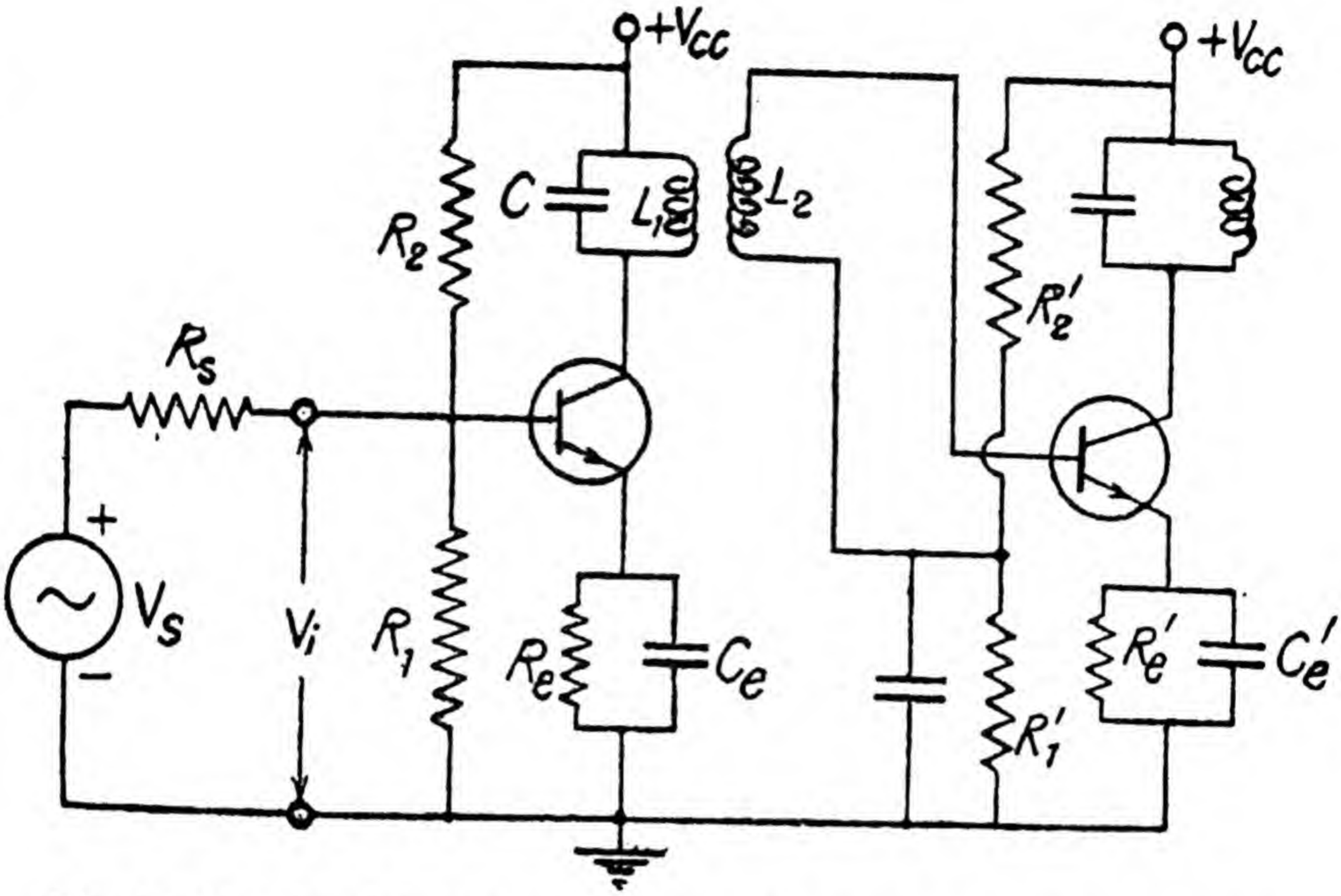


Fig. 51. Circuit of single tuned inductively coupled transistor amplifier.

The input and output equivalent circuits are shown in fig. 52. (*a*) and (*b*) respectively.

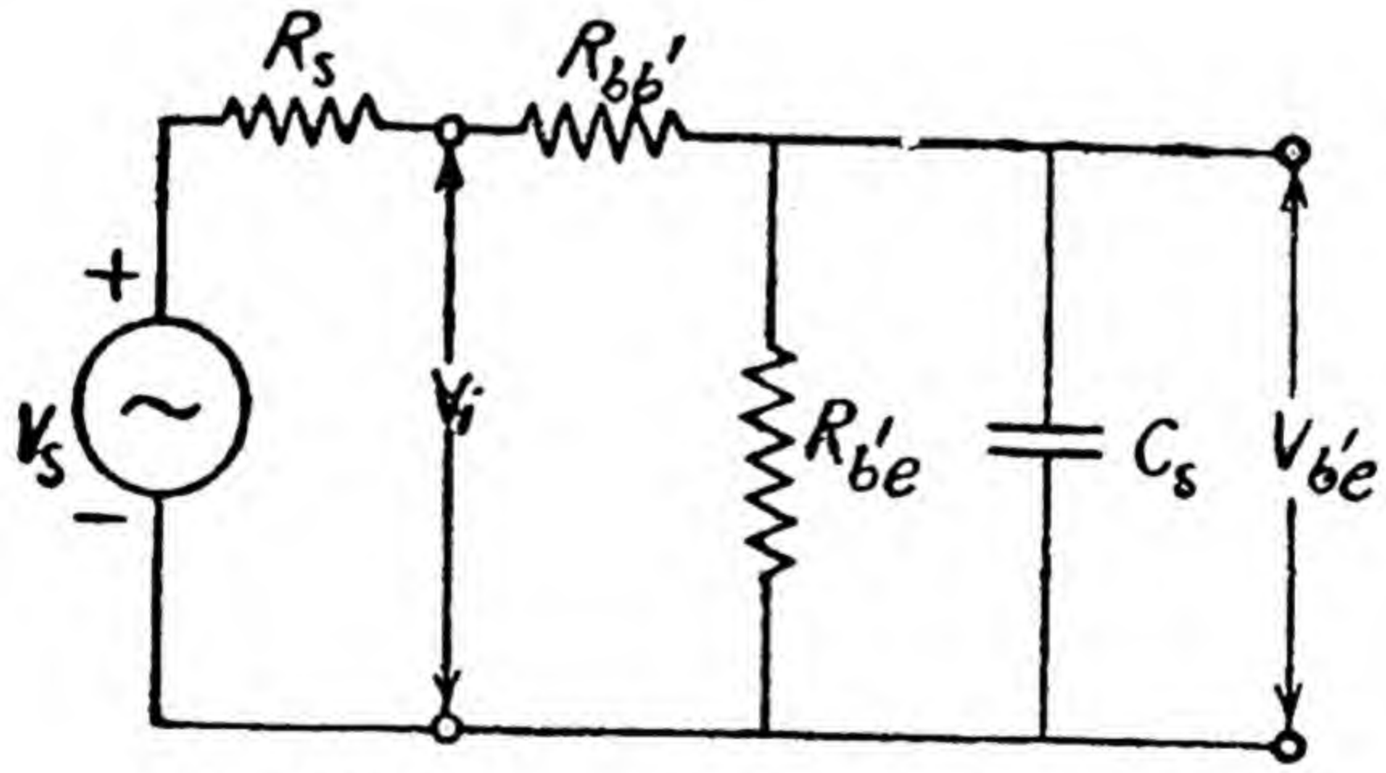


Fig. 52. (*a*) Input equivalent circuit.

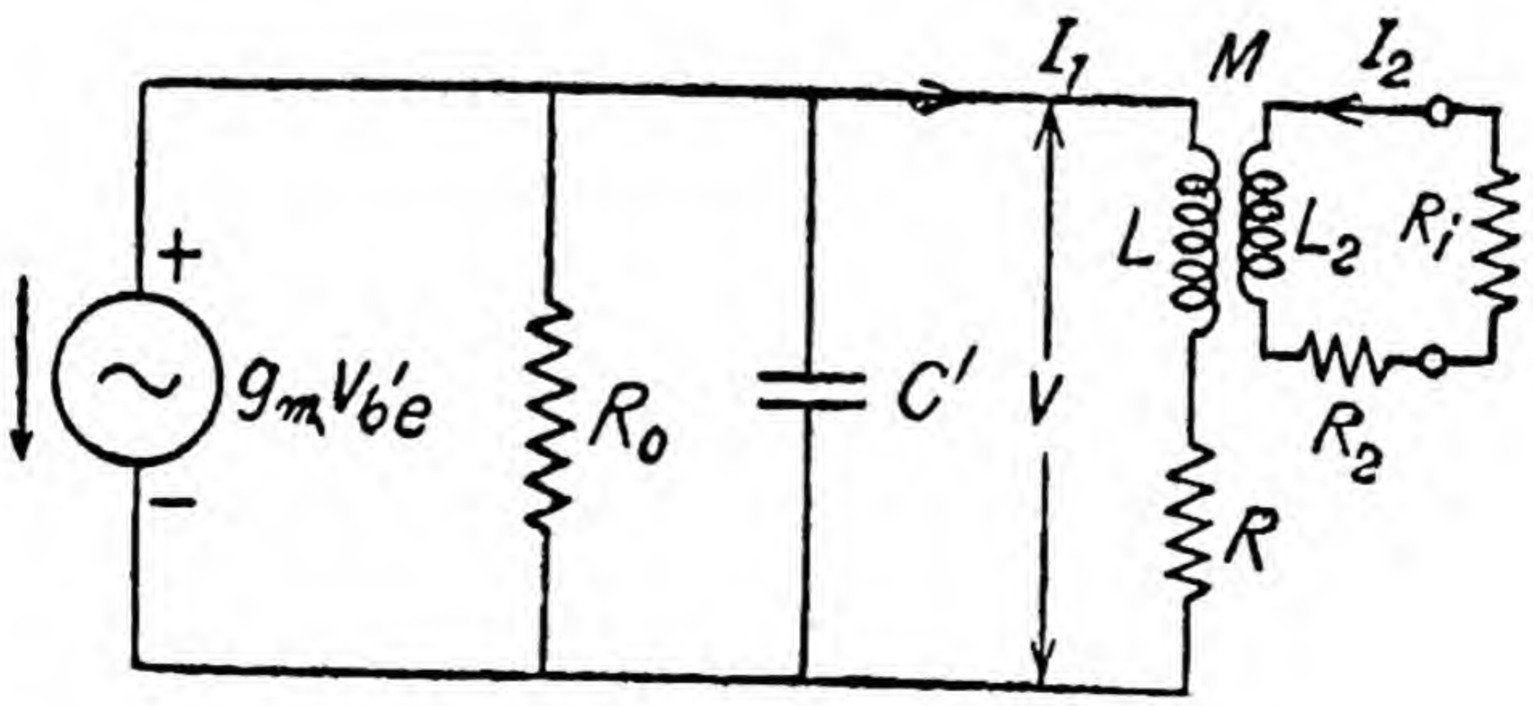


Fig. 52. (*b*) Output equivalent circuit.

$$Q_e = \frac{\text{Susceptance of inductance } L \text{ or capacitance } C'}{\text{Conductance of shunt resistance } R_{tt}}$$

or $$Q_e = \omega_0 C R_{tt} = \frac{R_{tt}}{\omega_0 L} \quad \ldots(12)$$

The impedance of C', L and R_{tt} in parallel is given by

$$\frac{1}{Z} = \frac{1}{R_{tt}} + \frac{1}{j\omega L} + j\omega C'$$

$$= \frac{1}{R_{tt}}\left[1 + \frac{R_{tt}}{j\omega L} + j\omega C' R_{tt}\right]$$

$$= \frac{1}{R_{tt}}\left[1 + j\frac{\omega\omega_0 C' R_{tt}}{\omega_0} + \frac{R_{tt}\,\omega_0}{j\omega_0\omega L}\right]$$

$$= \frac{1}{R_{tt}}\left[1 + j\,Q_e\left\{\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right\}\right] \quad \ldots(13)$$

Comparing eqn. (12) and (13), we have

$$Z = \frac{R_{tt}}{1 + jQ_e\left[\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right]} \quad \ldots(14)$$

Let δ indicate the fractional frequency variation, *i.e.*, variation in frequency expressed as a fraction of the resonant frequency

$$\delta = \frac{\omega - \omega_0}{\omega_0} = \frac{\omega}{\omega_0} - 1$$

or $$\frac{\omega}{\omega_0} = 1 + \delta$$

$$\therefore \quad Z = \frac{R_{tt}}{1 + jQ_e\left[(1+\delta) - \frac{1}{(1+\delta)}\right]} = \frac{R_{tt}}{1 + j\,2Q_e\,\delta\left[\frac{1+\delta/2}{1+\delta}\right]}$$

or $$Z = \frac{R_{tt}}{1 + j\,2Q_e\,\delta} \quad \ldots(15)$$

(Because at any frequency ω close to frequency of resonance ω_0, $\delta \ll 1$).

From fig. (36*a*), we have

$$V'_{be} = V_i \frac{R_{b'e}}{R_{bb'} + R_{b'e}} \quad \ldots(16)$$

The voltage is given by

$$V = -g_m V'_{be} Z$$

$$= -g_m V_i \frac{R_{b'e}}{R'_{bb} + R_{b'e}} \times \frac{R_{tt}}{1 + j\,2\delta\,Q_e} \quad \ldots(17)$$

The voltage gain at frequency of resonance ω_0 and at any other frequency ω close to ω_0 may be found as follows :

From equations (1) and (2), we have

$$I_2 = \frac{VZ_{21}}{Z_{21}Z_{12} - Z_{11}Z_{22}} \quad \ldots(18)$$

The output voltage V_0 is given by

$$V_0 = -I_2 R_i \quad \text{(see fig. 52b)} \quad \ldots(19)$$

From eqs. (18) and (19), we get

$$V_0 = -\frac{VZ_{21}}{Z_{21}Z_{12} - Z_{11}Z_{22}} \times R_i$$

$$= V R_i \frac{Z_{21}}{Z_{11}Z_{22} - Z_{12}^2} \quad \ldots(20)$$

Substituting the value of V from eq. (17) in eq. (20), we get

$$V_0 = -g_m V_i \frac{R_{b'e}}{R_{bb'} + R_{b'e}} \times \frac{R_{tt}}{1 + j\,2\delta\, Q_e} \times \frac{R_i Z_{21}}{Z_{11}Z_{22} - Z_{12}^2} \quad \ldots(21)$$

The voltage gain A at any frequency ω is given by

$$A = \frac{V_0}{V_i} = -g_m \frac{R_{b'e}}{R_{bb'} + R_{b'e}} \times \frac{R_{tt}}{1 + j\,2\delta\, Q_e} \times \frac{R_i Z_{21}}{Z_{11}Z_{22} - Z_{12}^2} \quad \ldots(22)$$

Voltage gain at resonance where $\omega = \omega_0$ and $\delta = 0$, we have

$$A_{res} = -g_m \frac{R_{b'e}}{R_{bb'} + R_{b'e}} \times \frac{R_{tt}}{1} \times \frac{R_i Z_{21}}{Z_{11}Z_{22} - Z_{12}^2} \quad \ldots(23)$$

Hence $$\frac{A}{A_{res.}} = \frac{1}{1 + j\,2\delta Q_e} \quad \ldots(24)$$

$3dB$ bandwidth is given by

$$Q_e = \frac{\omega_0}{B} \quad \ldots(25)$$

From eq. (12), $Q_e = \dfrac{R_{tt}}{\omega_0 L}$

Under condition of maximum transfer of power, the total resistance appearing in shunt with the coil equals $R_0/2$. Hence

$$Q_e = \frac{R_0/2}{\omega_0 L} = \frac{R_0}{2\omega_0 L} \quad \ldots(26)$$

From eqs. (25) and (26), we get

$$\frac{\omega_0}{B} = \frac{R_0}{2\omega_0 L} \quad \ldots(27)$$

Thus the value of L will be governed by the bandwidth B required for the amplifier.

16·29. DOUBLE TUNED TRANSISTOR AMPLIFIER :

Fig. (56) shows the circuit of a *CE* double tuned amplifier. The voltage developed across the tuned circuit in the collector circuit is inductively coupled to the another tuned circuit. The two circuits are tuned to the same frequency *i.e.*, the frequency of the signal.

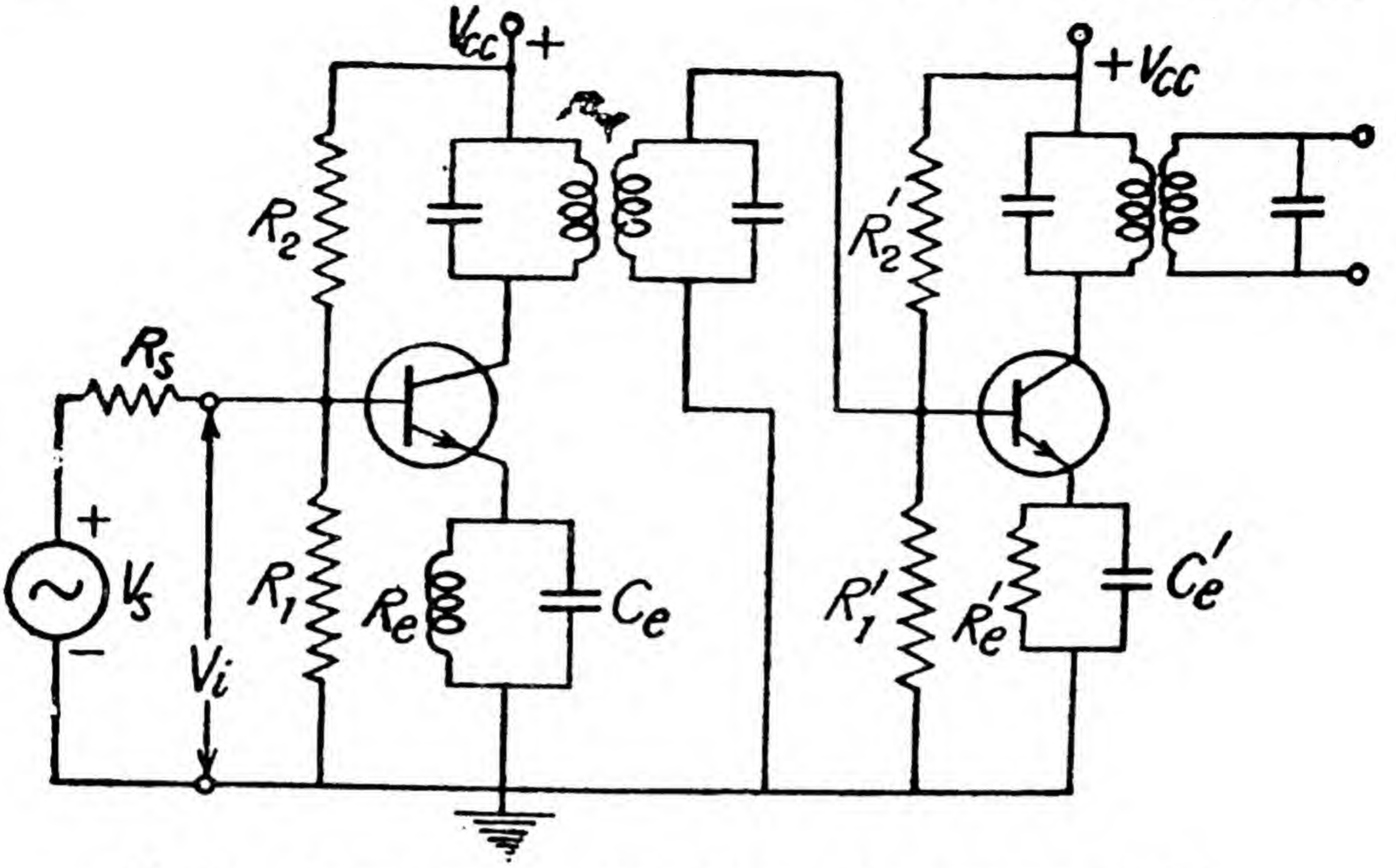


Fig. 56. Circuit diagram of double tuned transistor amplifier.

The equivalent output circuit of the amplifier is shown in fig. (57).

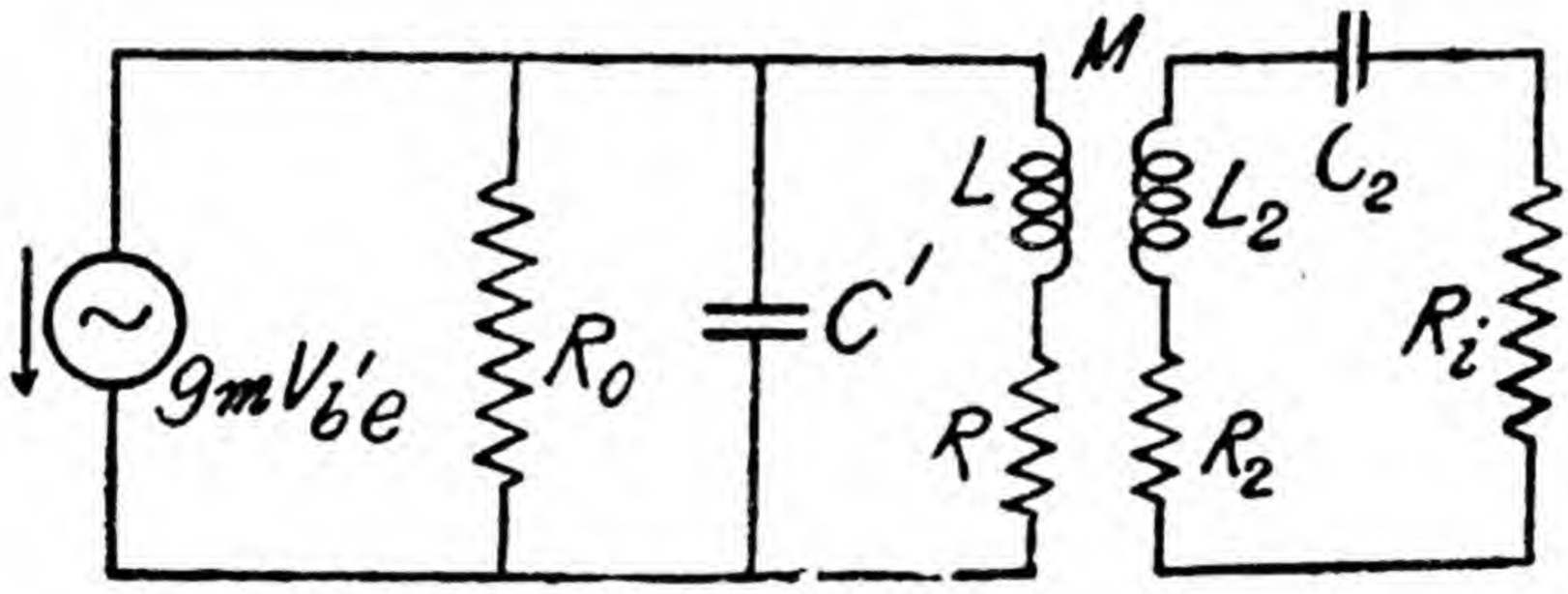


Fig. 57. Equivalent output circuit.

In the equivalent output circuit of fig. (57), R_i and R_2 may be combined to form R_i'. Similarly, R_0 in parallel with R may be brought in series with L to form R_0'. This modified equivalent circuit is shown in fig. (58).

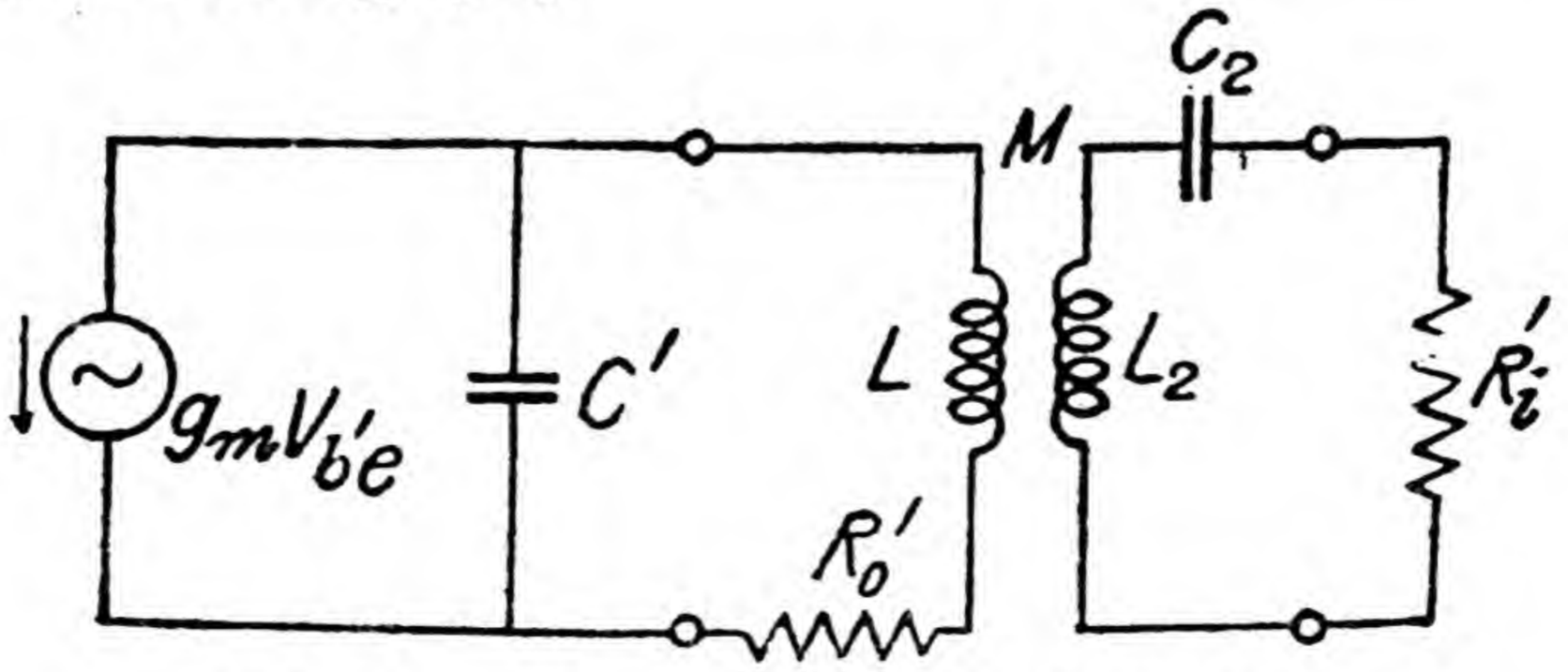


Fig. 58. Modified equivalent circuit.

Again the current source $g_m V_{b'e}$ in shunt with capacitance C' may be replaced by voltage generator V_1 in series with C'. In this way, the modified final equivalent circuit is shown in fig. (59).

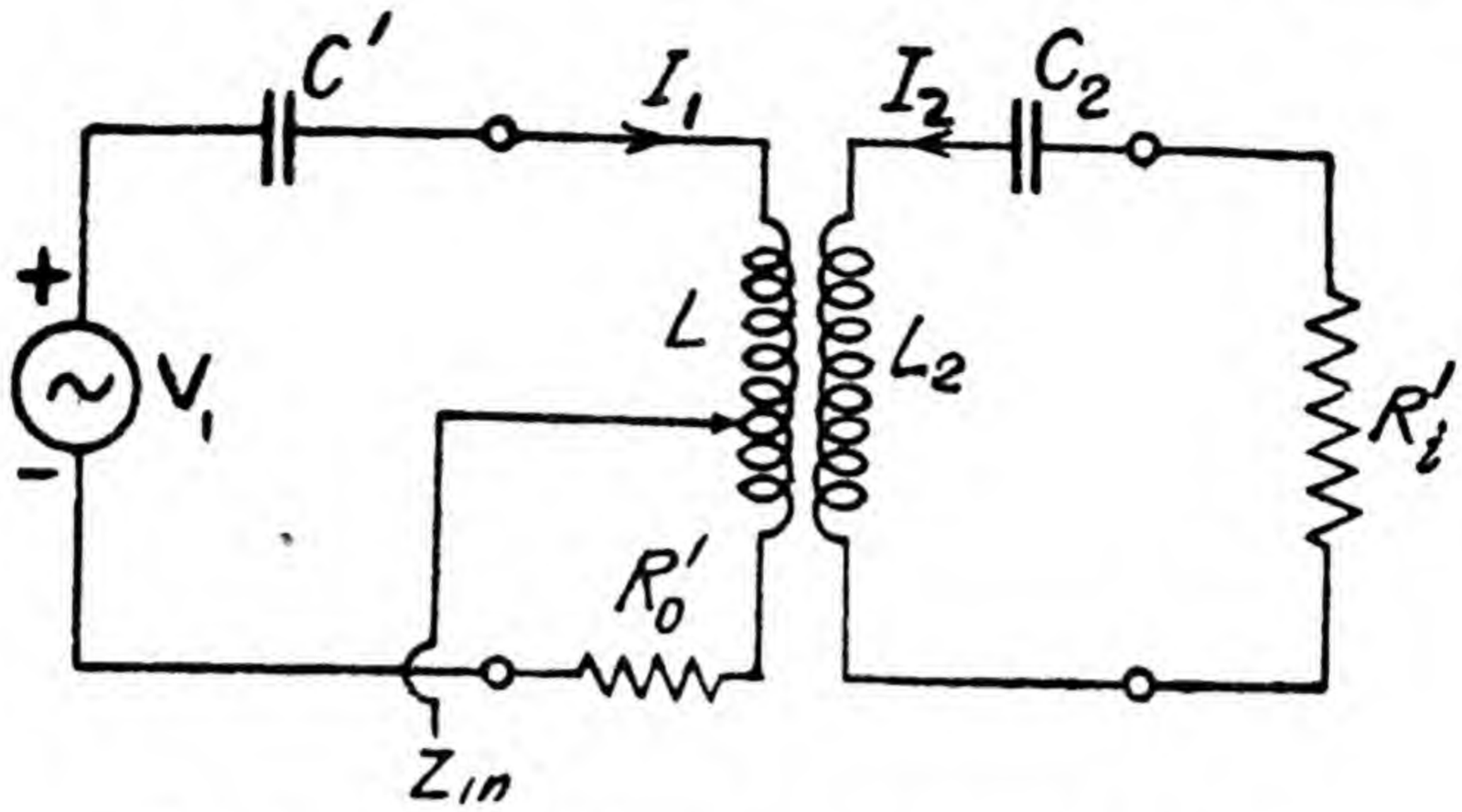


Fig. 59. Final modified equivalent circuit

The impedance Z_{in} seen by the primary in fig. (59) is given by

$$Z_{in}=\frac{\omega^2 M^2}{R_t'+j\left(\omega L_2-\frac{1}{\omega C_2}\right)} \qquad ...(1)$$

At resonance $\omega_0 L=\frac{1}{\omega_0 C}$ and $\omega_0 L_2=\frac{1}{\omega C_2}$. ...(2)

Hence at resonance, Z_{in} reduces to

$$Z_{in}=\frac{\omega_0^2 M^2}{R_t'} \qquad ...(3)$$

Hence at resonance, for maximum transfer of power R_0' must equal Z_t *i.e.*, is adjusted to critical value M_c such that

$$R_0'=\frac{\omega_0^2 M_c^2}{R_t'}$$

or $$R_0'\,R_t'=\omega_0^2 M_c^2 \text{ or } \sqrt{(R_0' R_t')}=\omega_0 M_c \qquad ...(4)$$

If K_c be the critical value of the cofficient of coupling corresponding to the critical value M_c of mutual inductance, then

$$\sqrt{(R_0' R_t')}=\omega_0 K_c\ \sqrt{(LL_2)}$$

$$K_c=\frac{\sqrt{(R_0' R_t')}}{\omega_0\sqrt{(LL_2)}}=\left(\frac{R_0'}{\omega_0 L}\right)^{1/2}\times\left(\frac{R_t'}{\omega_0 L_2}\right)^{1/2}$$

$$\therefore \quad K_c=\frac{1}{\sqrt{(Q_1 Q_2)}} \qquad ...(5)$$

From fig (59), we have

$$V_1=Z_{11}\,I_1+Z_{12}\,I_2 \qquad ...(6)$$

$$0=Z_{21}\,I_1+Z_{22}\,I_2 \qquad ...(7)$$

where $$Z_{11}=R_0'+j\left(\omega L-\frac{1}{\omega C}\right) \qquad ...(8)$$

$$Z_{22}=R_i'+j\left(\omega L_2-\frac{1}{\omega C_2}\right) \qquad ...(9)$$

$$Z_{12}=Z_{21}=j\omega_0 M \qquad ...(10)$$

At resonance $Z_{11}=R_0'$ and $Z_{22}=R_i'$

For maximum transfer of power at resonance

$$Z_{11}=R_0',\ Z_{22}=R_i' \text{ and } Z_{12}=j\omega_0 M_c \qquad ...(11)$$

From equations (6) and (7), we have

$$I_2=-\frac{V_1 Z_{21}}{Z_{11}Z_{22}-Z_{12}^2} \qquad ...(12)$$

Substituting the values of Z_{21}, Z_{11} and Z_{22} from eq. (11) in eq. (12) we have

$$(I_2)_{max}=-\frac{j\,V_1\omega_0 M_c}{R_0'R_i'+\omega_0^2 M_c^2} \qquad ...(13)$$

Substituting the value of M_c from eq. (4) in eq. (13) we get

$$(I_2)_{max}=-j\,\frac{V_1\sqrt{(R_0'R_i')}}{R_0'R_i'+R_0'R_i'}=-j\,\frac{V_1}{2\sqrt{(R_0'R_i')}} \qquad ...(14)$$

The magnitude of $(I_2)_{max}$ is given by

$$|\,(I_2)_{max}\,|=\frac{V_1}{2\sqrt{(R_0'R_i')}} \qquad ...(15)$$

Power transfer is maximum at two other frequencies. These two frequencies may be obtained by equating the value of $(I_2)_{max}$ as given by eq. (14) to the value of I_2 given by eq. (12) where Z_{11}, Z_{22} pertain to any frequency ω close to ω_0. Thus we get

$$\left|\frac{-j\,V_1}{2\sqrt{(R_0'R_i')}}\right|=\left|\frac{-jV_1\sqrt{(R_0'R_i')}}{\left(R_0'+j\omega L+\frac{1}{j\omega C'}\right)\left(R_i'+j\omega L_2+\frac{1}{j\omega C_2}\right)+b^2R_0'R_i'}\right| \qquad ...(16)$$

where the coefficient of coupling

$$K=bK_c$$

Let $L_2=L,\ C_2=C'=C$ and $R_0'=R_i'=R$

and $j\left(\omega L-\frac{1}{\omega C}\right)=jX$ } ...(17)

Now eq. (16) may be put as

$$\left|\frac{V_1}{2R}\right|=\frac{V_1 bR}{(R+jX)^2+b^2R^2} \qquad ...(18)$$

$$|\,2bR^2\,|=|-X^2+2j\,RX+R^2+b^2R^2\,|$$

or $$4b^2R^4=[R^2\,(1+b^2)-X^2]^2+4R^2X^2 \qquad ...(19)$$

Solving eq. (19) for X, we get

$$X=\pm R\sqrt{(b^2-1)}$$

or $$\left(\omega L-\frac{1}{\omega C}\right)=\pm R\sqrt{(b^2-1)}$$

or $$(\omega^2 LC-1)=\pm\omega CR\,[\sqrt{(b^2-1)}] \qquad \ldots(20)$$

Now $\omega CR \simeq \omega_0 CR=\frac{1}{Q}$ ($\because$ Q is large)

also $$\omega_0^2=\frac{1}{LC}$$

Equation (20) yields

$$\frac{\omega^2}{\omega_0^2}=1\pm\frac{\sqrt{(b^2-1)}}{Q}$$

Hence $$\omega\simeq\pm\omega_0\sqrt{\left(1\pm\frac{\sqrt{(b^2-1)}}{Q}\right)} \qquad \ldots(21)$$

It is clear from eq. (21) that if $b > 1$, there are two frequencies at which maximum power transfer takes place.

The $3dB$ frequency, *i.e.*, the frequency at which current I_2 reduces to 0·707 of its maximum value are obtained by equating I_2 as given by eq. (18) to $\frac{1}{\sqrt{2}}$ times $(I_2)_{max}$ as given by eq. (14).

Thus

$$\left|\frac{-jV_1 b\sqrt{(R_0'\,R_i')}}{\left(R_0'+j\omega L+\frac{1}{j\omega C}\right)\left(R_i'+j\omega L_2+\frac{1}{j\omega C_2}\right)+b^2R_0'R_i'}\right|$$

$$=\frac{1}{\sqrt{2}}\left|\frac{-jV_1}{2\sqrt{(R_0'R_i')}}\right| \qquad \ldots(22)$$

Using condition (17) and solving for X, we have

$$X=\pm R\sqrt{(b^2-1\pm 2b)}$$

or $$\omega L-\frac{1}{\omega C}=\pm R\sqrt{(b^2-1\pm 2b)}$$

$$\omega\omega_0 LC-\frac{\omega_0}{\omega}=\pm\omega_0 CR\sqrt{(b^2-1\pm 2b)} \qquad \ldots(23)$$

But $\omega_0 CR=Q$ and $\omega_0=\frac{1}{\sqrt{(LC)}}$

$\therefore$ $$\frac{\omega}{\omega_0}-\frac{\omega_0}{\omega}=\pm\frac{\sqrt{(b^2-1\pm 2b)}}{Q} \qquad \ldots(24)$$

Equation (24) gives two $3dB$ frequencies ω_2 and ω_1, one corresponding to positive sign and other corresponding to negative sign. Thus

$$\left(\frac{\omega_2}{\omega_0}-\frac{\omega_0}{\omega_2}\right)=-\left(\frac{\omega_1}{\omega_0}-\frac{\omega_0}{\omega_1}\right) \qquad \ldots(25)$$

This yields $$\omega_0=\sqrt{(\omega_1\omega_2)} \qquad \ldots(26)$$

Now taking positive sign in eq. (24), we have

$$\frac{\omega_2}{\omega_0}-\frac{\omega_0}{\omega_2}=\frac{\sqrt{\{(b^2-1)+2b\}}}{Q}$$

or $$\frac{\omega_2^2-\omega_0^2}{\omega_0\omega_2}=\frac{\sqrt{\{(b^2-1)+2b\}}}{Q}$$

or $$\frac{\omega_2^2-\omega_1\omega_2}{\omega_0\omega_2}=\frac{\sqrt{\{(b^2-1)+2b\}}}{Q}$$

or $$\frac{\omega_2-\omega_1}{\omega_0}=\frac{\sqrt{\{(b^2-1)+2b\}}}{Q}$$

Thus the $3dB$ bandwidth is given by

$$\omega_2-\omega_1=\frac{\omega_0}{Q}\sqrt{\{(b^2-1)+2b\}} \qquad ...(27)$$

This shows that $3dB$ bandwidth $(\omega_2-\omega_1)$ is proportional to $\frac{\omega_0}{Q}$. In case of a single tuned amplifier, the $3dB$ bandwidth is equal to ω_0/Q. This shows that $3dB$ bandwidth in case of double tuned amplifier exceeds that in a single tuned amplifier by a factor $\sqrt{\{(b^2-1)\pm 2b\}}$.

NUMERICALS

1. *Obtain a correlation between r and h parameters.*

The r and h parametric equations for the same two part network are

$$\left.\begin{aligned} v_1&=r_{11}\,i_1+r_{12}\,i_2\\ v_2&=r_{21}\,i_1+r_{22}\,i_2\end{aligned}\right\} \qquad ...(1)$$

and

$$\left.\begin{aligned} i_1&=h_{11}\,i_1+h_{12}\,v_2\\ i_2&=h_{21}\,i_1+h_{22}\,v_2\end{aligned}\right\} \qquad ...(2)$$

From equations (2)

$$-h_{11}\,i_1=-v_1+h_{12}\,v_2$$
$$i_2-h_{21}\,i_1=0.v_1+h_{22}\,v_2.$$

Hence $$v_1=\begin{vmatrix} -h_{11}\,i_1 & h_{12}\\ i_2-h_{21}\,i_1 & h_{22}\\ \hline -1 & h_{12}\\ 0 & h_{22}\end{vmatrix}=\frac{-h_{11}\,h_{22}\,i_1-h_{12}\,i_2+h_{12}\,h_{21}\,i_1}{-h_{22}}$$

or $$v_1=\frac{(h_{11}\,h_{22}-h_{12}\,h_{21})\,i_1+h_{12}\,i_2}{h_{22}}$$

Similarly $$v_2=\begin{vmatrix} -1 & h_{11} & i_1\\ 0 & i_2-h_{21} & i_1\\ \hline -1 & h_{12} & \\ 0 & h_{22} & \end{vmatrix}=\frac{-i_2+h_{21}\,i_1}{-h_{22}}$$

or $$v_2=-\frac{h_{21}\,i_1+i_2}{h_{22}}.$$

$$\therefore \quad r_{11}=\frac{h_{11}\,h_{22}-h_{12}\,h_{21}}{h_{22}}$$
$$r_{12}=\frac{h_{12}}{h_{22}}$$
$$r_{21}=-\frac{h_{21}}{h_{22}} \qquad \ldots (3)$$
$$r_{22}=\frac{1}{h_{22}}.$$

2. *The T equivalent circuit parameters of a transistor are :* $r_e=12{\cdot}5\ \Omega$: $r_b=550\ \Omega$: $r_c=220\times10^3\ \Omega$ *and* $\alpha=0{\cdot}95$. *Calculate input impedance, output impedance, current gain, voltage gain, and power gain in common base amplifier using this transistor. Given that load resistance* $R_L=10{,}000\ \Omega$ *and source resistance* $R_S=500\ \Omega$.

We know that $r_m=\alpha\ r_c-0{\cdot}95\times220\times10^5$
$=209\times10^3\ \Omega.$

(1) Input resistance $r_i=r_e+\dfrac{r_b\,(r_c-r_m+R_L)}{r_e+r_b+R_L}$

$=12{\cdot}5+\dfrac{550\,(220\times10^3-209\times10^3+10{,}000)}{220\times10^3+550+10{,}000}$

$=62{\cdot}5\ \Omega.$

(2) Output resistance $r_0=r_c+r_b-\dfrac{r_b\,(r_b+r_m)}{r_b+r_e+R_S}$

$=220\times10^3+550-\dfrac{550\,(550+209\times10^3)}{550+12{\cdot}5+500}$

$=112{\cdot}3\times10^3\ \Omega.$

(3) Current gain $A_i=\dfrac{r_b+r_m}{r_c+r_b+R_L}$

$=-\dfrac{(500+209\times10^3)}{220\times10^3+550+10{,}000}$

$=-0{\cdot}905.$

(4) Voltage gain $A_v=A_i\dfrac{R_L}{r_i}$

$=0{\cdot}905\times\dfrac{10{,}000}{62{\cdot}5}=145$

(5) Power gain $=A_i\times A_v=0{\cdot}905\times145=131.$

3. *In class B push pull amplifier battery supply is 5 volts, output power is 400 mW and maximum collector dissipation of each trasistor is 100 mW. Calculate collector to collector load and peak collector.*

Given $V_{max}=V_{cc}=5$ volts

$V=400\ \text{m}_1W=400\times10^{-3}$ watts.

The collector current is given by

$$P = \frac{I_{max.}}{\sqrt{2}} \cdot \frac{V_{max.}}{\sqrt{2}}$$

or $$400.10^{-3} = \frac{I_{max.}}{2} \times 5.$$

$$\therefore \quad I_{max.} = \mathbf{160\ mA}$$

and collector load can be obtained from the relation

$$R_L = 4. \frac{V_{max.}}{I_{max.}} = 4. \frac{5}{160 \cdot 10^{-3}} = \mathbf{125\ ohms.}$$

EXPERIMENTS

Exp. No. 1. To draw the characteristic curves of

(i) PN junction diode.

(ii) PNP junction transistor.

Procedure : (*i*) For the characteristics of *PN* junction diode, procedure has been described in Chapter 6.

(*ii*) *PNP junction Transistor Characteristics* : *PNP* junction transistor can be used in three ways as regards its method of connection. We shall draw here the characteristic curves for common base and common emitter connections.

(*a*) *Common base circuit* : The circuit is shown in fig 60.

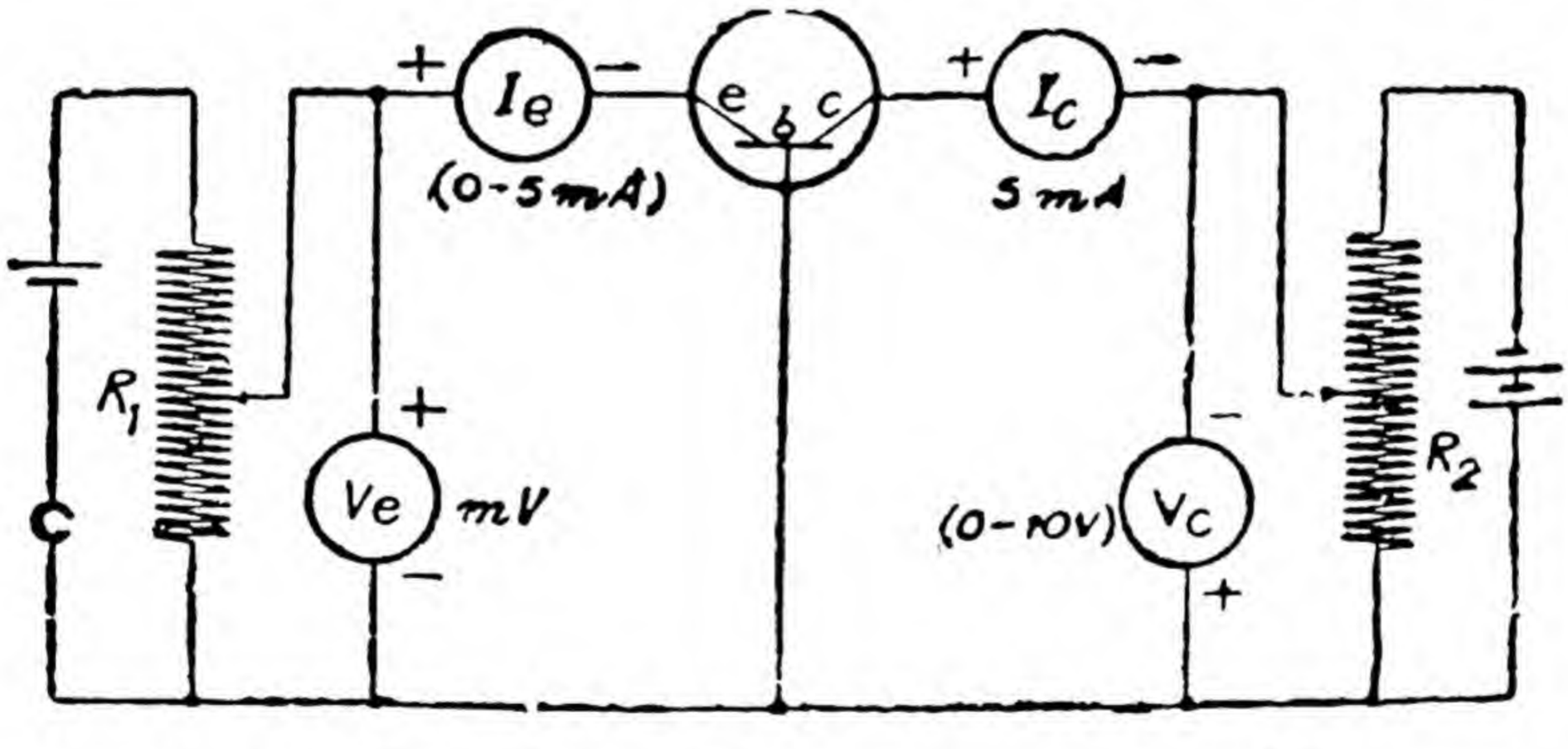


Fig. 60. Circuit for common base characteristics.

The method of obtaining the two common characteristics are described as follows :

(a_1) *Plot of collector current/collector voltage characteristics at different values of emitter current.* The emitter current is adjusted to 1 mA and maintained constant by adjusting variable resistor R_1. The collector voltage is adjusted to −5V, by means of resistor R_2 and the collector current is noted. The collector voltage is reduced in steps of 1 volt and with emitter current constant, simultaneous reading of collector voltage and current are taken.

Now the polarity of collector voltage is reversed and adjusted to 0·1 volt and collector current is noted. The collector voltage is increased in steps of 0·1 volt until the collector current reaches zero. Voltage and current readings are noted at each step. The test is repeated for constant emitter currents of 2, 3 and 4 mA. and a family of collector current/collector voltage curves are plotted as shown in fig. 61.

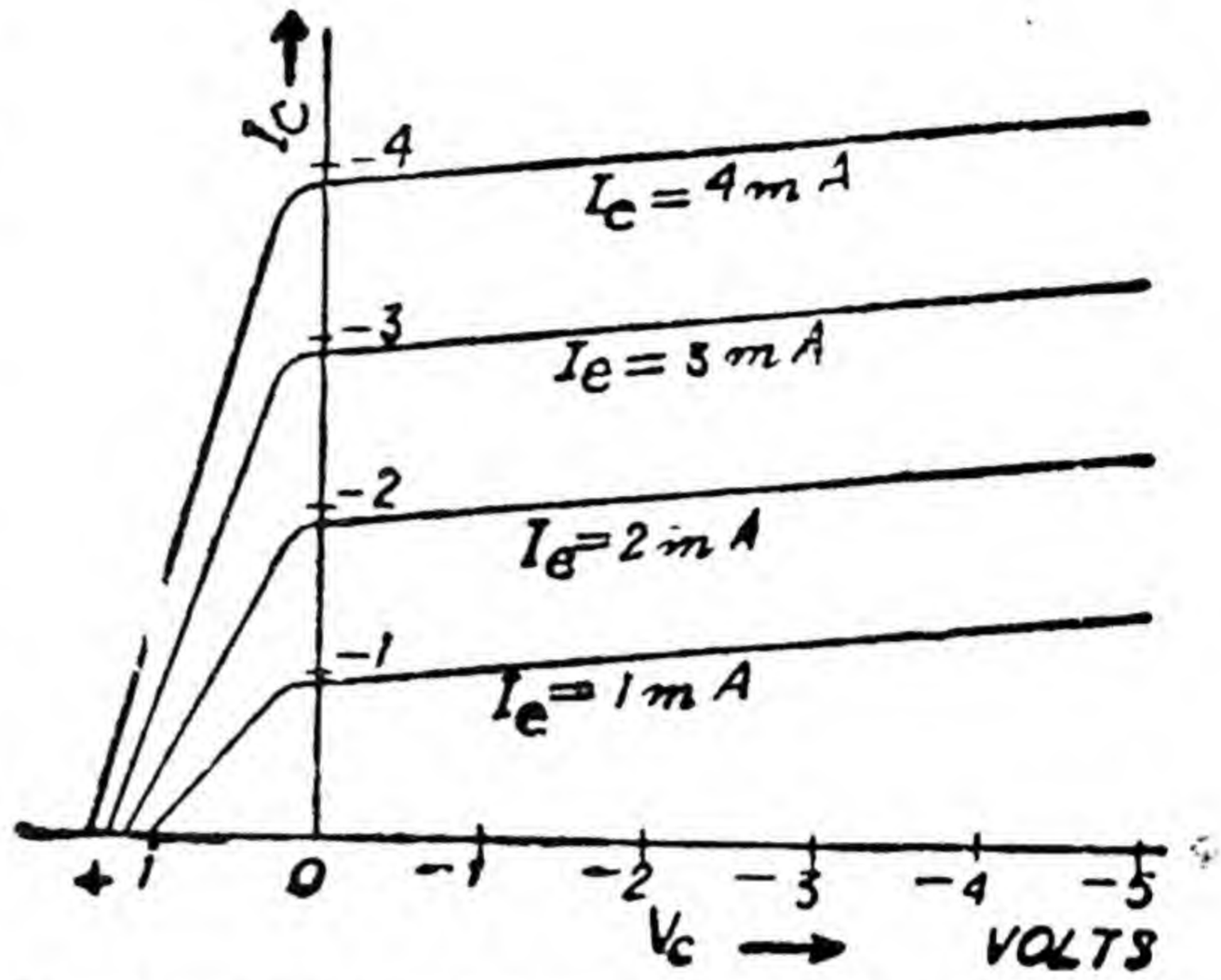


Fig. 61. Collector current/collector voltage characteristics.

(a_2) *Plot of collector current/emitter current characteristics* : The collector voltage is set to —5 volts and is kept constant throughout this test. The emitter current is adjusted to 100μA, and collector current is noted. By adjusting R_1 and R_2, the simultaneous ammeter readings are taken upto a maximum of 4 mA. of collector current. A graph is plotted in collector current to a base of emitter current as shown in figure 62.

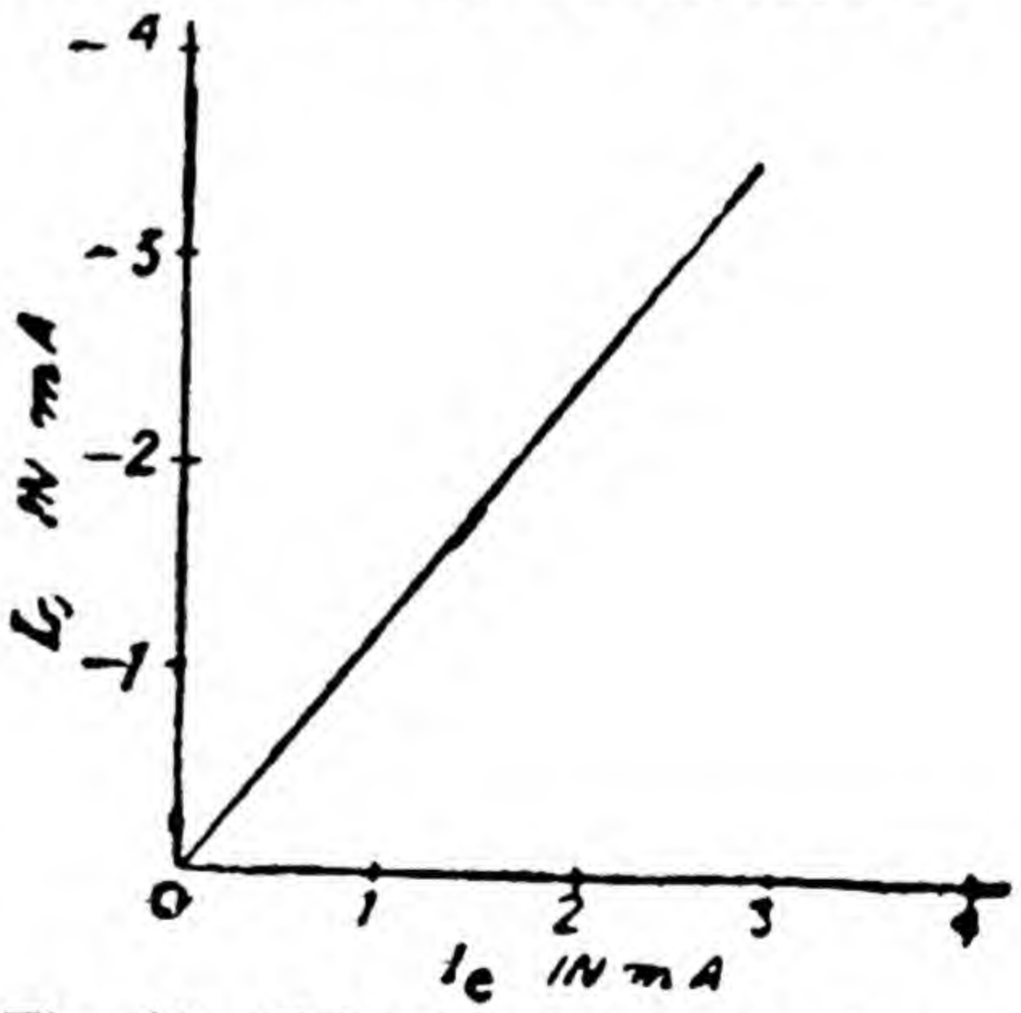


Fig. 62. Collector current/emitter current characteristics.

(b) *Grounded Emitter Circuit* : The circuit is shown in fig. 63.

The methods of obtaining the two general characteristics are given below ;

(b_1) *Plot of collector current/base current characteristics* : By means of R_2 the collector voltage ie adjusted to —2V which

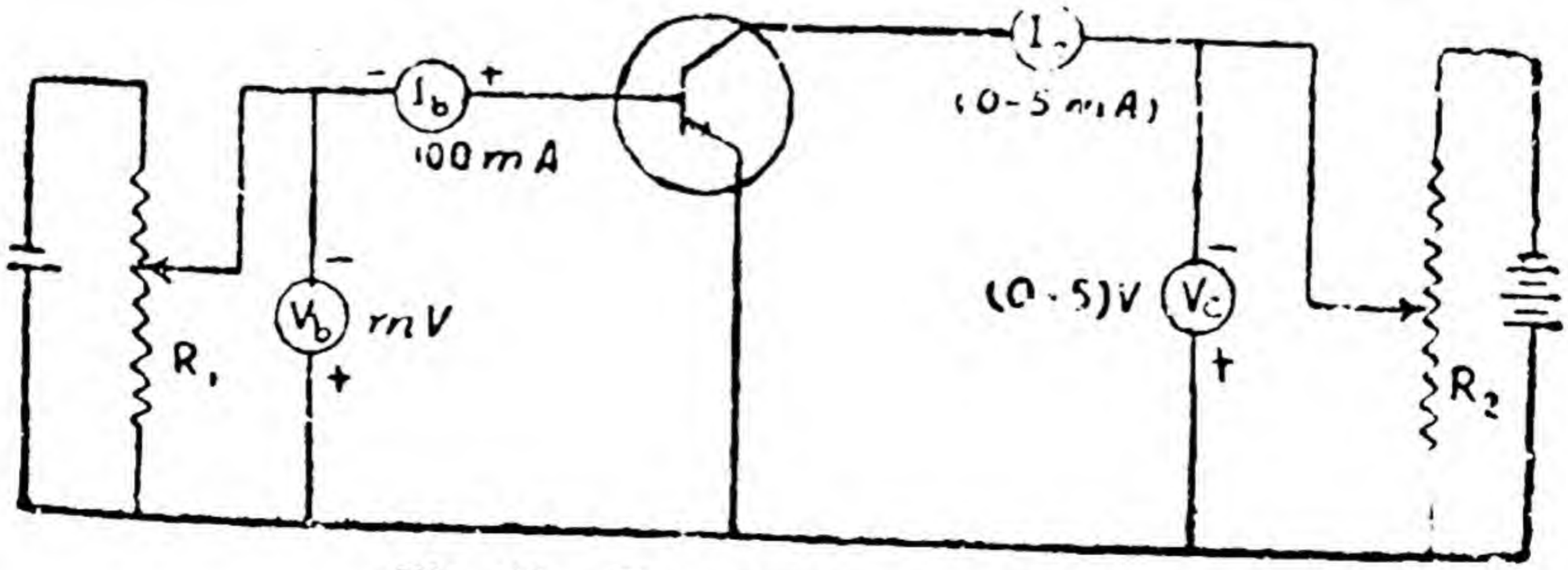


Fig. 63. Common emitter circuit.

is kept constant during this test. The base current is adjusted to $20\mu A$, by means of R_1 and collector current is noted. The readings are repeated for increasing values of base current and collector current is noted until the latter reaches about 4 mA.

Curve is drawn in collector current against the base current as shown in fig. 64.

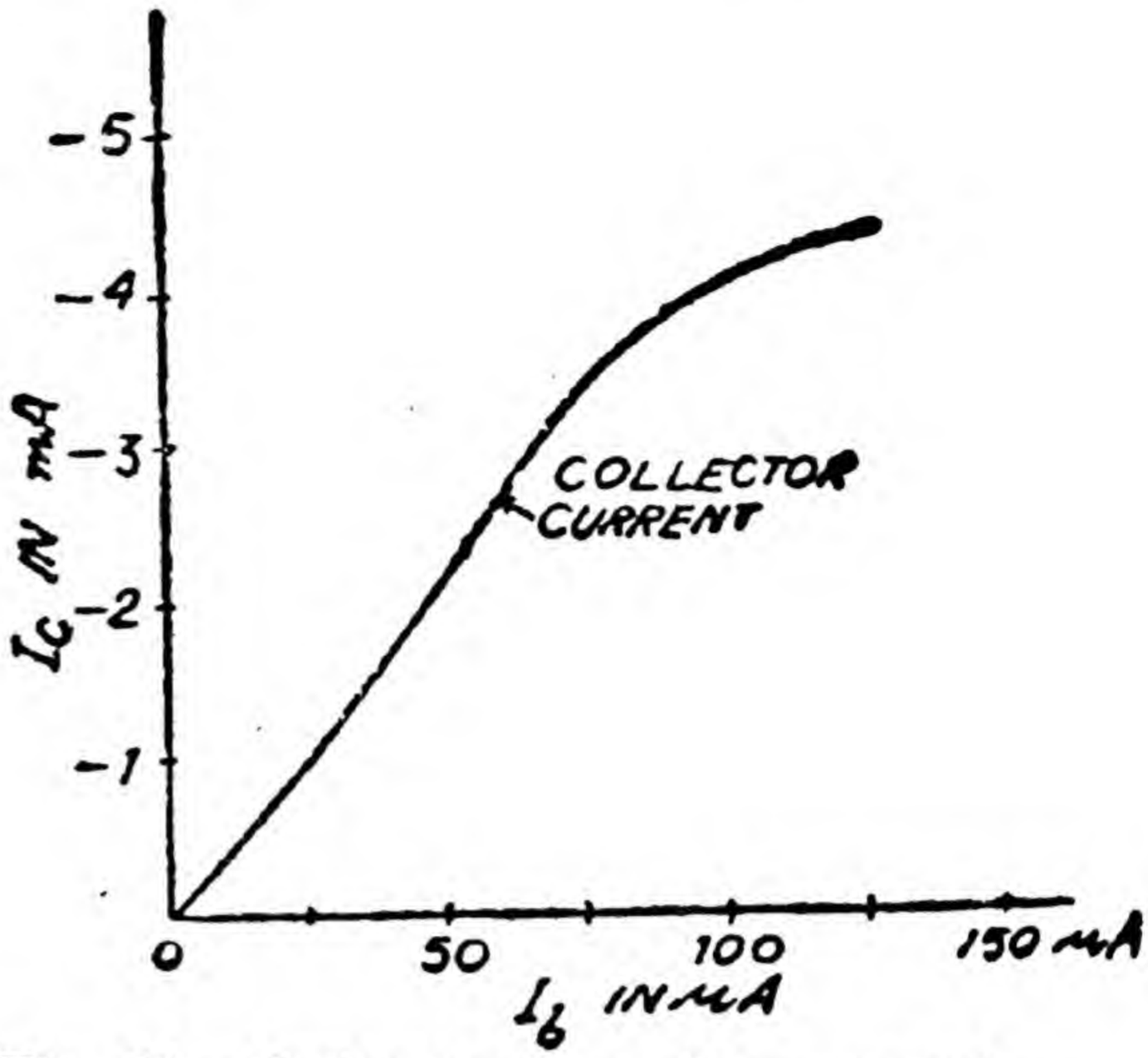


Fig. 64. Collector current/base current characteristics.

(b_2) *Plot of collector current/collector voltage characteristics for various base currents* : The circuit is same. Base current is adjusted to $20\mu A$ by means of R_1. The collector voltage is increased in steps by 0·1V by R_2 and collector current at each step is

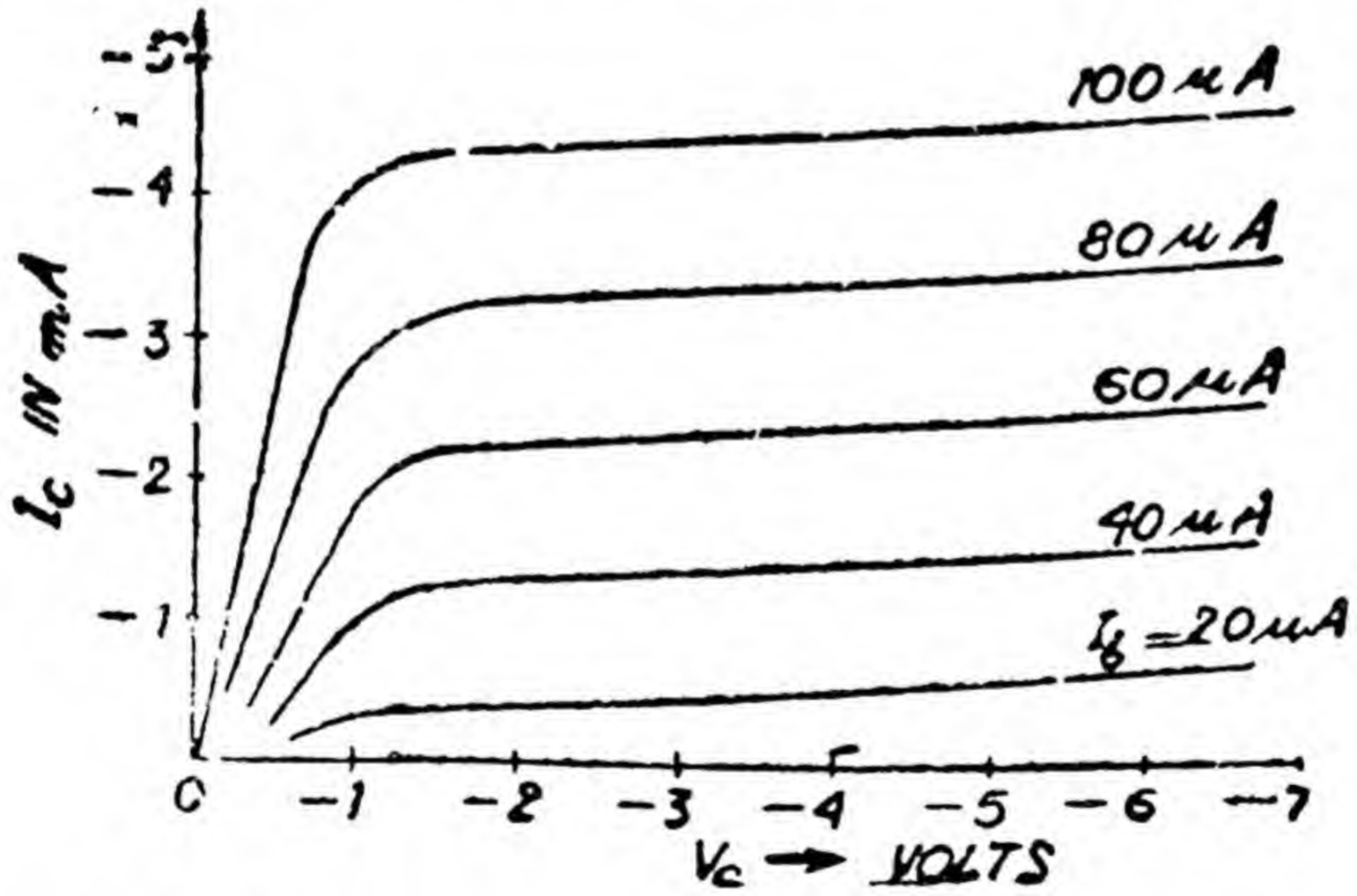


Fig. 65. Collector current/collector voltage characteristics.

noted. This is repeated for constant base currents of 40, 60, 100μA, and a family of characteristics is plotted as shown in fig. 65.

Expt. No. 2 : To draw frequency response curve of a common emitter-amplifier circuit and to determine :

(i) Input Impedance Z_i,
(ii) Output impedance Z_o,
(iii) Power gain.

Procedure : Circuit for common emitter-amplifier is shown in fig. 66.

The input is provided by means of a variable frequency oscillator. Voltage E_b and resistor R_1 are set to provide the correct d.c. potential. Capacitors applied are of high values. The oscillator voltage, which is maintained constant, is set to a suitable level with switch S_2 open and S_1 to the right. The oscillator

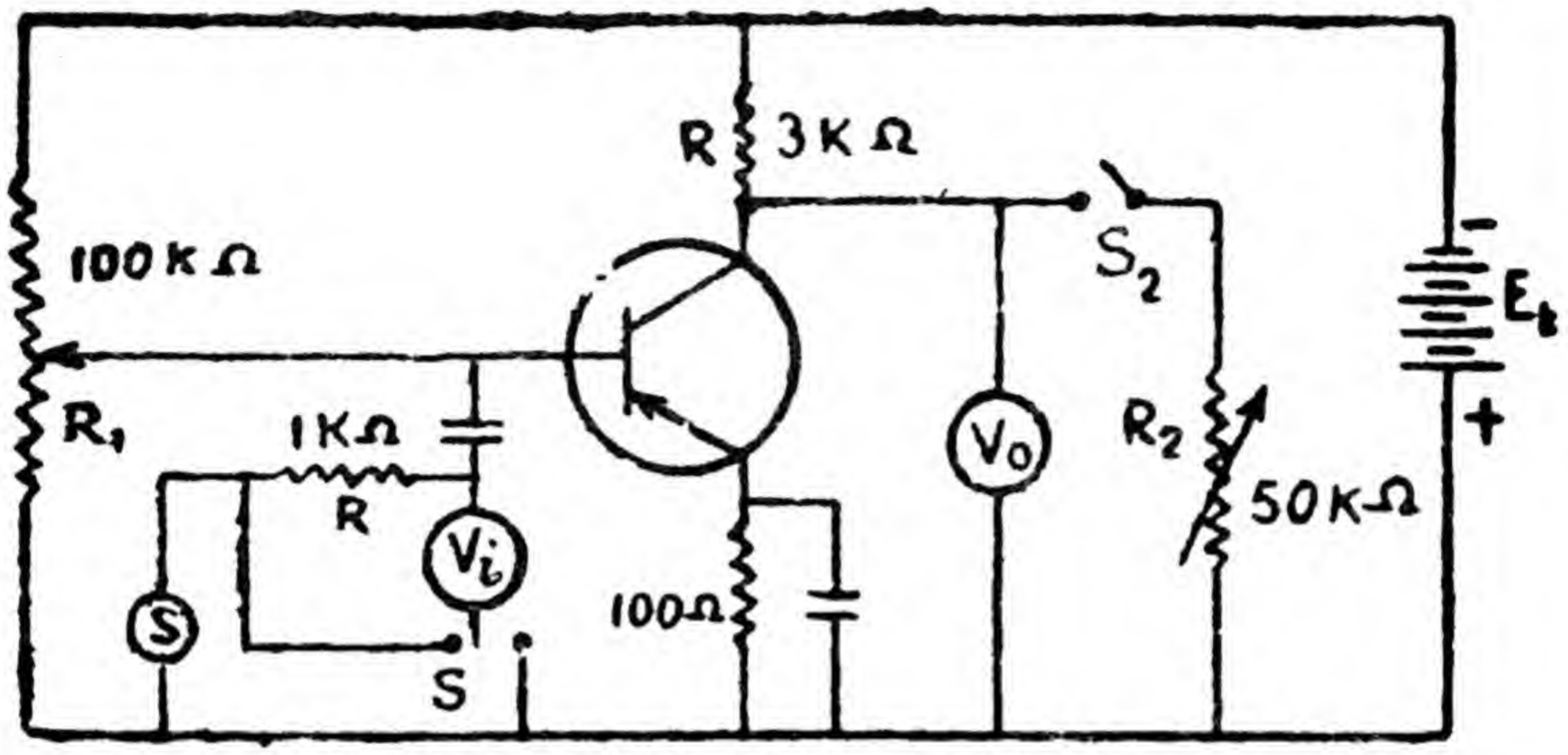


Fig. 66. Common emitter-amplifier circuit.

frequency is varied from low audio value upto 1 mc/s. and the reading of V_0 is noted at each frequency. Voltage gain of the stage is measured by the ratio of $\frac{v_0}{v_1}$. A curve of voltage gain to a logarithmic base of frequency is plotted, called the *frequency response curve of this type of amplifier* circuit. Now

(*i*) To measure the input impedance Z_i at given frequency, the input voltage v_i is divided by the reading v_r obtained when S_1 is switched to left. Ratio $\frac{v_i}{v_r}$ then gives the input impedance Z_i.

(*ii*) In order to measure the output impedance Z_0 for a given condition, we note the reading of V_0 for this condition and then with S_2 closed resistance R_2 is varied to have this reading of V_0. The resistance R_2 is calibrated to give its value directly.

(*ii*) From relation

$$\text{power gain} = (\text{voltage})^2 . \frac{Z_i}{Z_0}.$$

We can thus calculate power gain.

EXERCISES AND PROBLEMS

1. Define the terms (*a*) conductor, (*b*) insulator, (*c*) semi-conductor, Discuss what is meant by *P* and *N* type semi-conductors. Explain the movement of electrons in semi-conductor.
2. Explain the effect on a semi-conductor of adding (*a*) a donor impurity, (*b*) an acceptor impurity.
3. Describe the characteristics of a *PN* junction. Explain the meaning of (*a*) reverse bias, (*b*) potential hill, (*c*) reverse current.
4. Describe the static characteristics of transistor for (*a*) collector current versus collector voltage for a constant value of emitter current, (*b*) collector current versus collector voltage for a constant value of base current.

5. Describe the following characteristics of a grounded emitter transistor-amplifier circuit : (*a*) current gain, (*b*) voltage gain, (*c*) power gain.
6. Describe in detail the operation of *PNP* grounded emitter transistor amplifier circuit.
7. Describe in detail the operation of *PNP* of grounded base transistor amplifier circuit.
8. Compare the base operations of a transistor-amplifier circuit with the basic operations of a vacuum-tube amplifier circuit.
9. Explain why the common-emitter configuration is preferred for a transistor.
10. Compare a common emitter transistor and a common base transistor amplifier circuits. Explain why the former is widely used in amplifier circuit.
11. Explain the action of a *PN* junction rectifier.
Derive expression for the input and output resistance, current amplification and voltage gain of *PNP* grounded base transistor amplifier: Why is grounded emitter circuit preferred over the grounded base circuit ?
12. For a *PNP* transistor, the low frequency equivalent *T* net-work constant are

$r_e = 20$ ohms ; $r_b = 730$ ohms ; $r_c = 10^6$ ohms ; $\alpha = 0{\cdot}97$.

It is used as an amplifier with a source having an impedance $R_S = 500$ ohms. The load resistance $R_L = 300$ ohms. Find :

(*a*) voltage gain when used as a common amplifier

(*b*) output impedance when used as common emitter amplifier.

13. A certain *PNP* transistor (junction type) when used as a common emitter has the following circuit parameters.

$r_e = 21$ ohms, $r_b = 580$ ohm, $r_c = 3\ 44$ megohms,

$R_L = 20{,}000$ ohms, $\alpha = 0{\cdot}982$

Find the input resistance.

17 FIELD EFFECT TRANSISTOR AMPLIFIERS

17·1. BIASING THE FET :

Like a bipolar transistor or a vacuum tube, the FET or MOSFET must be properly biased before it is put to use as an amplifier. The consideration must be given in biasing the FET to place its operating voltage and current within the linear portion of its active region. The factors governing the selection of an appropriate operating point (V_{DS}, I_D, V_{GS}) for a FET are similar to those for tubes and transistors. So before discussing a particular FET amplifier, we shall first discuss the following biasing arrangements :

(1) Self Bias : In case of a FET, a reverse bias voltage is to be applied across the gate to source junction. So a simple self-biasing technique used for vacuum tubes may also be used in FET. Fig. (1) shows the self biasing. The conventional current flowing down through R_S produces a source to ground voltage given by

$$V_S = I_D R_S$$

Since negligible gate current flows, hence the gate terminal is at d.c. ground *i.e.*, $V_G = 0$.

So $$V_{GS} = -V_S = -I_D R_S$$

Thus biasing voltage V_{GS} is set up entirely by IR drop across R_S. Due to this fact the circuit is called self-biased. For a.c. signals, a capacitor having a large value is connected in parallel to R_S to avoid any degenerative feed back.

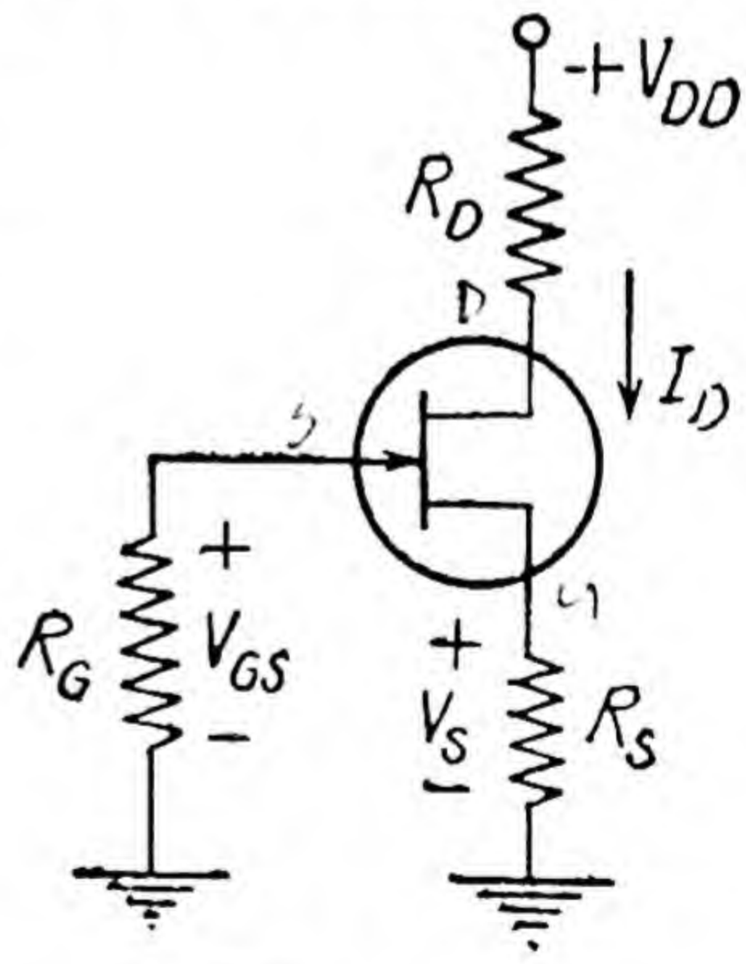


Fig. 1.

(2) Voltage divider Bias : Fig. (2) shows the voltage divider bias. In this case, the voltage developed across R_2 appears as the voltage between gate and source *i.e.*, V_G

$$\therefore \qquad V_G = \frac{R_2 . V_{DD}}{R_1 + R_2} \text{ and } V_{GS} = V_G - I_D R_S$$

The voltage divider bias is only used in JFETs when large supply voltages are available. With low voltage supplies, this bias is not preferred.

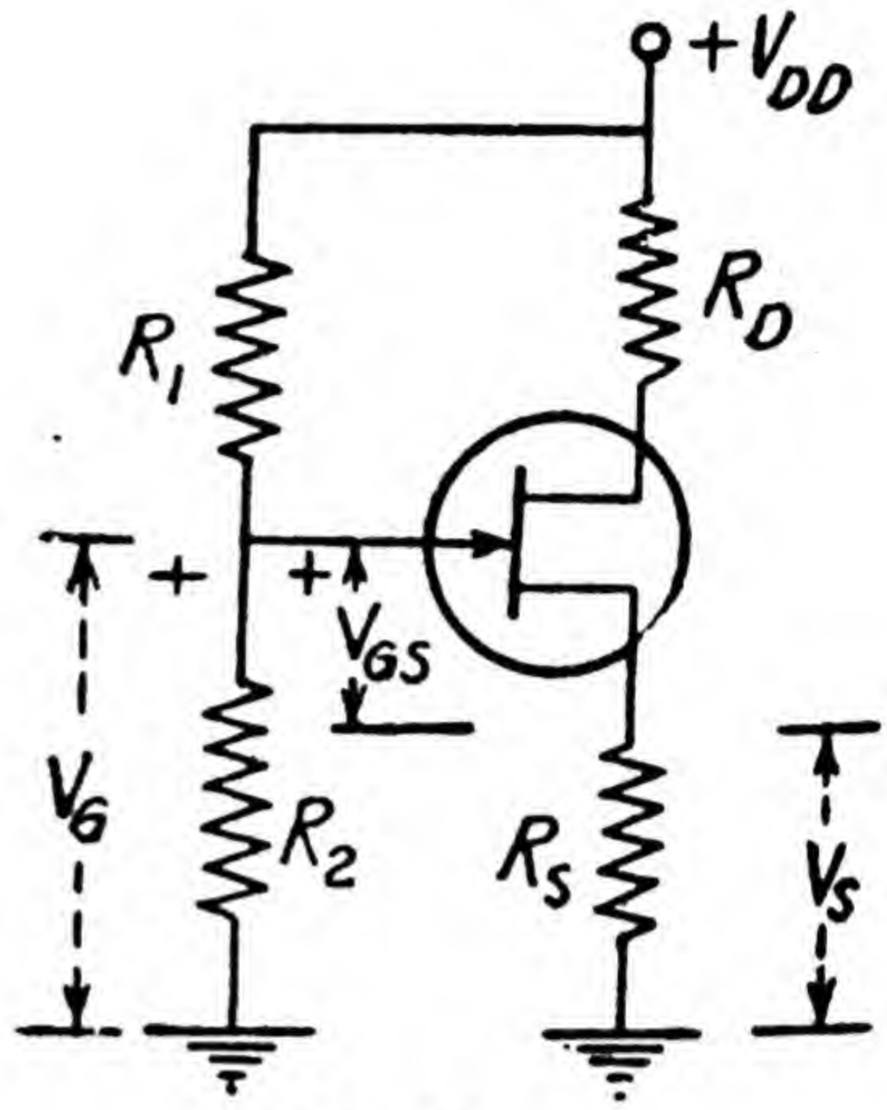


Fig. 2.

(3) Source Bias : The source bias is shown in fig. (3). Since most V_{SS} appears across R_S, and hence the drain current I_D is almost independent of FET characteristics as shown below :

Summing voltages around the gate, we have

$$I_G R_G + V_S + I_D R_{GS} - V_{SS} = 0$$

$$\therefore \quad I_D = \frac{V_{SS} - V_{GS} - I_G R_G}{R_S}$$

But

$$V_{SS} - V_{GS} \gg I_G R_G$$

(for a practical circuit)

$$\therefore \quad I_D = \frac{V_{SS} - V_{GS}}{R_S}$$

When V_{SS} is much greater than V_{GS}, then

$$I_D \approx \frac{V_{SS}}{R_S} \text{ (ideal)}$$

In an ideal circuit V_{SS} appears across R_S. As these quantities are independent of JFET characteristics, so I_D is almost constant.

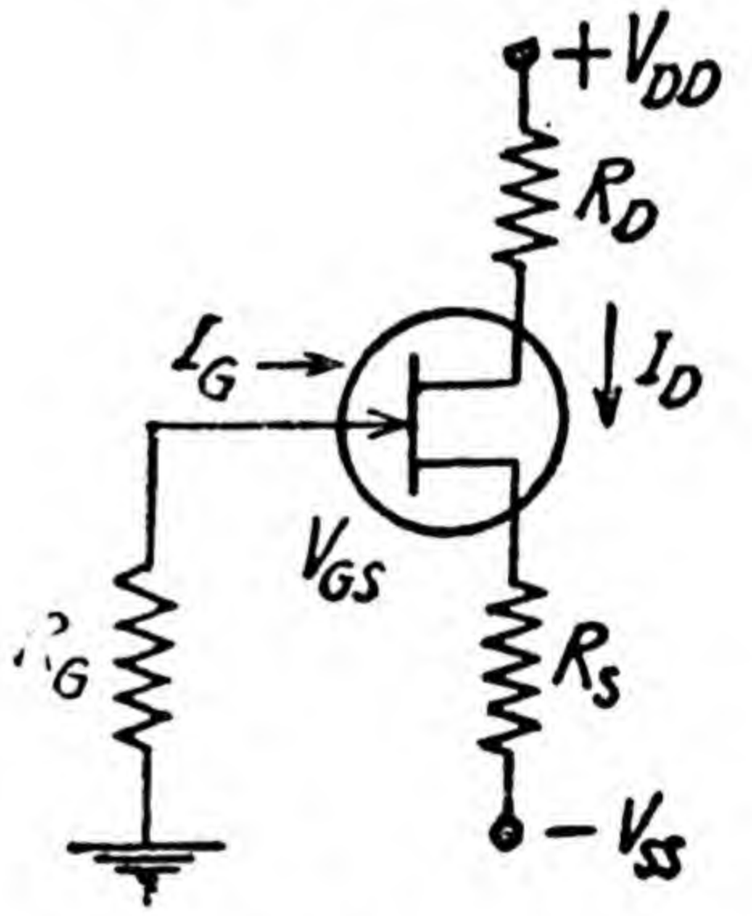


Fig. 3.

17·2. FET SMALL SIGNAL MODELS :

The complete a.c. equivalent circuit of a FET includes lead

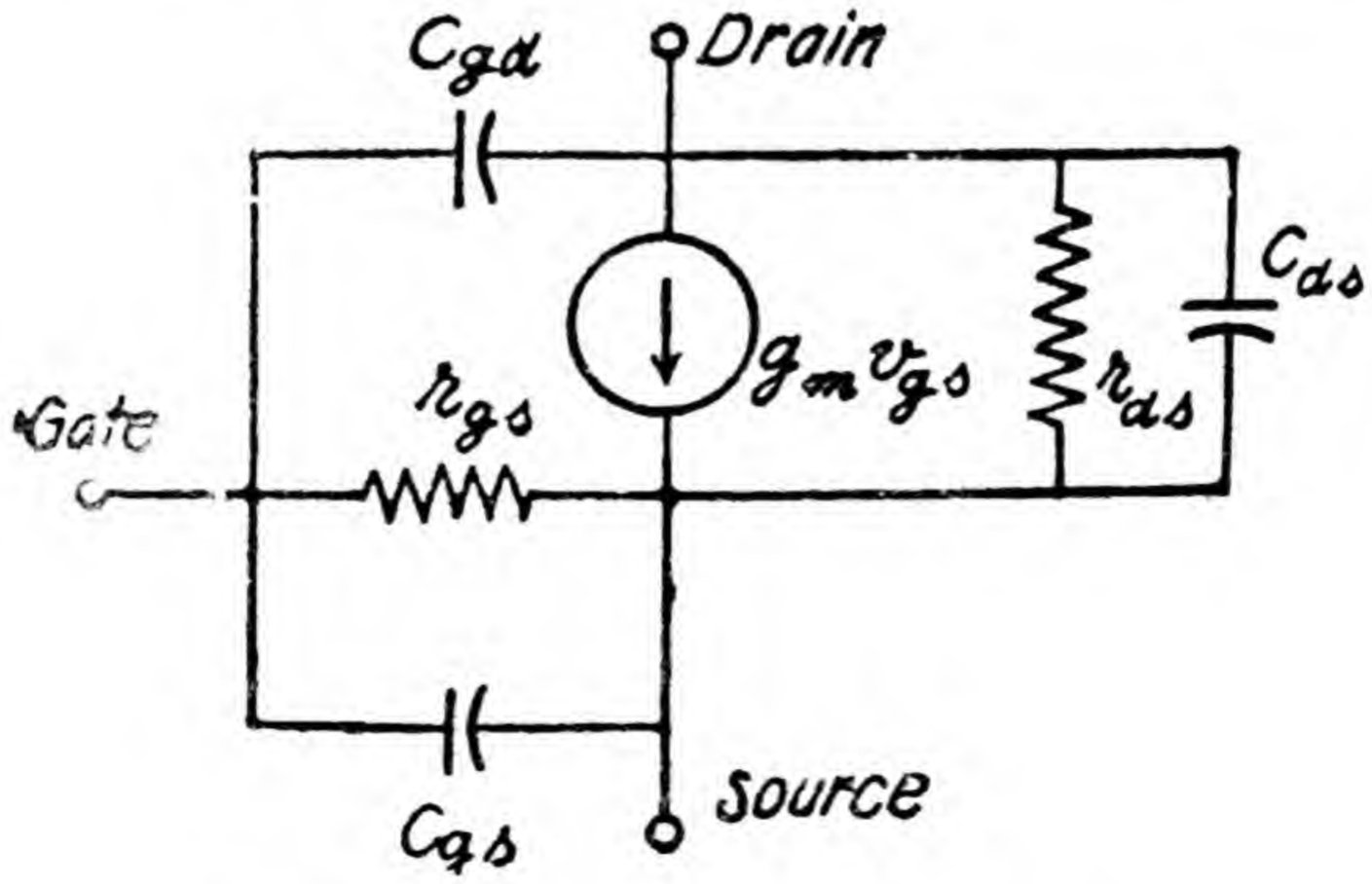


Fig 4. a.c. equivalent circuit.

inductances, internal resistances, internal capacitances etc. However, in practice, the lead inductances are neglected. Fig. (4) shows the equivalent FET circuit in Norton's form. On input side, C_{gd} and C_{gs} represent the capacitance between gate and drain and capacitance between gate and source respectively. r_{gs} is the a.c. resistance from gate to source. On the output side, C_{ds} represents the capacitance between drain and source while r_{ds} is the a c. resistance from drain to source.

Low frequency model : At low frequencies, all the capacitances becomes high enough to neglect. The input resistance between gate and source r_{gs} in infinite because reversed biased gate current is assumed to be zero. Similarly the resistance between gate and drain is infinite. The equivalent circuit is shown in fig. 5. If we compare the low frequency model of FET with a low frequency *h*-parameter model of BJT, we observe the following points :

(i) Both have Norton's output circuit,

(ii) In FET model, the generator current depends upon input voltage v_{gs} while in BJT model, the generator voltage depends upon input current.

(iii) In case of BJT, the feedback takes place from output circuit to input circuit through the parameter h_{rs} while there is no feed back in case of FET.

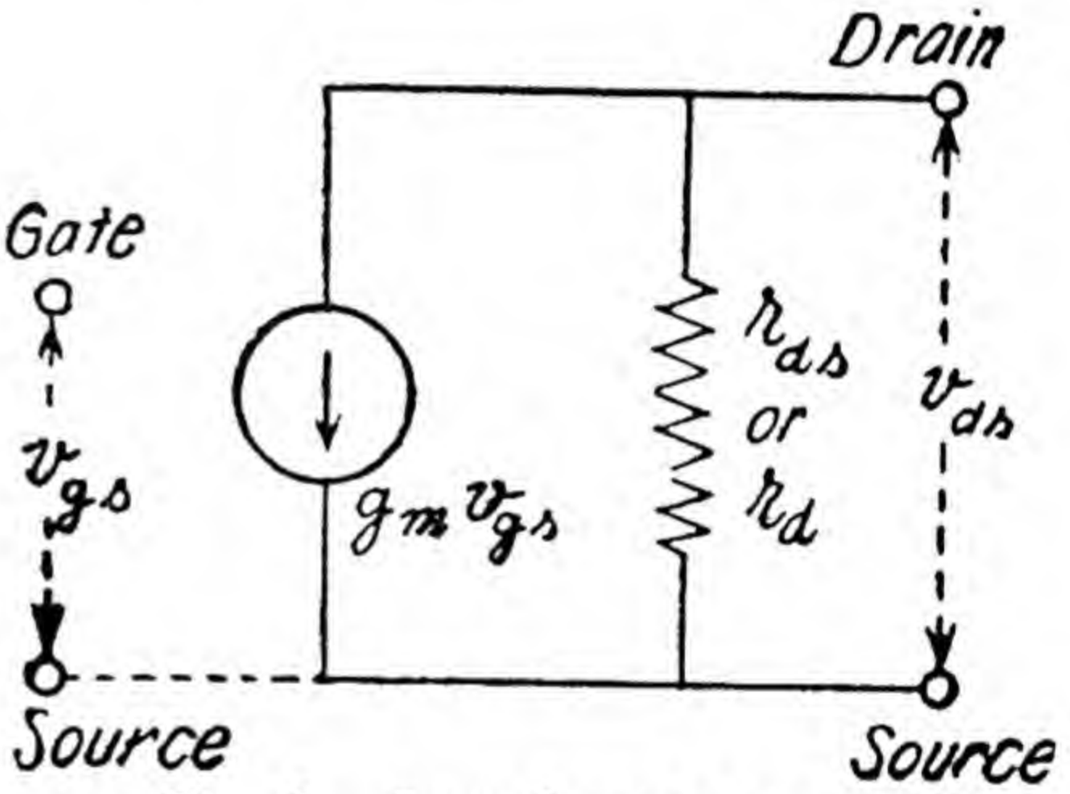


Fig. 5. Low frequency small signal model.

(iv) Input resistance of common emitter is about 1 $K\Omega$ while FET has almost infinite resistance.

The above facts show that low frequency FET forms a more ideal amplifier than conventional BJT. It should be remembered that the superiority of FET is restricted only to low frequencies and not to high frequencies

High frequency model : The high frequency small signal model of FET is shown in fig. 6. At high frequencies, the capacitances between different nodes can not be neglected. These capacitances produce the following effects :

(i) There is a feed back from output circuit to input circuit,

(ii) voltage gain drops rapidly as frequency increases.

For diffused junction FET, the different capacitances are

$$C_{ds} = 0{\cdot}1 \text{ to } 1\ pF$$

and

$$C_{gs} \text{ and } C_{gd} = 1\text{-}10\ pF.$$

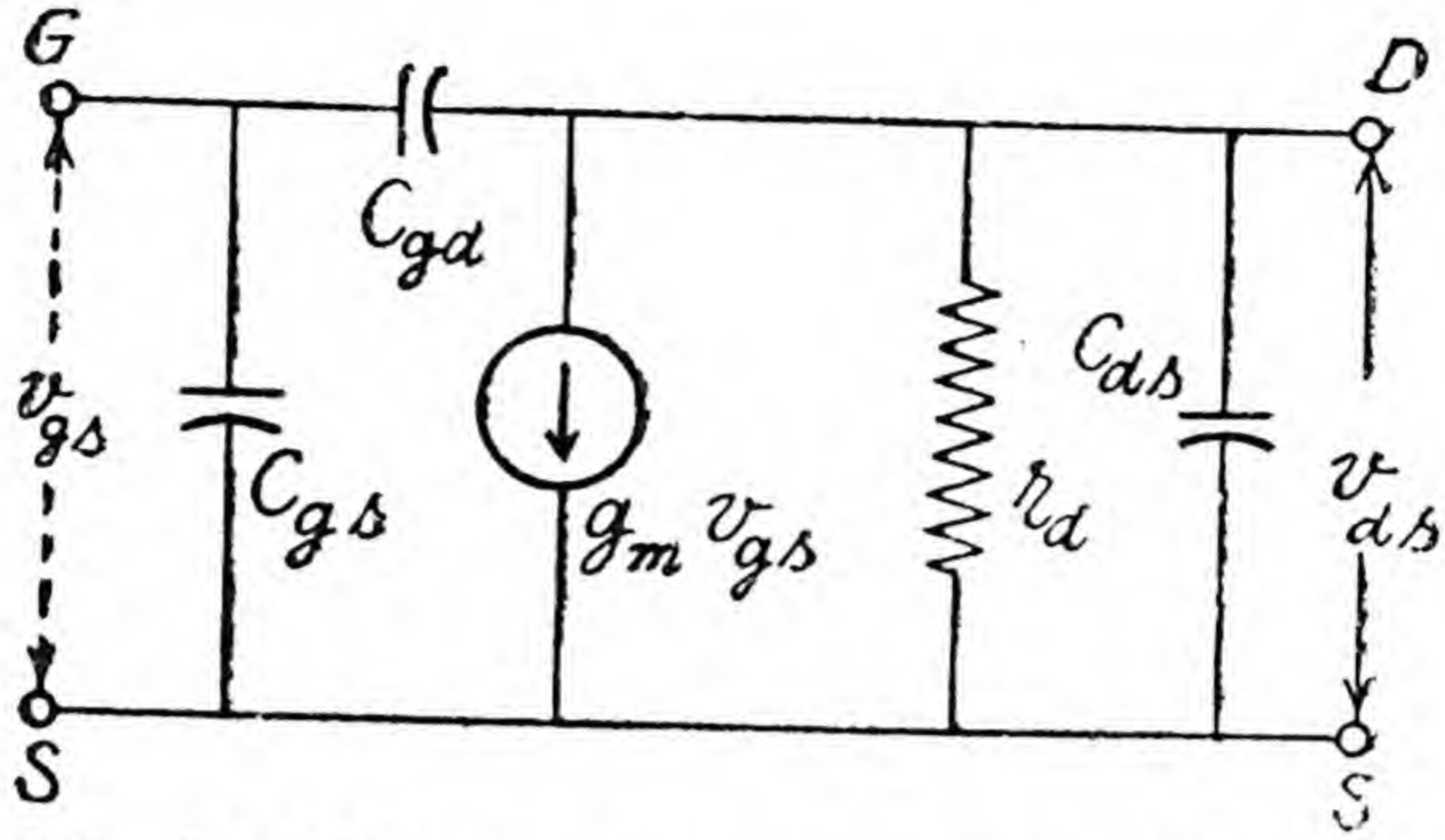


Fig. 6. High frequency small signal model.

17·3. COMMON SOURCE A.C. AMPLIFIER :

Figure 7 shows the common source a.c. amplifier. In fig. 7

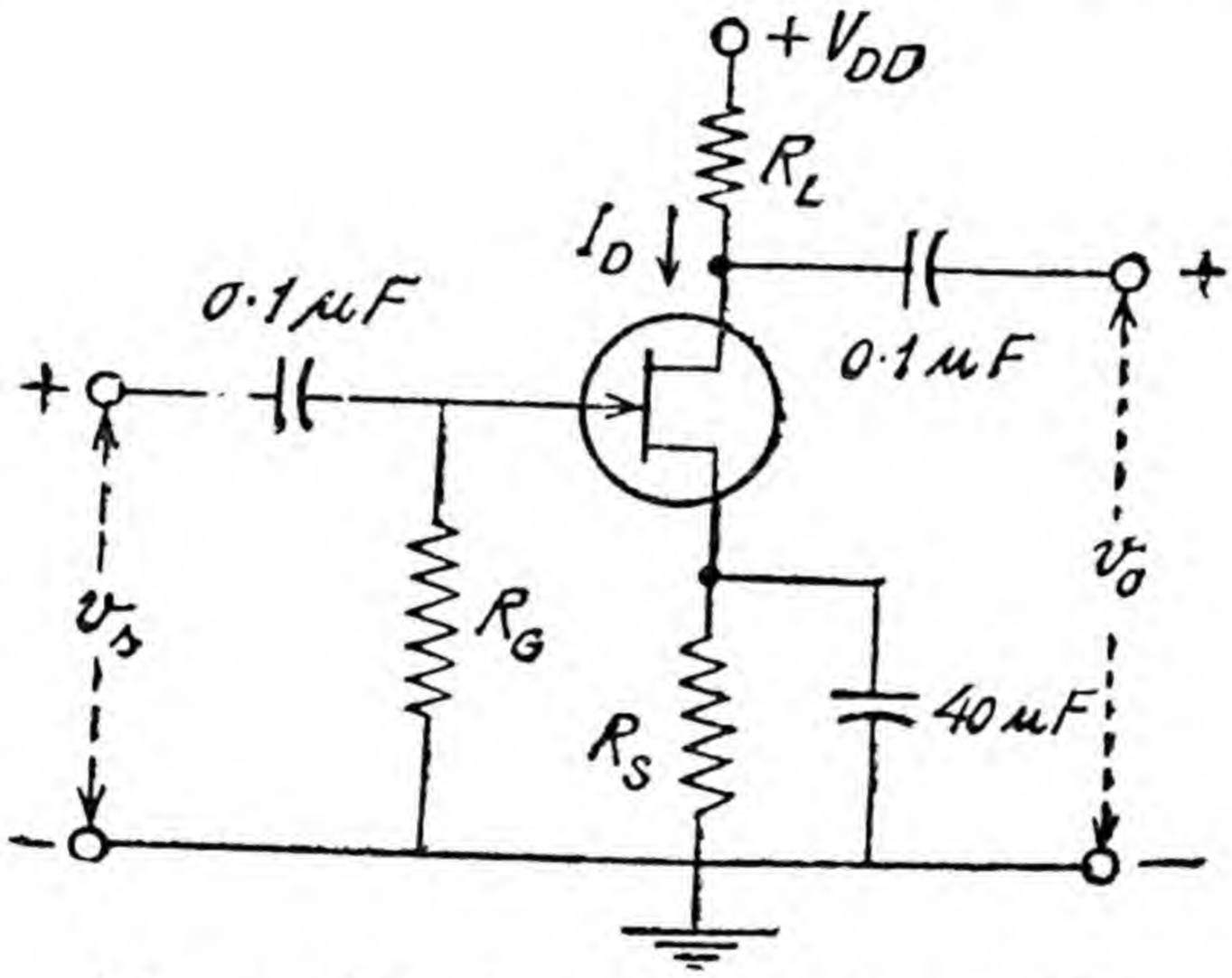


Fig. 7. Commom source a.c. amplifier.

the drain current flows through R_S and develops a voltage which reverse-biases the gate junction. R_G is provided for a d.c. return path The Thevenin equivalent circuit for small signal common source a.c. amplifier is shown in fig. 8.

Since the same equivalent circuit can be used as for triode and hence the voltage gain and other quantities can be calculated similarly.

Voltage gain A_v—The voltage gain is given by

$$A_v = \frac{-\mu R_L}{r_a + R_L} \qquad \ldots(1)$$

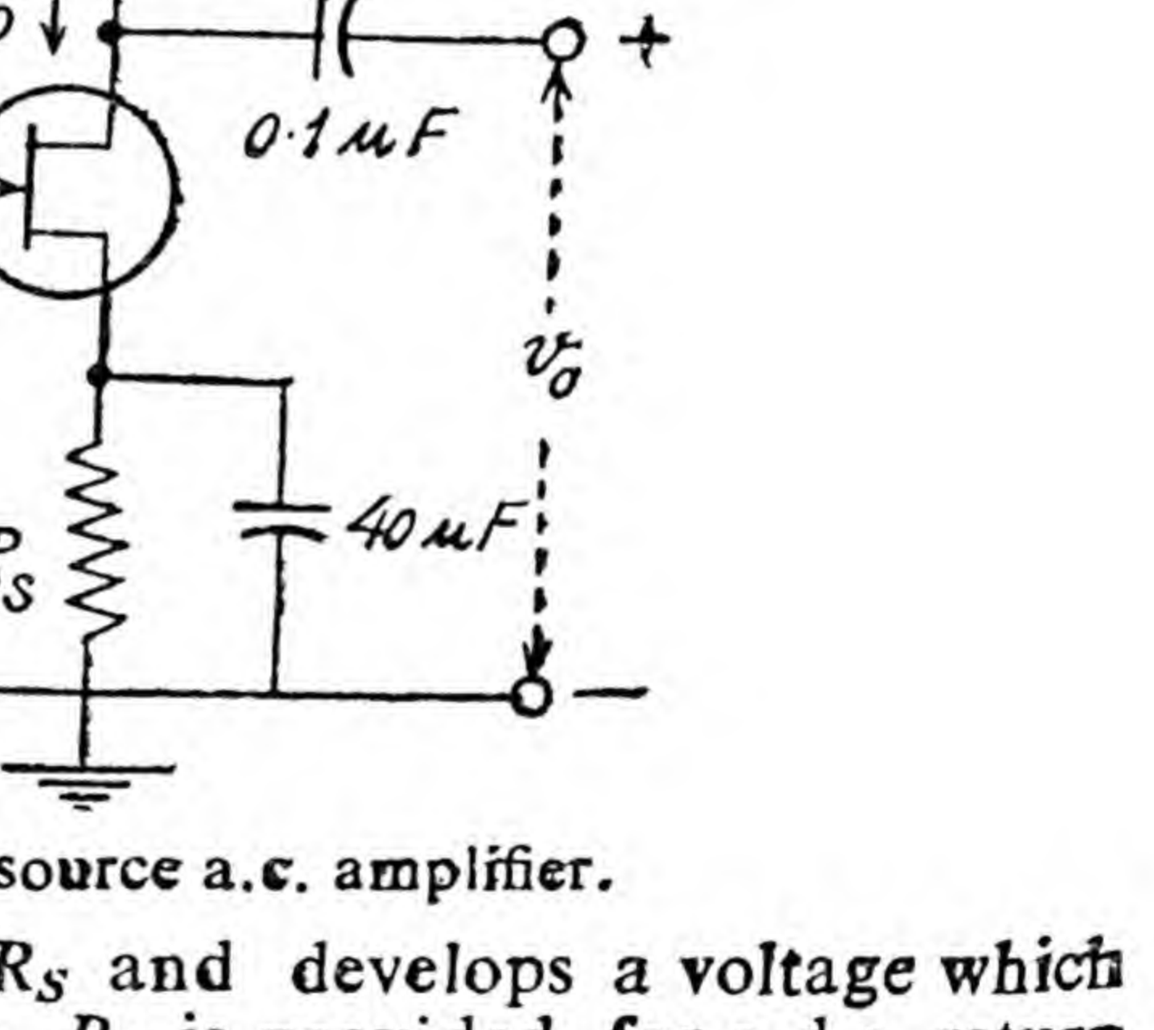


Fig. 8. Thevenin equivalent circuit.

or $$A_V = -g_m \frac{r_d R_L}{r_d + R_L} = -g_m \frac{R_L}{1 + (R_L/r_d)}$$

If $r_d \gg R_L$ (not always), then $A_V \simeq -g_m R_L$.

For example using a FET 2N3819 for which g_m varies from 2000 to 6500μ mhos and for $R_L = 10K$ ohms and $r_d = 20K$ ohms, we have $A_V = -13$ to -42.

Output resistance—The output resistance is a parallel combination of r_d and R_L which is given by $[(R_L r_d/(R_L + r_d)]$. For the above example the output resistance is approximately $7K$ ohm.

Input resistance—The input resistance to the amplifier is R_G.

17·4. THE COMMON DRAIN OR SOURCE FOLLOWER :

The basic circuit of a *P*-channel *FET* source follower is shown in fig. (9). The following expressions can easily be derived :

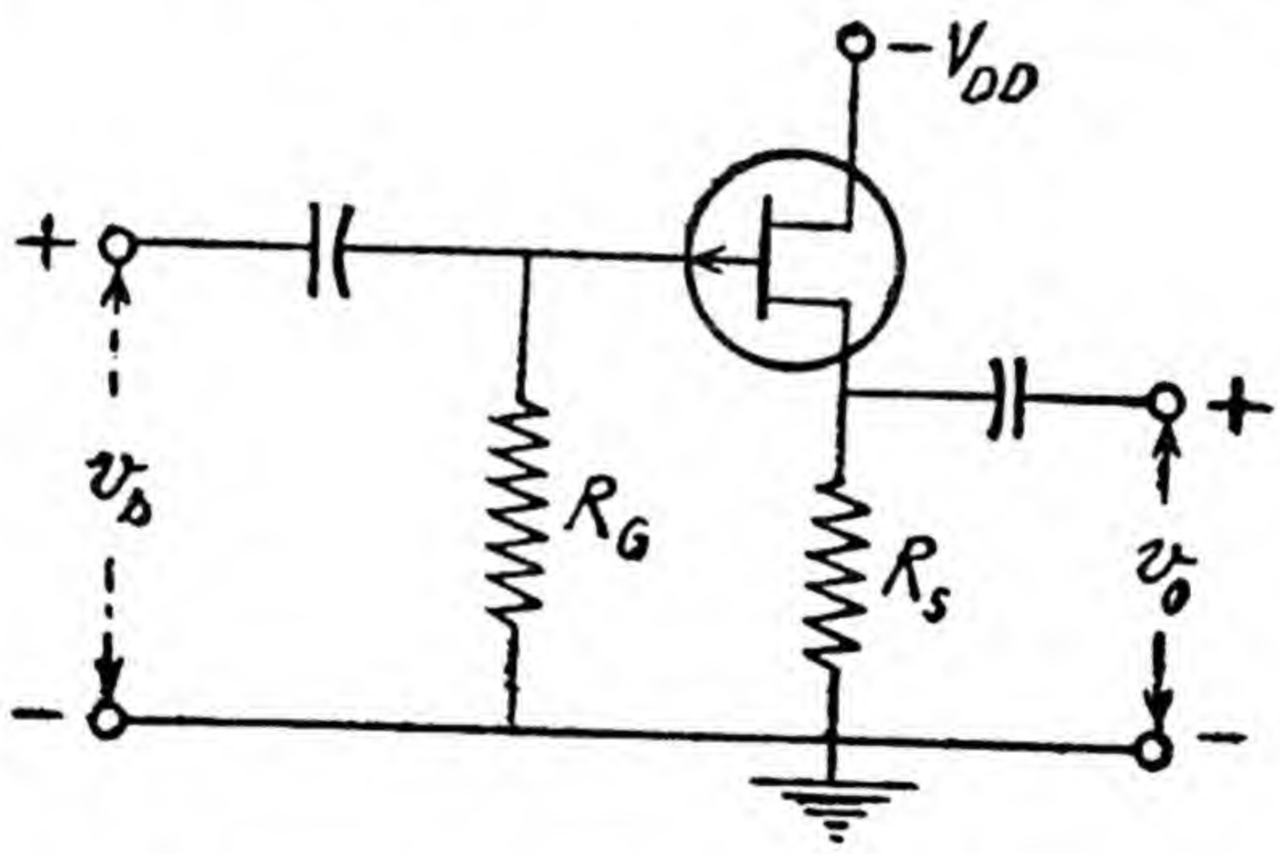


Fig. 9. The common drain or source follower.

Voltage gain $$A_V = \frac{R_S}{\frac{(\mu+1)}{\mu} R_S + \frac{1}{g_m}}$$

$$\approx \frac{g_m R_S}{1 + g_m R_S} \text{ when } \mu \gg 1$$

Output resistance—Output resistance is a parallel combination of R_S and $r_d/(1+\mu)$. Hence

$$R_O = \frac{R_S \times r_d/(1+\mu)}{R_S + r_d/(1+\mu)} \approx \frac{R_S}{1 + g_m R_S} \text{ when } \mu \gg 1$$

Input resistance—The input resistance will be determined by R_G. Following are the characteristics of a source follower FET like emitter follower :

(i) Voltage gain is less than unity.
(ii) There is no phase reversal.
(iii) High input impedance.
(iv) Low output impedance.

17·5. COMMON GATE AMPLIFIER :

The basic circuit of a common gate *N*-channel *FET* amplifier is shown in fig. (10). The circuit is just like a grounded grid amplifier.

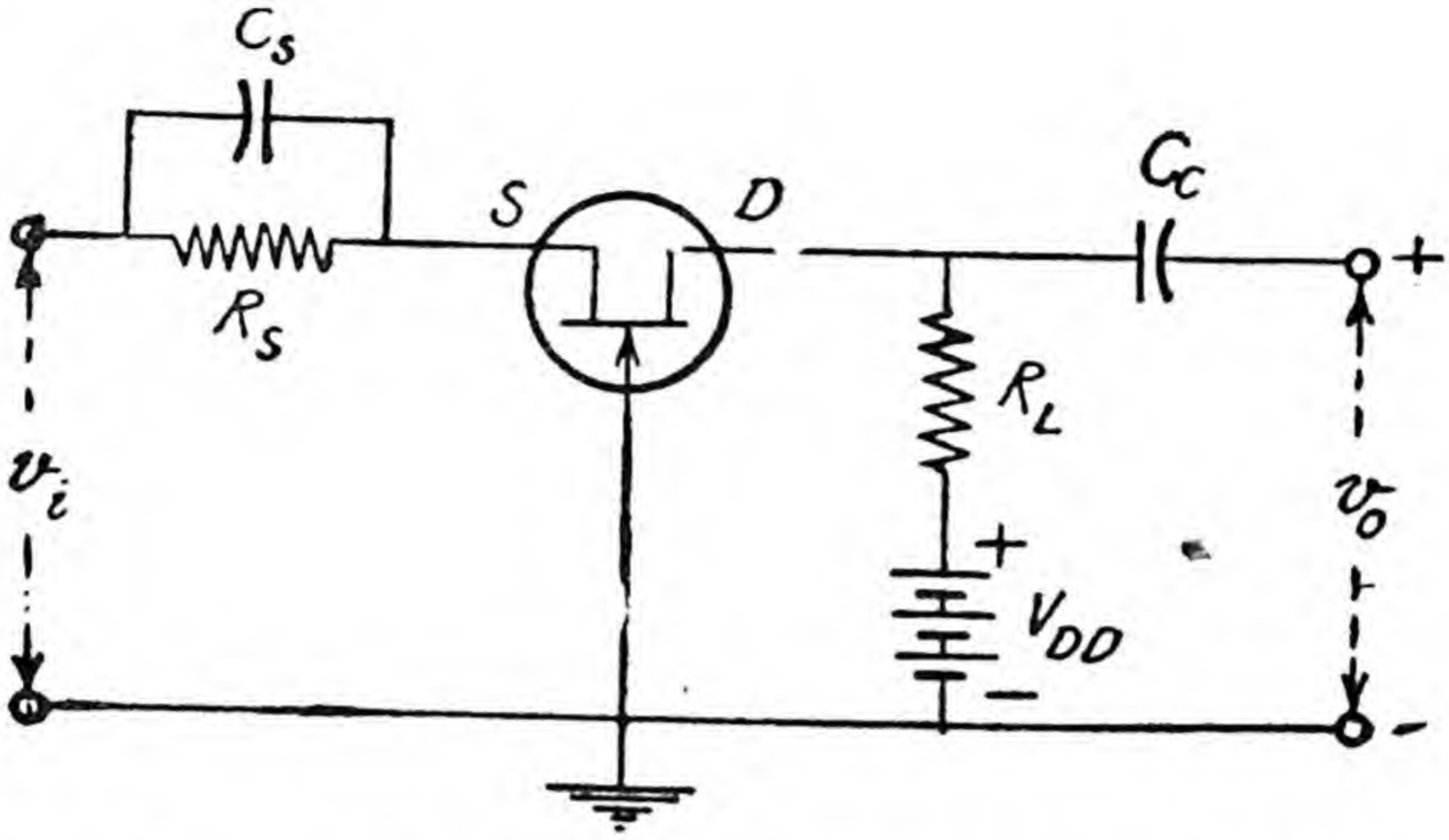


Fig. 10. Basic circuit of common gate *N*-channel *FET* amplifier.

In this amplifier, we have

$$\text{Voltage gain} = \frac{(\mu+1)\ R_L}{r_d+R_L} = \frac{(r_d g_m+1)\ R_L}{r_d+R_L}$$

$$\text{Input resistance} = \frac{r_d+R_L}{(\mu+1)}$$

$$\text{Output resistance} = r_d+R_S\ (1+\mu)$$

The amplifier has the following characteristics :

(i) High voltage gain (if R_S is small).
(ii) Low input impedance.
(iii) High output impedance.
(iv) No phase reversal.

Example 1. *Using the equivalent circuit show that if two identical FETs are connected in parallel, then g_m is double and r_d is half of that of individual FET. If two FETs are not identical show that*

$$\frac{1}{r_d} = \frac{1}{r_{d1}} + \frac{1}{r_{d2}}$$

and

$$\mu = \frac{\mu_1 r_{d2} + \mu_2 r_{d1}}{r_{d1}+r_{d2}}$$

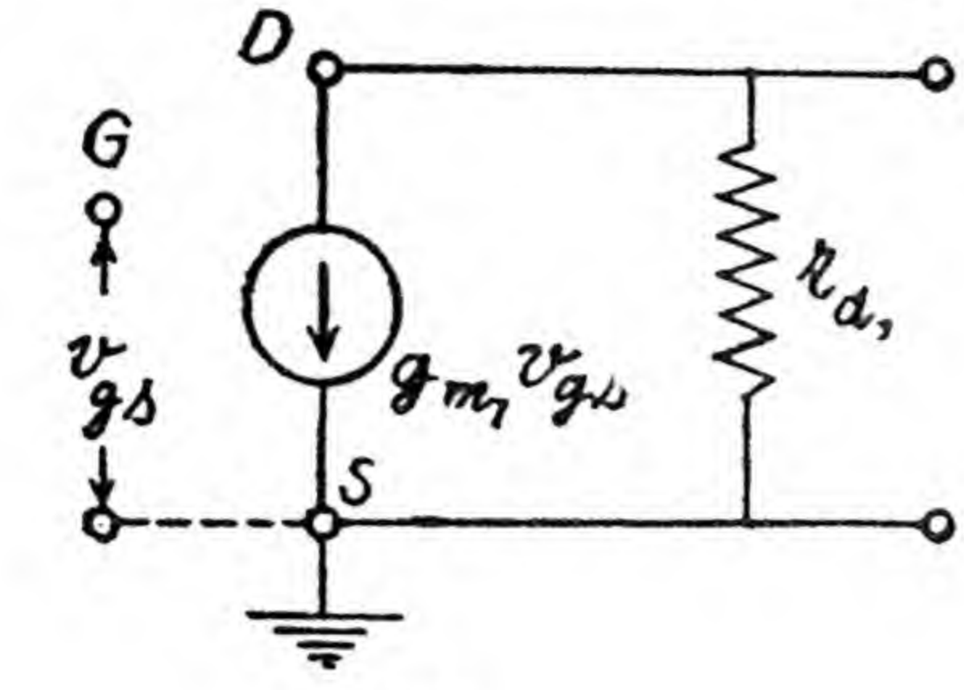
Let two *FETs* with mutual conductance g_{m1} and g_{m2} and resistances r_{d1} and r_{d2} be connected in parallel as shown in fig. (11). The two drain resistances are in parallel and total current flowing in the circuit is $g_{m1}\ v_{gs} + g_{m2}\ v_{gs}$.

Considering the equivalent circuit, we have

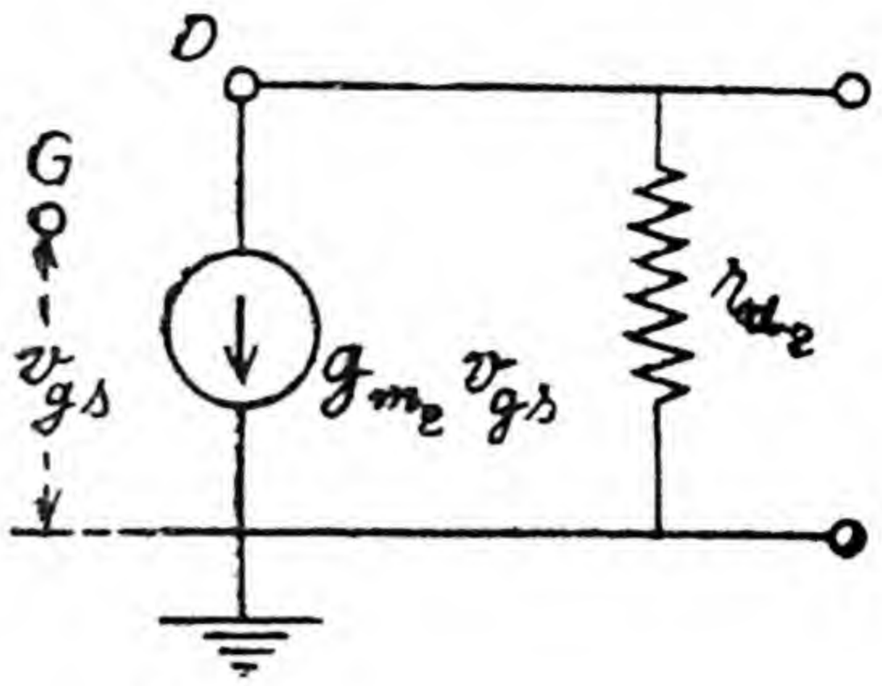
$$g_m v_{gs} = g_{m1} v_{gs} + g_{m2} v_{gs}$$

and

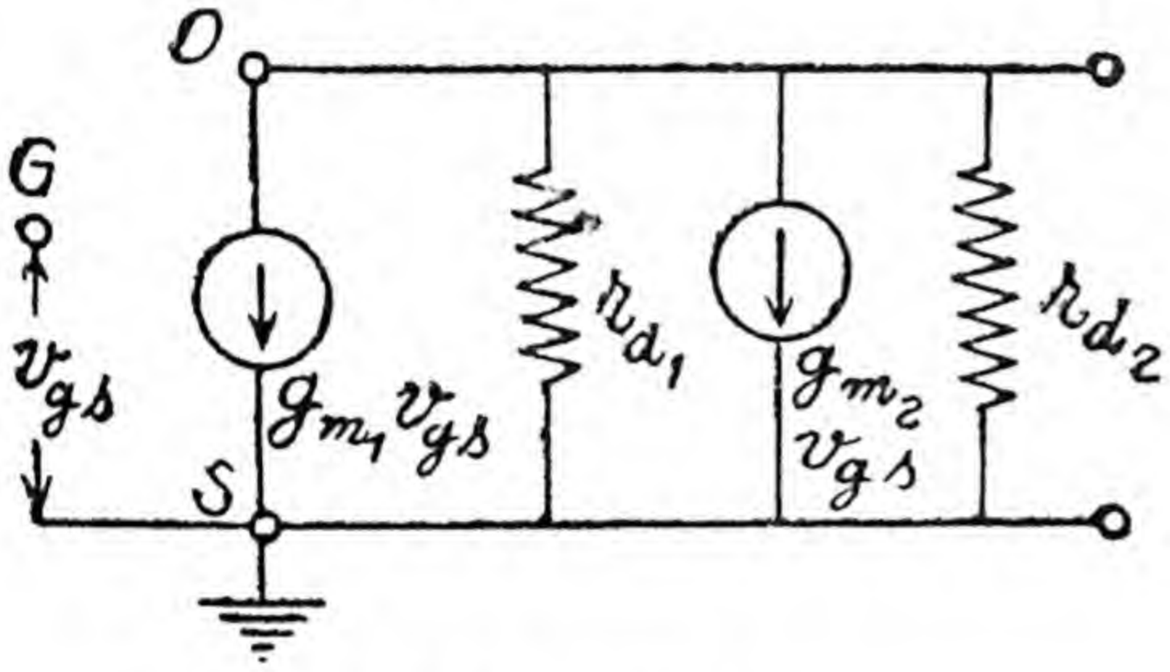
$$r_d = \frac{r_{d1} . r_{d2}}{r_{d1}+r_{d2}}$$



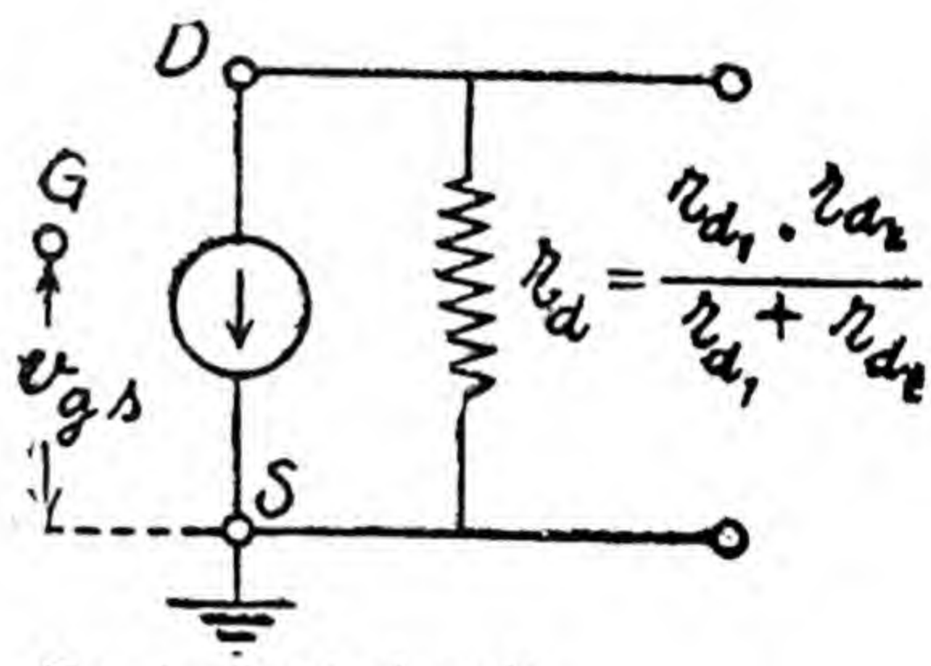
FET-1



FET-2



Two FETs in parallel



Equivalent circuit

Fig. 11. Two FETs connected in parallel

For identical FETs, $g_{m1}=g_{m2}$ and $r_{d1}=r_{d2}$

Thus $g_m=g_{m1}+g_{m2}=2g_{m1}$ and $r_d=\frac{r_{d1}.r_{d1}}{r_{d1}+r_{d1}}=\frac{r_{d1}}{2}$

when the FETs are not identical

$$r_d=\frac{r_{d1}.r_{d2}}{r_{d1}+r_{d2}} \text{ and } g_m=g_{m1}+g_{m2}$$

Now $\mu=r_d\, g_m=\frac{r_{d1}.r_{d2}}{r_{d1}+r_{d2}}\,(g_{m1}+g_{m2})$

On simplification, we have $\mu=\frac{\mu_1 r_{d2}+\mu_2 r_{d1}}{r_{d1}+r_{d2}}$

17·6. GENERAL TREATMENT OF LOW FREQUENCY COMMON SOURCE AND COMMON DRAIN AMPLIFIER :

For analysing common source and common drain amplifier simultaneously, we draw a generalized circuit as shown in fig. (12*a*). In the circuit, we have not shown the biasing arrangement for the sake of clarity but it has been assumed that the circuits are properly biased for linear operation. The equivalent circuit is shown in fig. (12*b*). For common source amplifier $R_S=0$ and the output $(v_0)_1$ is picked up at drain terminal D. In case of a common drain amplifier, $R_L=0$ and output $(v_0)_2$ is picked up at source terminal S.

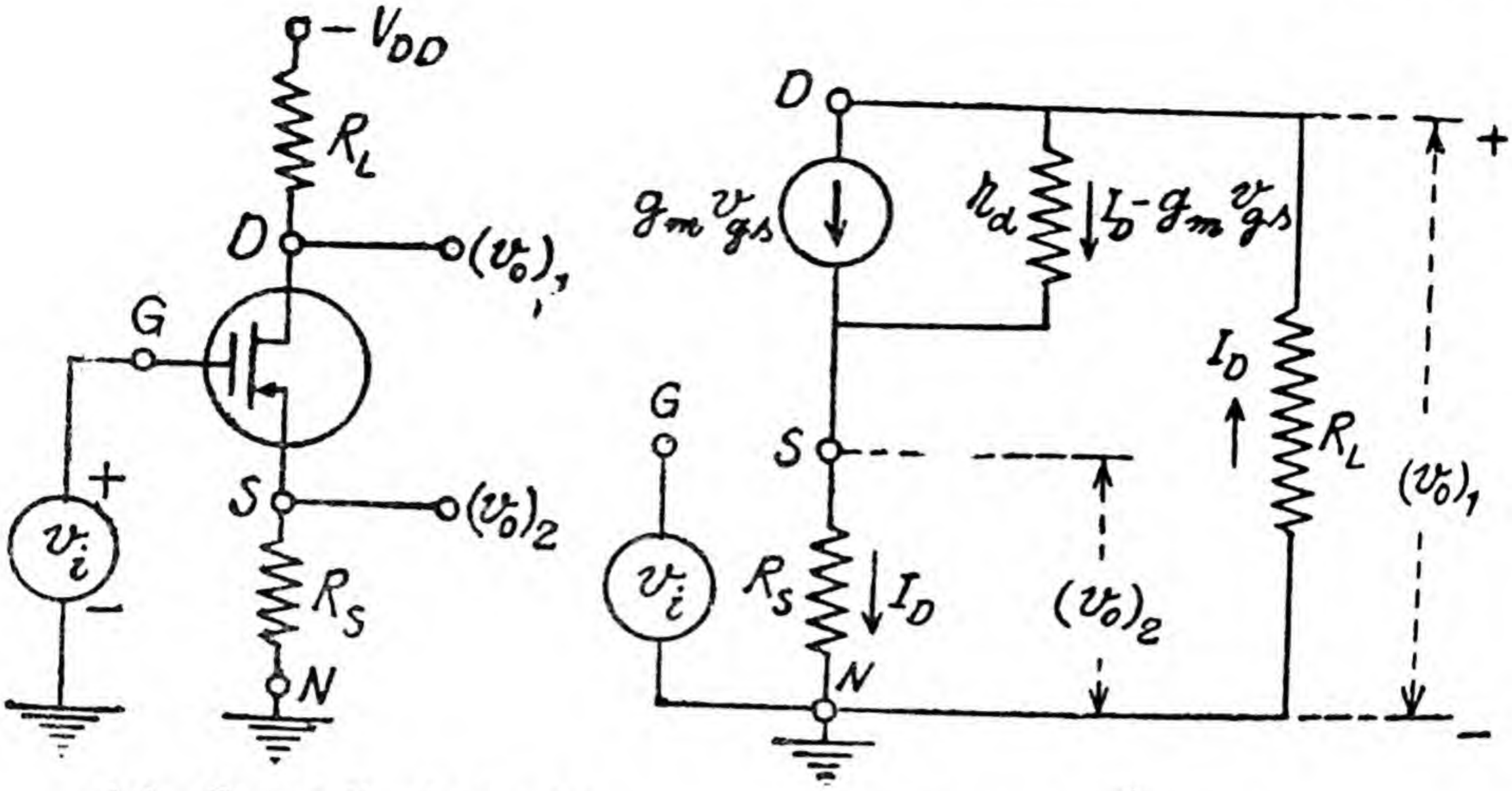
Applying Kirchoff's voltages law to the output, we have

$$I_D R_L+(I_D-g_m v_{gs})\, r_d+I_D R_S=0 \qquad \ldots(1)$$

But $$v_{gs}=v_i-I_D R_S \qquad \ldots(2)$$

Substituting the value of the v_{gs} from eq. (2) in eq. (1), we get

$$I_D R_L+\{I_D-g_m\,(v_i-I_D R_S)\}\, r_d+I_D R_S=0$$



(*a*) Generalized circuit (*b*) Equivalent circuit.

Fig. 12.

Solving we get $I_D = \dfrac{\mu v_i}{r_a + R_L + (\mu+1)R_S}$ ...(3)

We now consider the following two cases :

(1) Common source amplifier with unbypassed source resistance R_S. In this amplifier, R_L constitute the load impedance. The voltage $(v_0)_1$ developed across load is given by

$$(v_0)_1 = -I_D \times R_L = \frac{-\mu v_i R_L}{r_a + R_L + (\mu+1)\,R_S} \quad \ldots(4)$$

The negative sign indicates that the output voltage is out of phase with the input voltage. Corresponding to eq. (4), the Thevenin's equivalent circuit is shown in fig. 13. Now voltage gain is given by

$$A_V = \frac{(v_0)_1}{v_i} = \frac{-\mu R_L}{r_a + R_L + (\mu+1)R_S} \quad \ldots(5)$$

when R_S is bypassed with a large value capacitor in parallel or S is grounded (*i.e.*, $R_S = 0$), then voltage gain is given by

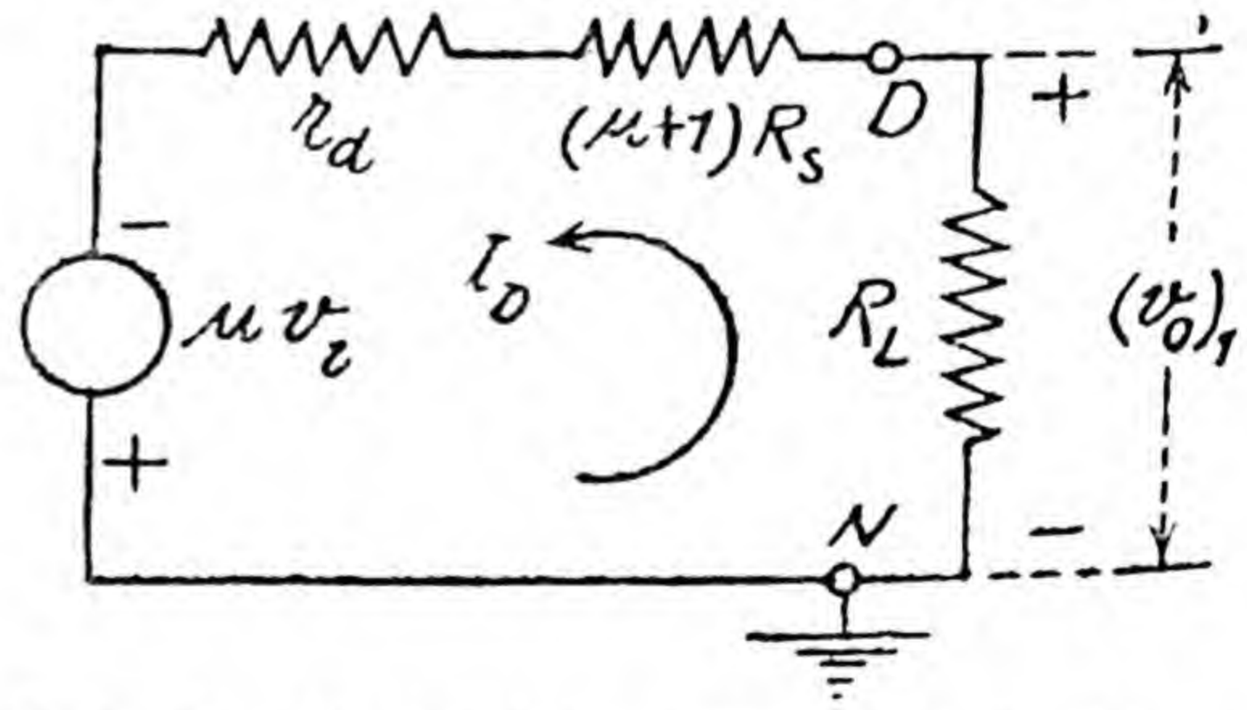


Fig. 13. Thevenin's equivalent circuit.

$$A_V = \frac{-\mu R_L}{r_a + R_L} \quad \ldots(6)$$

(2) Common drain amplifier with drain resistor R_L. In this case, output is taken across R_S. The output voltage $(v_0)_2$ is given by

$$(v_0)_2 = I_n R_S = \frac{\mu v_i R_S}{r_a + R_L + (\mu+1)R_S}$$

$$= \frac{(\mu/\mu+1)\, v_i R_S}{(r_d+R_L)/(\mu+1)+R_S} \qquad \ldots(7)$$

The Thevenin's equivalent circuit corresponding to eq. (7) is shown in fig. (14).

The voltage gain A_V is given by

$$A_V = \frac{(v_0)_2}{v_i} = \frac{\frac{\mu}{\mu+1}.R_S}{\frac{(r_d+R_L)}{(\mu+1)}+R_S} \qquad \ldots(8)$$

The output voltage is in phase with the input voltage.

When $R_L=0$, then

$$A_V = \frac{[\mu/(\mu+1).R_S]}{\frac{r_d}{(\mu+1)}+R_S} \qquad \ldots(9)$$

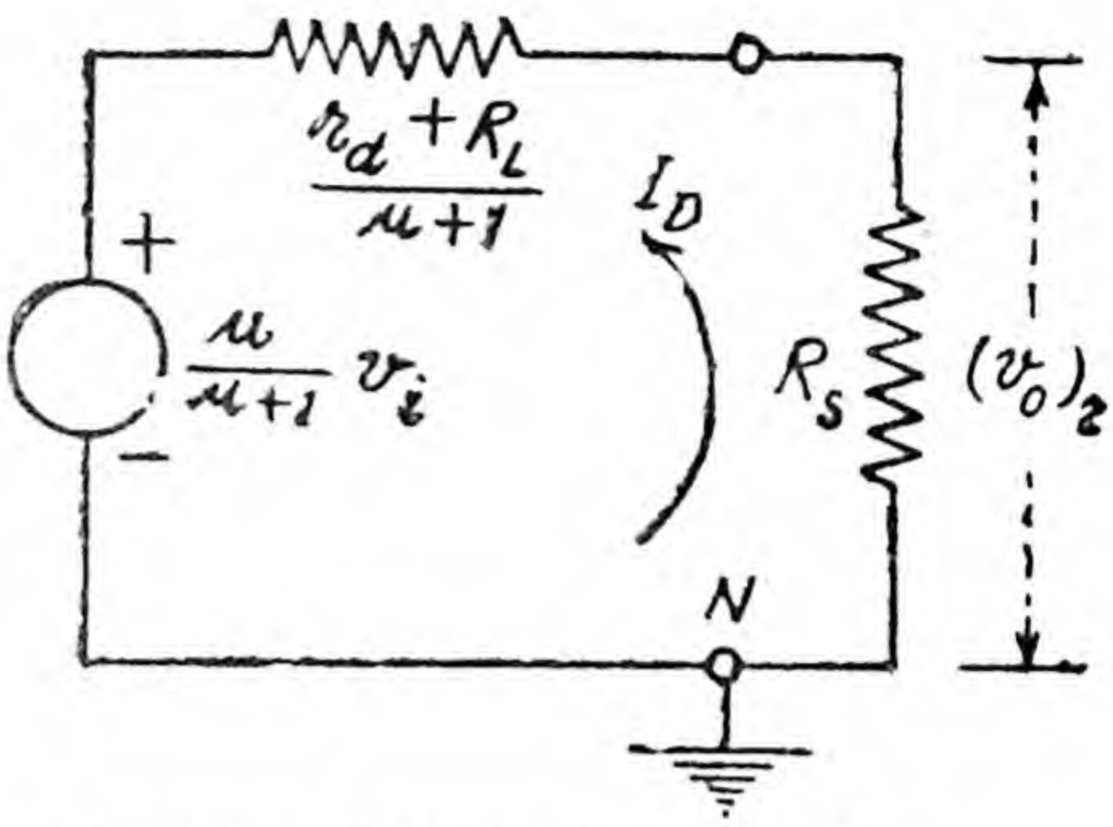


Fig. 14. Thevenin's equivalent circuit.

Expression (9) can be written as

$$A_V = \frac{i.R_S}{r_d+(\mu+1)R_S} \approx \frac{\mu}{(\mu+1)} \text{ when } (\mu+1)\, R_S \gg r_d$$

For $\mu \gg 1$, A_V is approximately equal to one. This signifies that the output voltage at the source S follows the input voltage at the gate. Due to this fact common drain amplifier is referred to as source follower.

Ex. 1. *A common source amplifier uses FET having drain resistance $r_d=150$ KΩ ana $\mu=15$. Calculate the voltage gain for load resistance R_L equal to (i) 150 KΩ and (ii) 900 KΩ.*

The voltage gain of a common source amplifier is given by

$$A_V = -\frac{\mu R_L}{r_d+R_L}$$

(i) Here $\mu=15$, $r_d=150$ KΩ and $R_L=150$ KΩ.

$$\therefore \quad A_V = -\frac{15\times(150\times10^3)}{(150+150)\ 10^3} = -\frac{15\times150}{300} = -7{\cdot}5$$

(ii) Here $\mu=15$, $r_d=150$ KΩ and $R_L=900$ KΩ.

$$\therefore \quad A_V = -\frac{15\times(900\times10^3)}{(150+900)\ 10^3} = -\frac{15\times900}{1050} = -12{\cdot}8$$

Ex. 2. *A common source FET amplifier uses load resistance $R_L=200$ KΩ and unbypassed resistor R_S in the source circuit. The FET has drain resistance $r_d=100$ KΩ and $\mu=20$. Compute the voltage gain and output impedance for $R_S=5$KΩ.*

The voltage gain and output impedance of a common source amplifier with unbypassed resistor are given by

$$A_V = \frac{-\mu R_L}{r_d+R_L+(\mu+1)R_S} \text{ and } R_o = r_d+(\mu+1)\, R_S$$

Here $R_L=200\ K\Omega$, $R_S=5K\Omega$, $r_d=100K\Omega$ and $\mu=20$

$$\therefore \quad A_V=\frac{-20\times200\times10^3}{[100+200+21\times5]\ 10^3}=-\frac{20\times200}{405}$$

$$=-9{\cdot}876$$

Again $R_o=[100+21\times5]\ 10^3=205\times10^3\Omega$

$=205\ K\Omega$.

Ex. 3. *A common drain amplifier uses FET having* $r_d=200\ K\Omega$ *and* $\mu=20$. *Compute voltage gain and output impedance for a load resistance* $R_S=300\ K\Omega$.

For a common drain amplifier without R_L, the voltage gain and output impedance are given by

$$A_V=\frac{\mu R_S}{r_d+(\mu+1)\ R_S} \quad \text{and} \quad R_o=\frac{r_d}{(\mu+1)}$$

Here $\mu=20$, $R_S=300\ K\Omega$ and $r_d=200\ K\Omega$.

$$A_V=\frac{20\times300\times10^3}{[200+21\times300]\ 10^3}=\frac{20\times300}{6500}$$

$=0{\cdot}923$.

And $$R_o=\frac{200\times10^3}{21}=9{\cdot}52\ K\Omega$$

Ex. 4. *A common drain amplifier uses FET having* $r_d=200\ K\Omega$ *and* $\mu=20$. *The load resistance* R_S *is* 150 $K\Omega$. *A resistor* R_L *is placed in the drain circuit. Compute the output impedance and voltage gain for* $R_L=100\ K\Omega$.

The voltage gain and output impedance of a common drain amplifier with drain resistor R_L are given by

$$A_V=\frac{\mu R_S}{r_d+R_L+(\mu+1)R_S} \quad \text{and} \quad R_o=\frac{r_d+R_L}{\mu+1}$$

Here $r_d=200\ K\Omega$, $\mu=20$, $R_S=150\ K\Omega$ and $R_L=100\ K\Omega$

$$\therefore \quad A_V=\frac{20\times150\times10^3}{[200+100+21\times150]\ 10^3}=\frac{20\times150}{3450}$$

$=0{\cdot}87$

And $$R_o=\frac{200\times10^3+100\times10^3}{21}\Omega$$

$$=\frac{340\times10^3}{21}=14{\cdot}3\ K\Omega$$

17·7. COMMON SOURCE AMPLIFIER AT HIGH FREQUENCIES :

Here we shall describe the analysis of common source *FET* amplifier at high frequencies. The analysis can equally be applied to MOSFETs and the only difference lies in the biasing arrangement. Fig. (15) shows the common source amplifier without biasing arrangement. In the circuit the different interelectrode capacitances are also shown. The Norton's equivalent circuit at high frequencies is shown in fig. (16). Using the equivalent circuit, we

hall calculate (i) voltage gain, (ii) input admittance (iii), input apacitance and (iv) output admittance.

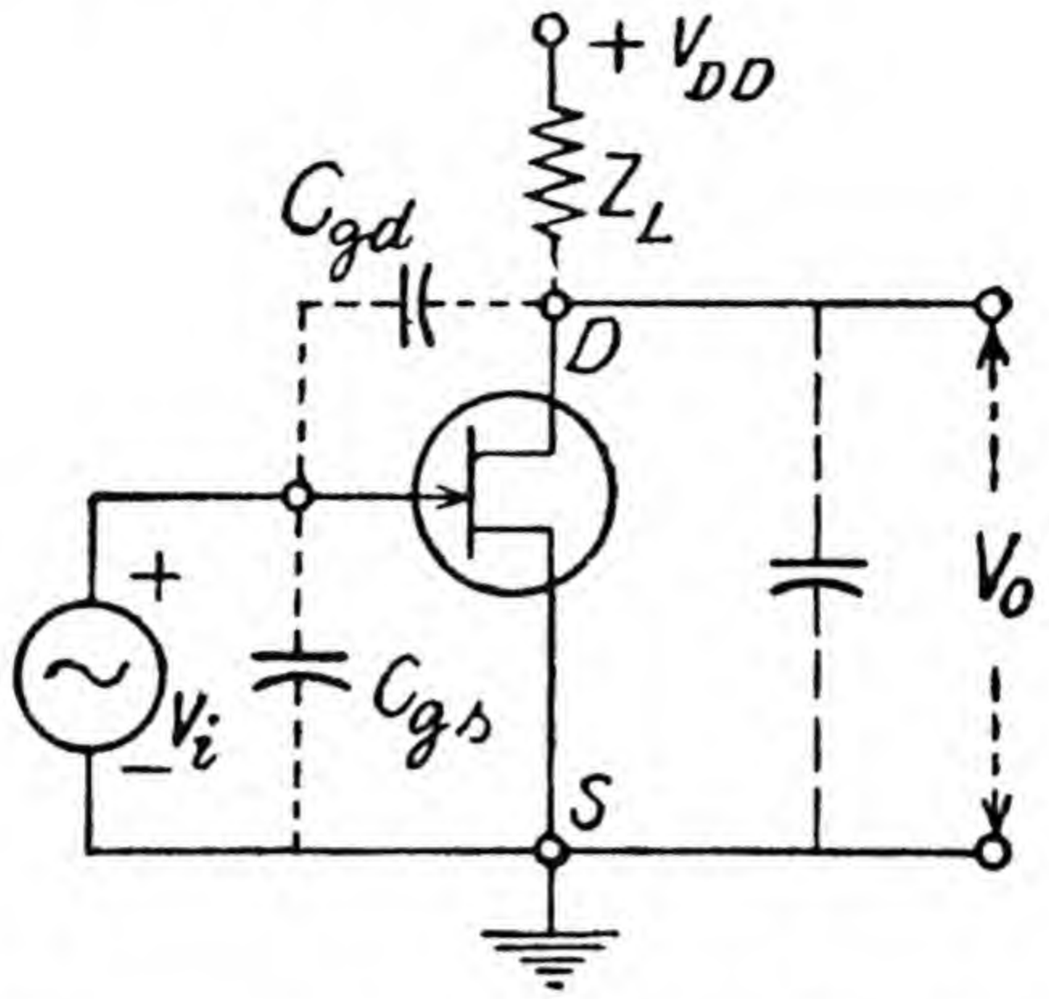


Fig. 15. Common source amplifier taking into account the different interelectrode capacitances.

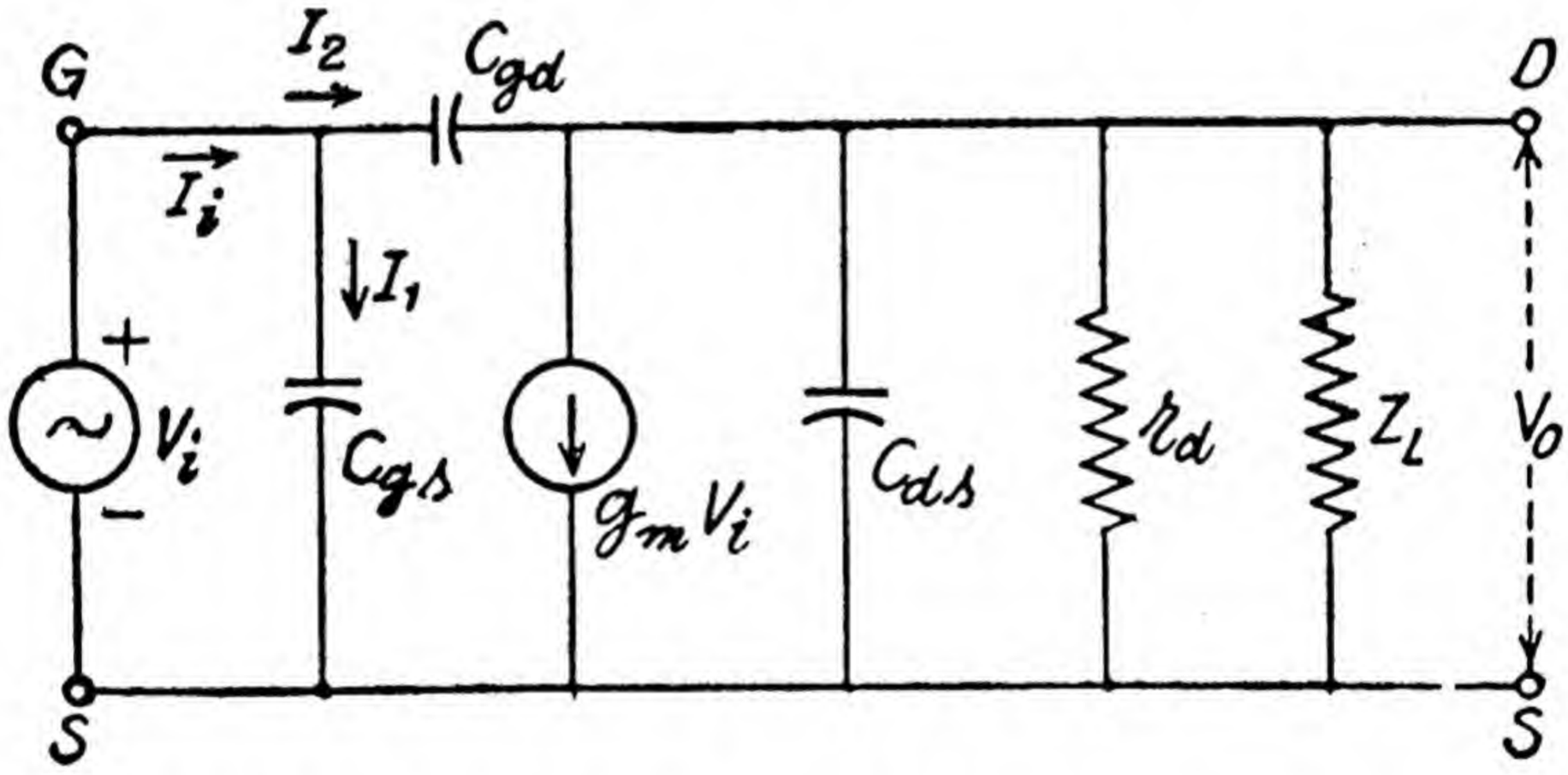


Fig. 16. Norton's small single equivalent circuit of common source amplifier at high frequencies.

(i) Voltage gain : The output voltage V_o is given by

$$V_o = IZ \qquad \ldots(1)$$

where I=current in the lead short circuiting terminals D and S, and

Z=Impedance seen between terminals D and S on short circuiting all independent voltage sources and open circuiting all the independent current sources.

The value of I is given by

$$I = -g_m V_i + V_i/(1/j\omega C_{gd}) = -g_m V_i + V_i (j\omega C_{gd}) \qquad \ldots(2)$$

The value of Z is the impedance of parallel combination of Z_L, r_d, C_{ds} and C_{gd}. So

$$\frac{1}{Z}=\frac{1}{Z_L}+\frac{1}{r_d}+\frac{1}{(1/j\omega C_{gd})}+\frac{1}{(1/j\omega C_{ds})}$$

or $$\frac{1}{Z}=\frac{1}{Z_L}+\frac{1}{r_d}+j\omega C_{gd}+j\omega C_{ds}$$

or $$Y=\frac{1}{Z}=Y_L+g_d+Y_{gd}+Y_{ds} \qquad ...(3)$$

where $Y_L=\frac{1}{Z_L}$ = admittance corresponding to Z_L

$g_d=\frac{1}{r_d}$ = conductance corresponding to r_d

$Y_{gd}=j\omega C_{gd}$ = admittance corresponding to C_{gd}

$Y_{ds}=j\omega C_{ds}$ = admittance corresponding to C_{ds}

The voltage gain is given by

$$A_V=\frac{V_o}{V_i}=\frac{IZ}{V_i}=\frac{-g_m V_i+V_i(Y_{gd})}{V_i\times Y}$$

or $$A_V=\frac{-g_m+Y_{gd}}{Y_L+g_d+Y_{gd}+Y_{ds}} \qquad ...(4)$$

(ii) Input Admittance : Let the input signal supplies a current I_i which is divided into two parts *i.e.*, I_1 and I_2 as shown in fig. (16). Then

$$I_1=V_i\times Y_{gs} \text{ where } Y_{gs}=j\omega C_{gs} \qquad ...(5)$$

and I_2 = (Voltage across C_{gd}) $\times Y_{gd}$ where $Y_{gd}=j\omega C_{gd}$

$=(V_i+V_{SD})\times Y_{gd}=(V_i-V_{DS})\ Y_{gd}$

But $V_{DS}=A_V\times V_i$

$\therefore \quad I_2=(V_i-A_V\ V_i)\ Y_{gd}=(1-A_V)\ V_i.Y_{gd} \qquad ...(6)$

Now $I_i=I_1+I_2=[Y_{gs}+(1-A_V)\ Y_{gd}]\ V_i \qquad ...(7)$

The input admittance is given by

$$Y_i=\frac{I_i}{V_i}=\frac{[Y_{gs}+(1-A_V)Y_{gd}]\ V_i}{V_i}$$

or $$Y_i=Y_{gs}+(1-A_V)\ Y_{gd} \qquad ...(8)$$

It is obvious from eq. (8) that for a FET to possess negligible input admittance over a wide range of frequencies, the gate to source capacitance C_{gs} and gate drain capacitance C_{gd} must be negligible.

(iii) Input capacitance : Let the drain circuit impedance Z_L be a pure resistance Z_L. Now within the audio-frequency range, the gain of the amplifier is given by

$$A_V=\frac{-\mu\ R_L}{r_a+R_L}=-g_m\ R'_a \quad \text{where} \quad R'_a=\frac{r_a.R_L}{r_a+R_L} \qquad ...(9)$$

Now eq. (8) yields the following expression for input capacitance

$$C_i=\frac{Y_i}{j\omega}=C_{gs}+(1-A_V)\ C_{gd} \qquad ...(10)$$

This increase in input capacitance C_i as caused by the presence of the output-to-input capacitance C_{gd} is called the *Miller Effect*.

(iv) Output admittance : The output admittance is obtained by looking back from output terminals into drain with V_i made zero. Making $V_i=0$, the output admittance is a parallel combination of r_d, C_{ds} and C_{gd} (fig. 16). The output admittance is given by

$$Y_o=g_d+Y_{ds}+Y_{gd} \qquad ...(11)$$

17·8. COMMON DRAIN AMPLIFIER (SOURCE FOLLOWER) AT HIGH FREQUENCIES :

Fig. (17) shows the basic circuit of common drain amplifier taking into account the different inter-electrode capacitances at high frequencies without biasing arrangement In figure R_s is the load impedance. Fig. (18) shows the Norton's equivalent of this circuit for small signal.

(i) Voltage gain : The output voltage V_0 is given by

$$V_o=IZ \qquad ...(1)$$

where I=current in the lead short circuiting terminals S and N.

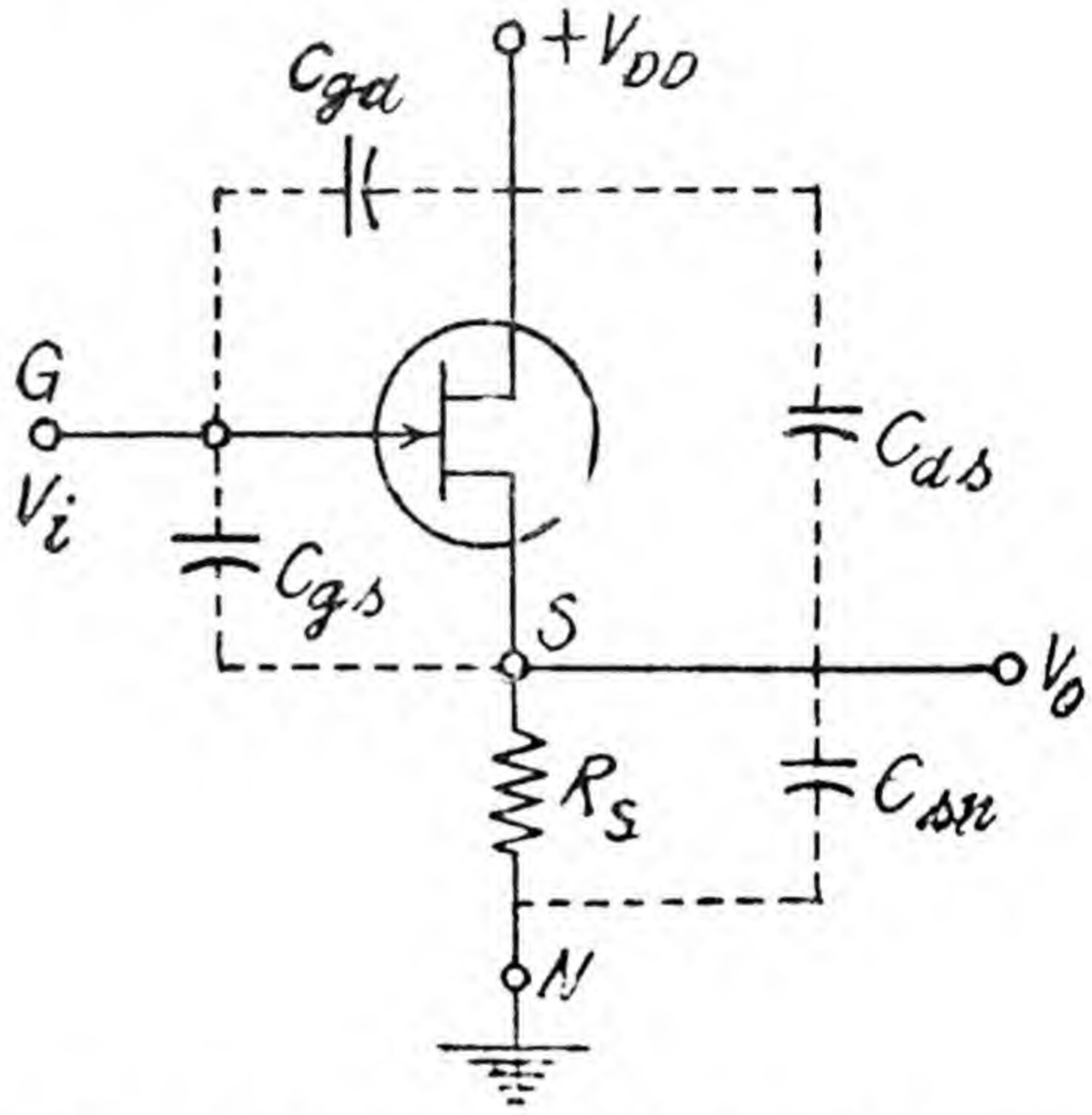


Fig. 17. Basic circuit of source follower with interelectrode capacitances.

Z=Impedance between terminals S and N on short circuiting all independent voltage sources and open circuiting all independent current sources.

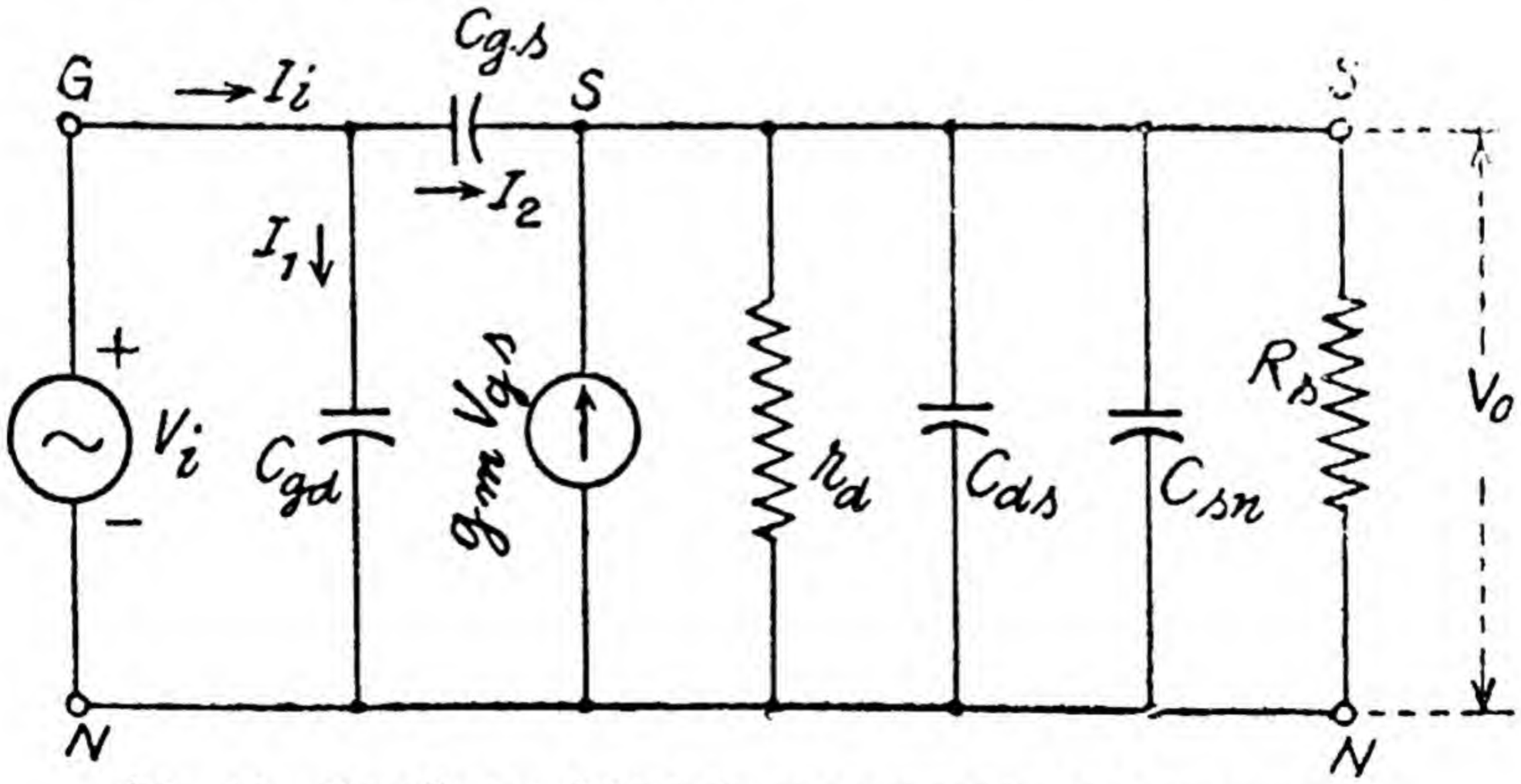


Fig. 18. Small signal high frequency equivalent circuit.

The current I is given by

$$I=g_m V_{gs}+V_i/(1/j\omega C_{gs})=g_m V_i+V_i (j\omega C_{gs})$$

or $$I=g_m V_{gs}+V_i Y_{gs} \text{ where } Y_{gs}=j\omega C_{gs} \quad (2)$$

The value of Z is the impedance of a parallel combination of R_s, C_{sn}, r_d, C_{ds} and C_{gs}. Hence

$$\frac{1}{Z}=\frac{1}{R_S}+\frac{1}{r_d}+\frac{1}{(1/j\omega C_{sn})}+\frac{1}{(1/j\omega C_{ds})}+\frac{1}{(1/j\omega C_{gs})}$$

or $$\frac{1}{Z}=\frac{1}{R_S}+\frac{1}{r_d}+j\omega C_{sn}+j\omega C_{ds}+j\omega C_{gs}$$

or $$Y=g_s+g_d+Y_{sn}+Y_{ds}+Y_{gs} \quad ...(3)$$

where $$Y=\frac{1}{Z},\ g_s=\frac{1}{R_S},\ g_d=\frac{1}{r_d}$$

$$Y_{sn}=j\omega C_{sn},\ Y_{ds}=j\omega C_{ds} \text{ and } Y_{gs}=j\omega C_{gs}$$

From fig. (18), $$V_{gs}=V_i-V_o \quad ...(4)$$

From eq. (1) and (2), we have

$$V_o=[g_m V_{gs}+V_i Y_{gs}] Z$$
$$=[g_m (V_i-V_o)+V_i Y_{gs}] Z \quad (\because V_{gs}=V_i-V_o)$$
$$=[g_m V_i-g_m V_o+V_i Y_{gs}] Z$$

Solving for, V_o, we get

$$V_o=\frac{V [g_m+Y_{gs}] Z}{[1+g_m Z]} \quad ...(5)$$

Hence the voltage gain is given by

$$A_V=\frac{V_o}{V_i}=\frac{g_m+Y_{gs}}{Y+g_m}$$

$$=\frac{g_m+j\omega C_{gs}}{(g_s+g_d+j\omega C_{sn}+j\omega C_{ds}+j\omega C_{gs})+g_m}$$

$$=\frac{g_m+j\omega C_{gs}}{(g_m+g_d+j\omega C_T)+(1/R_s)} \text{ where } C_T=C_{sn}+C_{ds}+C_{gs}$$

$$=\frac{(g_m+j\omega C_{gs}) R_s}{1+(g_m+g_d+j\omega C_T) R_s} \quad ...(6)$$

(ii) Input admittance : From fig. (18) we find that

$$I_1=V_i (j\omega C_{gd})$$

and $$I_2=(V_i-V_o) (j\omega C_{gs})=V_i (1-A_V) (j\omega C_{gs})$$

where A_V is the amplifier gain.

The input admittance is given by

$$Y_i=\frac{I}{V_i}=\frac{I_1+I_2}{V_i}$$

or $$Y_i=\frac{V_i (j\omega C_{gd})+V_i (1-A_V) (j\omega C_{gs})}{V_i}$$

or $$Y_i=j\omega C_{gd}+j\omega C_{gs} (1-A_V) \quad .. (7)$$

This expression shows that Y_i contains a resistive as well as a capacitive component. If $A_V \approx 1$, then

$$Y_i \approx j\omega C_{gd} \qquad \ldots(8)$$

(iii) Input capacitance : If the frequency is low enough so that A_V may be considered a real number, then input impedance consists of a capacitance C_i, then

$$Y_i = j\omega\, C_i$$

From equation (7)

$$j\omega\, C_i = j\omega\, C_{gd} + j\omega\, C_{gs}\,(1 - A_V)$$

or
$$C_i = C_{gd} + C_{gs}\,(1 - A_V) \qquad \ldots(9)$$

(iv) Output admittance : The output impedance can be found by looking back from the output terminals into the source when $V_i = 0$. From the equivalent circuit [fig. (18)], the output impedance consists of parallel combination of g_m, g_d $(= 1/r_d)$ and C_T. Thus the output admittance is given by

$$Y_o = g_m + g_d + j\omega\, C_T$$

where
$$C_T = C_{sn} + C_{ds} + C_{gs}.$$

Example 1. *A common source amplifier uses a MOSFET having following parameters :* $g_m = 1{\cdot}8$ *mA/volt;* $r_d = 50$ *k*Ω, $C_{gs} = 5$ *PF,* $C_{ds} = 2$ *PF and* $C_{gd} = 3$ *PF. The drain load resistance* $R_L = 150$ *k*Ω, *and the amplifier operates at* 50 *KHz. Calculate the voltage gain and input capacitance.*

Here
$$Y_{gs} = j\omega C_{gs} = j \times 2\pi \times (50 \times 10^3) \times 5 \times 10^{-12}$$
$$= j.\pi.50 \times 10^{-7} \text{ mho.}$$
$$Y_{ds} = j\omega C_{ds} = j \times 2\pi \times (50 \times 10^3) \times 2 \times 10^{-12}$$
$$= j \times \pi \times 2 \times 10^{-7} \text{ mho.}$$
$$Y_{gd} = j\omega C_{gd} = j \times 2\pi \times (50 \times 10^3) \times 3 \times 10^{-12}$$
$$= j \times \pi \times 3 \times 10^{-7} \text{ mho.}$$
$$g_d = \frac{1}{r_d} = \frac{1}{50 \times 10^3} = 2 \times 10^{-5} \text{ mho.}$$
$$Y_L = \frac{1}{R_L} = \frac{1}{150 \times 10^3} = 6{\cdot}66 \times 10^{-6} \text{ mho.}$$
$$g_m = 1{\cdot}8 \times 10^{-3} \text{ mho.}$$

The gain of the common source amplifier is given by

$$A_V = \frac{-g_m + Y_{gd}}{Y_L + g_d + Y_{gd} + Y_{ds}}$$
$$= \frac{-1{\cdot}8 \times 10^{-3} + j.\pi.3 \times 10^{-7}}{6{\cdot}66 \times 10^{-6} + 2 \times 10^{-5} + j\pi.3 \times 10^{-7} + j \times \pi \times 2 \times 10^{-7}}$$
$$= \frac{-1{\cdot}8 \times 10^{-3} + j.\pi.3 \times 10^{-7}}{2{\cdot}66 \times 10^{-5} + j.\pi.5 \times 10^{-7}}$$

The j terms are negligibly small in comparison with real terms. Neglecting the j terms, we have

$$= \frac{-1{\cdot}8 \times 10^{-3}}{2{\cdot}66 \times 10^{-5}} = -67{\cdot}67$$

The input capacitance is given by

$$C_i = C_{gs} + (1 - A_V)\, C_{gd}$$
$$= 5 + (1 + 67{\cdot}67)\, 3 = \mathbf{211{\cdot}01\ PF.}$$

Example 2. *A common drain amplifier uses MOSFET having the following parameters :* $g_m = 1{\cdot}5$ *mA/volt.* $r_d = 50\ K\Omega$, $C_{gs} = 3$ *PF,* $C_{ds} = 1$ *PF,* $C_{gd} = 2$ *PF and source to ground capacitance* $C_{sn} = 1$ *PF. The source circuit resistor* $R_s = 100\ K\Omega$. *The amplifier operates at 50 K. Hz. Calculate (i) voltage gain (ii) input admittance (iii) input capacitance and (iv) output admittance.*

Here

$$g_m = 1{\cdot}5 \times 10^{-3}\ \text{mho.},$$
$$g_d = 1/r_d = 1/(50 \times 10^3) = 2 \times 10^{-5}\ \text{mho},$$
$$Y_{gs} = j\omega C_{gs} = j.2\pi \times 50 \times 10^3 \times \cdot \times 10^{-12}$$
$$= j.\pi.3 \times 10^{-7}\ \text{mho.}$$
$$C_T = C_{sn} + C_{ds} + C_{gs} = 1 + 1 + 3 = 5\ \text{PF}$$
$$j\omega C_T = j\,(2\pi)\ 50 \times 10^{-3} \times 5 \times 10^{-12}$$
$$= j.\pi.5 \times 10^{-7}\ \text{mho.}$$

(i) Voltage gain $A_V = \dfrac{(g_m + j\omega C_{gs})\, R_s}{1 + (g_m + g_d + j\omega C_T)\, R_s}$

$$= \frac{(1{\cdot}5 \times 10^{-3} + j.\pi.3 \times 10^{-7})\ 100 \times 10^3}{1 + (1\ 5 \times 10^{-3} + 2 \times 10^{-5} + j.\pi.5 \times 10^{-7})\ 100 \times 10^3}$$
$$= \frac{1{\cdot}5 \times 10^{-3} + j.\pi.3 \times 10^{-7}}{10^{-5} + (150 \times 10^{-5} + 2 \times 10^{-5}) + j.\pi.5 \times 10^{-7}}$$

The j terms are extremely small and may be neglected. Hence

$$A_V = \frac{1{\cdot}5 \times 10^{-3}}{153 \times 10^{-5}} = \mathbf{0{\cdot}98}$$

(ii) Input Admittance $Y_L = j\omega C_{gd} + j\omega C_{gs}\ (1 - A_V \approx j\omega C_{gd}$

$$\approx j.2\pi.\ 50 \times 10^3 \times 2 \times 10^{-12}$$
$$\approx \mathbf{2}j\pi \times \mathbf{10^{-7}\ mho.}$$

(iii) Input capacitance $C_i = C_{gd} + C_{gs}\ (1 - A_V)$

$$= 2 + 3\ (1 - 0\ 98)\ \text{P.F.}$$
$$= \mathbf{2{\cdot}06\ PF.}$$

(iv) Output admittance $Y_o = g_m + g_d + j\omega C_T$

$$= 1{\cdot}5 \times 10^{-3} + 2 \times 10^{-5} + j.\pi.5 \times 10^{-7}$$
$$\approx \mathbf{152 \times 10^{-5}\ mho.}$$

———

18 TRANSISTOR OSCILLATORS

The theory of oscillations developed in vacuum tube oscillators is equally valid for transistorized oscillators. The network associated with the transistor determines the frequency of oscillations, while characteristics of transistor with circuit determine the conditions of oscillations. The oscillators may be classified as :

(*i*) sinusoidal oscillators, and

(*ii*) relaxation oscillators.

Sinusoidal oscillators are those oscillators which operate on the linear portion of its characteristics, whereas the relaxation oscillators operate on non-linear region of its characteristics. They are again classified as :

(*i*) feed back type,

(*ii*) negative resistance type.

In feed back oscillators, part of the output is fed back to the input in proper phase and magnitude. In negative resistance oscillators, the transistor provides the negative resistance which cancels the positive resistance of the associated circuit and thus provides for oscillations.

FEED BACK OSCILLATORS ;

18·1. TUNED COLLECTOR OSCILLATOR :

Figure (1*a*) shows the tuned collector oscillator using a junction transistor with an external feed back path. The oscillator contains a tuned circuit in the collector circuit. As in case of vacuum tube feed back oscillators, a part of the output voltage is fed back to the input ; similarly in transistor feed back oscillators, a part of the output current is fed back to the input.

The output developed across the tuned circuit is inductively coupled to the base circuit through L_2. The winding direction of the two coils are so adjusted that the positive feed back takes place from the collector circuit to the base circuit.

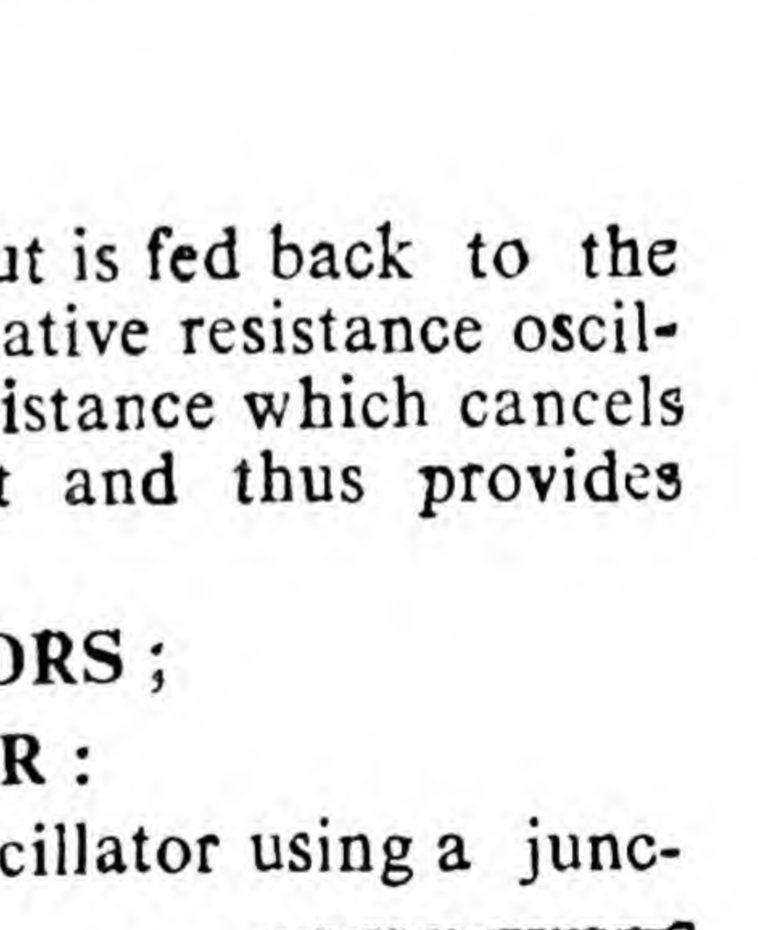


Fig. 1. (*a*) Tuned collector oscillator.

$$= \frac{-1{\cdot}8 \times 10^{-3}}{2{\cdot}66 \times 10^{-5}} = -\mathbf{67{\cdot}67}$$

The input capacitance is given by

$$C_i = C_{gs} + (1 - A_V)\, C_{gd}$$
$$= 5 + (1 + 67{\cdot}67)\, 3 = \mathbf{211{\cdot}01\ PF.}$$

Example 2. *A common drain amplifier uses MOSFET having the following parameters :* $g_m = 1{\cdot}5$ *mA/volt.* $r_d = 50\ K\Omega$, $C_{gs} = 3$ *PF,* $C_{ds} = 1$ *PF,* $C_{gd} = 2$ *PF and source to ground capacitance* $C_{sn} = 1$ *PF. The source circuit resistor* $R_s = 100\ K\Omega$. *The amplifier operates at 50 K. Hz. Calculate (i) voltage gain (ii) input admittance (iii) input capacitance and (iv) output admittance.*

Here

$$g_m = 1{\cdot}5 \times 10^{-3}\ \text{mho.},$$
$$g_d = 1/r_d = 1/(50 \times 10^3) = 2 \times 10^{-5}\ \text{mho},$$
$$Y_{gs} = j\omega C_{gs} = j.2\pi \times 50 \times 10^3 \times \cdot \times 10^{-12}$$
$$= j.\pi.3 \times 10^{-7}\ \text{mho.}$$
$$C_T = C_{sn} + C_{ds} + C_{gs} = 1 + 1 + 3 = 5\ \text{PF}$$
$$j\omega C_T = j\ (2\pi)\ 50 \times 10^{-3} \times 5 \times 10^{-12}$$
$$= j.\pi.5 \times 10^{-7}\ \text{mho.}$$

(i) Voltage gain $A_V = \dfrac{(g_m + j\omega C_{gs})\ R_s}{1 + (g_m + g_d + j\omega C_T)\ R_s}$

$$= \frac{(1{\cdot}5 \times 10^{-3} + j.\pi.3 \times 10^{-7})\ 100 \times 10^3}{1 + (1\ 5 \times 10^{-3} + 2 \times 10^{-5} + j.\pi.5 \times 10^{-7})\ 100 \times 10^3}$$
$$= \frac{1{\cdot}5 \times 10^{-3} + j.\pi.3 \times 10^{-7}}{10^{-5} + (150 \times 10^{-5} + 2 \times 10^{-5}) + j.\pi.5 \times 10^{-7}}$$

The j terms are extremely small and may be neglected. Hence

$$A_V = \frac{1{\cdot}5 \times 10^{-3}}{153 \times 10^{-5}} = \mathbf{0{\cdot}98}$$

(ii) Input Admittance $Y_L = j\omega C_{gd} + j\omega C_{gs}\ (1 - A_V \approx j\omega C_{gd}$

$$\approx j.2\pi.\ 50 \times 10^3 \times 2 \times 10^{-12}$$
$$\approx \mathbf{2}j\pi \times \mathbf{10^{-7}\ mho.}$$

(iii) Input capacitance $C_i = C_{gd} + C_{gs}\ (1 - A_V)$

$$= 2 + 3\ (1 - 0\ 98)\ \text{P.F.}$$
$$= \mathbf{2{\cdot}06\ PF.}$$

(iv) Output admittance $Y_o = g_m + g_d + j\omega C_T$

$$= 1{\cdot}5 \times 10^{-3} + 2 \times 10^{-5} + j.\pi.5 \times 10^{-7}$$
$$\approx \mathbf{152 \times 10^{-5}\ mho.}$$

18 TRANSISTOR OSCILLATORS

The theory of oscillations developed in vacuum tube oscillators is equally valid for transistorized oscillators. The network associated with the transistor determines the frequency of oscillations, while characteristics of transistor with circuit determine the conditions of oscillations. The oscillators may be classified as :

(*i*) sinusoidal oscillators, and

(*ii*) relaxation oscillators.

Sinusoidal oscillators are those oscillators which operate on the linear portion of its characteristics, whereas the relaxation oscillators operate on non-linear region of its characteristics. They are again classified as :

(*i*) feed back type,

(*ii*) negative resistance type.

In feed back oscillators, part of the output is fed back to the input in proper phase and magnitude. In negative resistance oscillators, the transistor provides the negative resistance which cancels the positive resistance of the associated circuit and thus provides for oscillations.

FEED BACK OSCILLATORS ;

18·1. TUNED COLLECTOR OSCILLATOR :

Figure (1*a*) shows the tuned collector oscillator using a junction transistor with an external feed back path. The oscillator contains a tuned circuit in the collector circuit. As in case of vacuum tube feed back oscillators, a part of the output voltage is fed back to the input ; similarly in transistor feed back oscillators, a part of the output current is fed back to the input.

The output developed across the tuned circuit is inductively coupled to the base circuit through L_2. The winding direction of the two coils are so adjusted that the positive feed back takes place from the collector circuit to the base circuit.

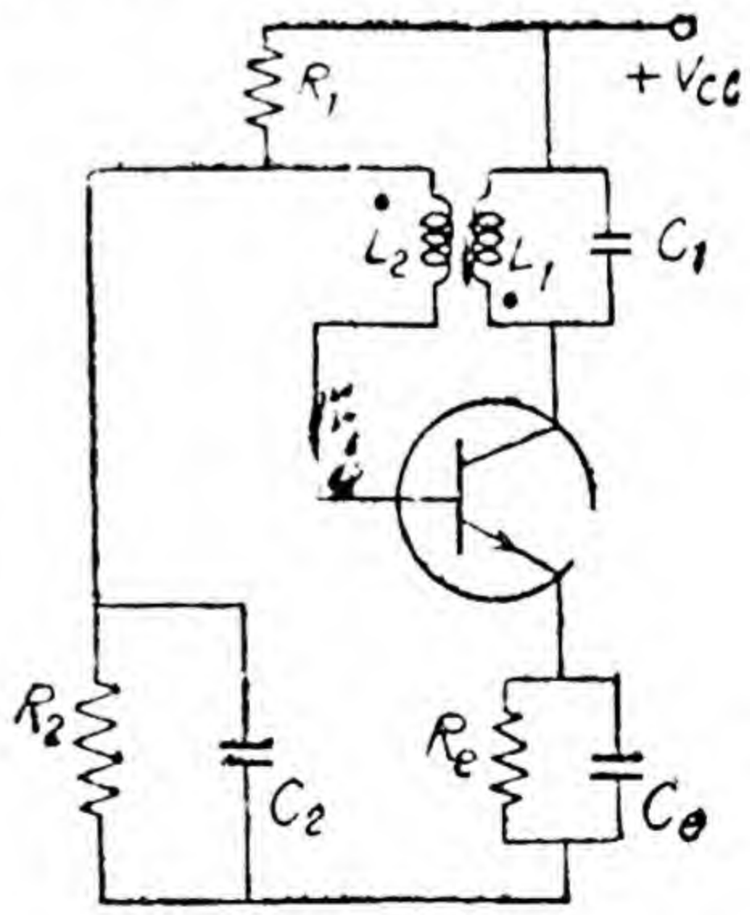


Fig. 1. (*a*) Tuned collector oscillator.

Analysis : Figure (1 *b*) shows the equivalent circuit of the tuned collector oscillator using *CE* hybrid model. The transformer winding resistances have been neglected It is also assumed that $h_{re}=0$. R is the resistance of the secondary reflected in the primary.

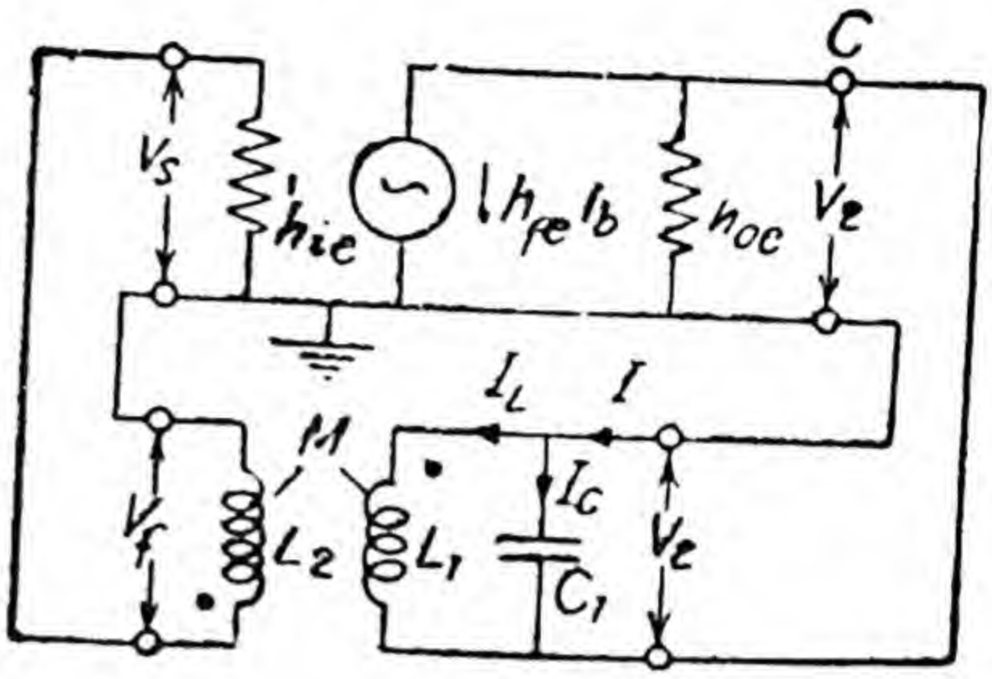


Fig. 1. (*b*) Equivalent circuit of tuned collector oscillator using *CE* hybrid model

The feed back voltage V_f is given by

$$V_f=-j\omega M I_L. \qquad \ldots(1)$$

The voltage V_2 across the tuned circuit is given by

$$V_2=-(R+j\omega L_1)\ I_L=-\frac{I_c}{j\omega C_1} \qquad \ldots(2)$$

Hence for feed back network,

$$\beta=\frac{V_f}{V_2}=\frac{-j\omega M I_L}{-(R+j\omega L_1)I_L}=\frac{j\omega M}{R+j\omega L_1}. \qquad \ldots(3)$$

Further

$$V_S=-V_f. \qquad \ldots(4)$$

The current I is given by

$$I=h_{fe}\ I_b+h_{oe}\ V_2. \qquad \ldots(5)$$

Then transfer impedance Z_T is given by

$$Z_T=\frac{V_f}{I}=\frac{-j\omega M I_L}{I_L+I_c}=\frac{-j\omega M I_L}{I_L+I_L\ (R+j\omega L_1)}. \qquad \ldots(6)$$

Also

$$I_b\approx\frac{V_S}{h_{ie}}. \qquad \ldots(7)$$

Equation (5) may be put as

$$-\frac{V_S}{Z_T}=h_{fe}.\frac{V_S}{h_{ie}}+h_{oe}\ V_2$$

or

$$V_2=-\frac{V_S}{Z_T\ h_{oe}}-\frac{h_{fe}\ V_S}{h_{ie}\ h_{oe}}.$$

Hence the gain of the amplifier is

$$A=\frac{V_2}{V_S}=-\left[\frac{h_{fe}}{h_{ie}.h_{oe}}+\frac{1}{Z_T\ h_{oe}}\right]. \qquad \ldots(8)$$

For sustained oscillations, according to Barkhausen criterion, $-A\beta$ must equal one.

Hence from equations (3) and (8), we get

$$\frac{j\omega M}{(R+j\omega L_1)}\left[\frac{h_{fe}}{h_{ie}.h_{oe}}+\frac{1}{Z_T\ h_{oe}}\right]=1. \qquad \ldots(9)$$

Substituting the value of Z_T from equation (6), we have

$$\frac{j\omega M}{(R+j\omega L_1)}\left[\frac{h_{fe}}{h_{ie}.h_{oe}}+\frac{1}{h_{oc}\left\{\frac{-j\omega M}{1+(R+j\omega L_1)\,j\omega C_1}\right\}}\right]=1$$

or $$\frac{j\omega M}{(R+j\omega L_1)}\left[\frac{h_{fe}}{h_{ie}.h_{oe}}-\frac{\{1+(R+j\omega L_1)\;j\omega C_1\}}{h_{oe}.\;j\omega M}\right]=1$$

or $$\frac{j\omega M.h_{fe}}{h_{ie}.h_{oe}}-\frac{1}{h_{oe}}-\frac{R.j\omega C_1}{h_{oe}}-\frac{j\omega L_1\times j\omega C_1}{h_{oe}}=R+j\omega L_1.$$

Multiplying by $h_{oe}/j\omega C_1$ on both sides, we have

$$\frac{h_{fe}}{h_{ie}}.\frac{M}{C_1}-\frac{1}{j\omega C_1}-R-j\omega L_1=\frac{(R+j\omega L_1)\,h_{oe}}{j\omega C_1}$$

or $$\frac{h_{fe}}{h_{ie}}.\frac{M}{C_1}=R+j\left(\omega L_1-\frac{1}{\omega C_1}\right)+\left[\frac{-jh_{oe}R}{\omega C_1}+\frac{L_1\,h_{oe}}{C_1}\right]. \qquad ...(10)$$

Equating the real parts in equation (10), we get

$$\frac{h_{fe}\,M}{h_{ie}\,C_1}=R+\frac{L_1\,h_{oe}}{C_1}$$

or $$\frac{h_{fe}}{h_{ie}}=\frac{RC_1+h_{oe}\,L_1}{M} \qquad ...(11)$$

Equation (11) gives the condition for sustained oscillations. Equating the imaginary part in equation (10), we get

$$\omega L_1-\frac{1}{\omega C_1}-\frac{h_{oe}\,R}{\omega C_1}=0$$

or $$\omega^2=\frac{1}{L_1C_1}[1+h_{oe}\,R]$$

$$=\omega_0^2\,(1+h_{oe}\;R). \qquad ...(12)$$

Equation (12) gives the frequency of oscillations. This equation shows that the frequency of oscillation exceeds the frequency of resonance of the tuned circuit.

18·2. TUNED EMITTER OSCILLATOR :

Figure (2*a*) and (2*b*) show the tuned emitter oscillator and its equivalent circuit.

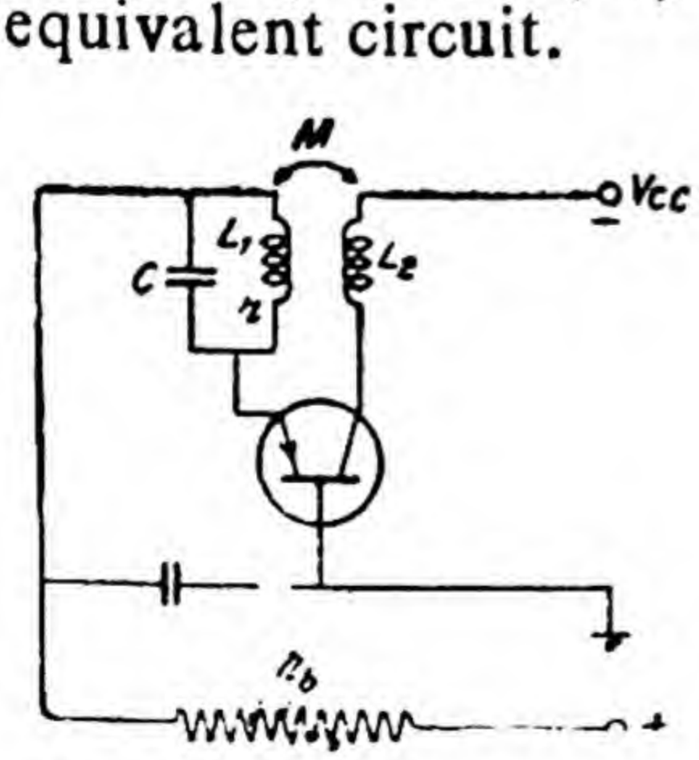


Fig. 2 (*a*). Tuned emitter oscillator.

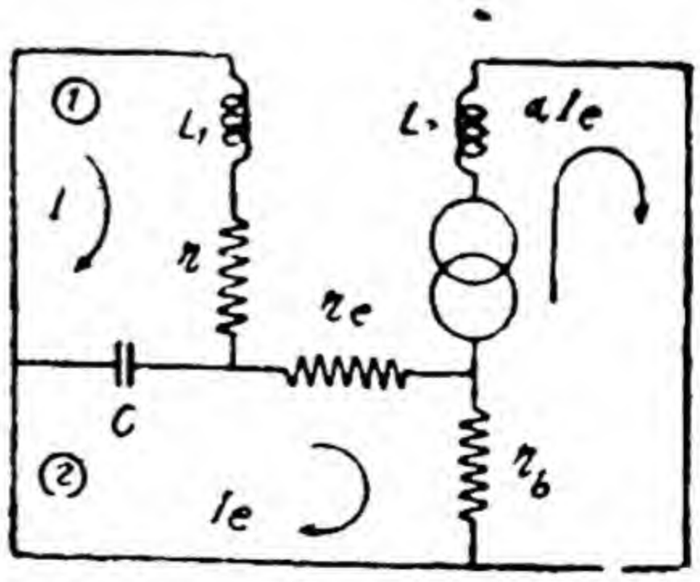


Fig. 2 (*b*) Equivalent circuit.

In the analysis, the effect of bias circuit is ignored. The collector resistance being high and the resistance of coil L_2 being low, their effect are neglected.

The loop potential equations for meshes (1) and (2) are

$$I\,(r+j\omega L_1+1/j\omega C)-I_e/j\omega C\pm j\omega M\alpha I_e=0 \qquad ...(1)$$

and $$I_e\,(r_c+r_b+1/j\omega C)-I/j\omega C-\alpha I_c r_b=0. \qquad ...(2)$$

Putting $R=r_c+r_b\,(1-\alpha)$, in equation (2), we have

$$I\,(r+j\omega L_1+1/j\omega C=I_e\,[1/j\omega C\mp j\omega M\alpha], \qquad ...(3)$$

and $$1/j\omega C=I_e\,[R+1/j\omega C]. \qquad ...(4)$$

From equations (3) and (4), we have

$$\frac{1}{j\omega C}\left(\frac{1}{j\omega C}\mp j\omega M\alpha\right)=\left(R+\frac{1}{j\omega C}\right)\left(r+j\omega L_1+\frac{1}{j\omega C}\right)$$

Equating real parts

$$-\frac{1}{\omega^2C^2}\mp\frac{M\alpha}{C}=Rr+L_1/C-\frac{1}{\omega^2C^2}.$$

Solving it, we have

$$M=\pm[(CRr/\alpha)+(L_1/\alpha)].$$

This is the maintenance condition and $\omega^2=1/(L_1C)$.

Equating the imaginary parts,

$$-\frac{r}{\omega C}+\omega L_1R-R/\omega C=0$$

$$\therefore \quad \omega^2L_1CR=R+r$$

so

$$f=\frac{1}{2\pi L_1C}\sqrt{\left(1+\frac{r}{R}\right)}.$$

This is the generated frequency.

18·3. HARTLEY OSCILLATOR :

Figure 3 (*a*) shows the circuit of Hartley oscillator using a transistor. The collector supply voltage is appled to the collector

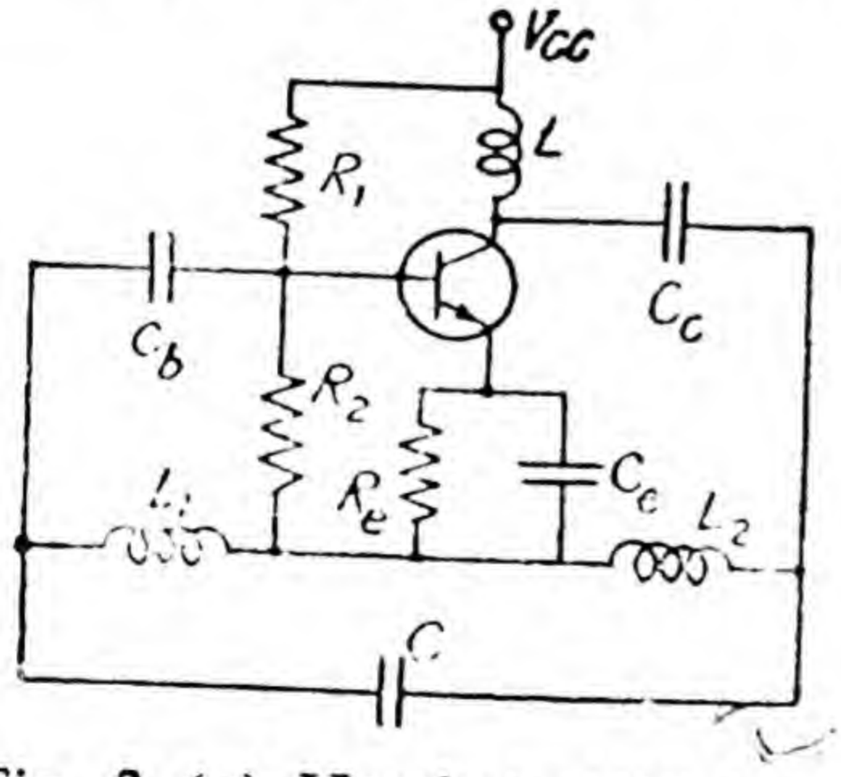


Fig. 3. (*a*) Hartley oscillator.

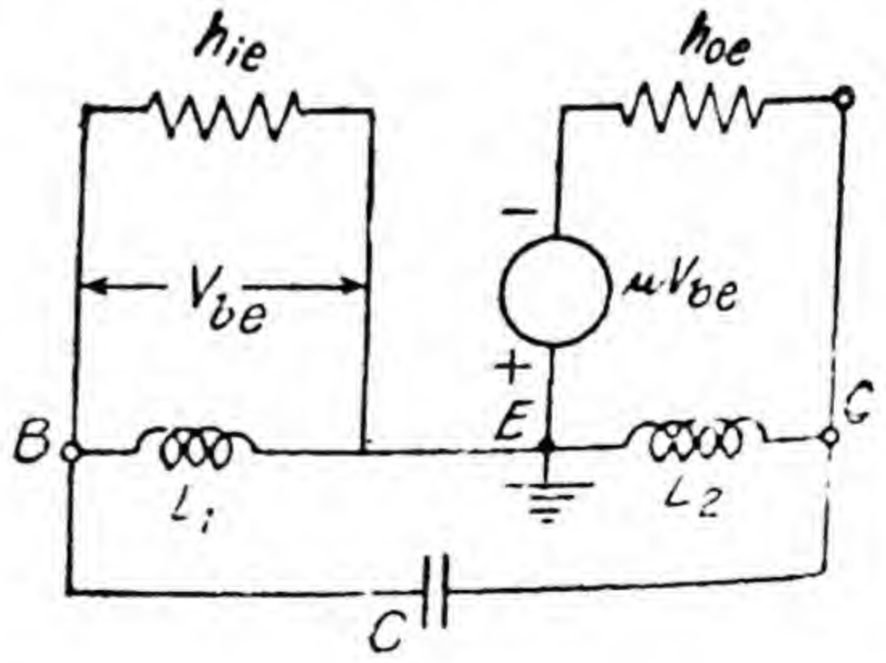


Fig. 3 (*b*) Equivalent circuit of Hartley oscillator.

through inductor L. The reactance of L is higher than L_2 and hence may be omitted from the equivalent circuit. The capacitor C_c has a low reactance at the frequency of oscillation and also may be omitted from the equivalent circuit. However, C_c acts as an open circuit at zero frequency. Therefore, it prevents collector

supply V_{cc} from getting short circuited by L in series with L_2. The parallel combination of R_e and C_e in conjunction with R_1 and R_2 combination provides stabilized self bias. The frequency of oscillation is determined by the tuned circuit L_1-L_2 and C. Since the circuit extends from the collector circuit to the base circuit, the feed back needed for sustained oscillations takes place through the tuned circuit itself.

Analysis :

The equivalent circuit of Hartley oscillator is shown in fig. (3*b*). Now we proceed to calculate the frequency of oscillation and the value of transistor parameters needed for sustained oscillations.

The effective load impedance Z_L is given by

$$Z_L = \frac{j\omega L_2 \left(\frac{1}{j\omega C} + \frac{h_{ie}\, j\omega L_1}{h_{ie} + j\omega L_1} \right)}{j\omega L_2 + \frac{1}{j\omega C} + \frac{h_{ie}.\, j\omega L_1}{h_{ie} + j\omega L_1}} \quad \ldots(1)$$

The voltage gain without feed back is given by

$$A = \frac{-\mu Z_L}{j\omega L_1 + h_{oe}} \quad (2)$$

$$\beta = \frac{\frac{-j\omega L_1 . h_{ie}}{j\omega L_1 + h_{ie}}}{\frac{j\omega L_1 . h_{ie}}{j\omega L_1 + h_{ie}} + \frac{1}{j\omega C}} \quad \ldots(3)$$

$$\therefore \quad -A\beta = \frac{-\mu Z_L}{j\omega L_1 + h_{oe}} \times \frac{j\omega L_1 . h_{ie}/(h_{ie} + j\omega L_1)}{\frac{j\omega L_1\, h_{ie}}{j\omega L_1 + h_{ie}} + \frac{1}{j\omega C}} \quad \ldots(4)$$

From equation (1), substituting the value of Z_L and rearranging, we get

$$-A\beta = \frac{\mu\omega^2\, L_1 L_2\, h_{ie}}{h_{oe}\left[h_{ie}\left(\omega L_1 + \omega L_2 - \frac{1}{\omega C} \right) - \frac{\omega L_1 . \omega L_2\, (-1/\omega C)}{h_{ie}} \right]} \quad \ldots(5)$$

$$-h_{oe}\, \omega L_1 \left(\omega L_2 - \frac{1}{\omega C} \right) - \omega L_2 h_{ie} \left(\omega L_1 - \frac{1}{\omega C} \right)$$

The loop gain to be real, the imaginary term in the denominator must be zero.

$$\therefore \quad h_{ie} \left(\omega L_1 + \omega L_2 - \frac{1}{\omega C} \right) = \frac{\omega L_1\, \omega L_2 \left(-\frac{1}{\omega C} \right)}{h_{oe}}$$

or $$\omega L_1 + \omega L_2 - \frac{1}{\omega C} = \frac{\omega L_1\, \omega L_2 \left(-\frac{1}{\omega C} \right)}{h_{oe} . h_{ie}}$$

or $$\omega^2=\frac{1}{(L_1+L_2)\,C}\left[1-\frac{\omega^2 L_1 L_2}{h_{ie}.h_{oe}}\right]$$

$$\therefore \quad \omega=\omega_0\sqrt{\left(1-\frac{X_1X_2}{h_{ie}.h_{oe}}\right)}, \qquad \ldots(6)$$

where $\omega L_1=X_1$ and $\omega L_2=X_2$.

Thus the frequency of oscillation ω is slightly less than the frequency of series resonance ω_0 of L_1, L_2 and C.

Equating imaginary term to zero in equation (5) and applying Barkhausen criterion ($-A\beta=1$) for sustained oscillations, we get

$$-1=\frac{\mu h_{ie}\omega L_1.\,\omega L_2}{h_{oe}\,\omega L_1\left(\omega L_2-\frac{1}{\omega C}\right)+h_{ie}\,\omega L_2\left(\omega L_1-\frac{1}{\omega C}\right)} \qquad \ldots(7)$$

But $$\omega L_1+\omega L_2-\frac{1}{\omega C}\approx 0.$$

$$\therefore \quad -1\approx\frac{\mu h_{ie}\,\omega L_1\,\omega L_2}{h_{ce}\,\omega L_1(-\omega L_1)+h_{ie}\,\omega L_2\,(-\omega L_2)}$$

$$\approx\frac{h_{fe}\,h_{oe}\,\omega L_1.\omega L_2}{h_{ce}\,(-\omega^2 L_1^{\,2})+h_{ie}\,(-\omega^2 L_2^{\,2})}$$

$$\approx\frac{h_{fe}\,h_{oe}\omega^2\,L_1L_2}{-h_{oe}\,\omega^2 L_1^{\,2}}$$ since $h_{ce}\,(-\omega^2 L_1^{\,2})\gg h_{ie}\,(-\omega^2 L_2^{\,2})$

or $$h_{fe}=\frac{\omega L_1}{\omega L_2}=\frac{L_1}{L_2}. \qquad \ldots(8)$$

Equation (8) gives the condition for sustained oscillations. If there exists mutual inductance M between coils L_1 and L_2, then equation (8) gets modified as below

$$h_{fe}=\frac{L_1+M}{L_2+M}. \qquad \ldots(9)$$

18·4. COLPITT'S OSCILLATOR :

Figure 4 (*a*) shows the circuit diagram of a Colpitt oscillator.

The circuit of Colpitt oscillator is the same as that of Hartley oscillator except that the emitter tap is connected between the capacitances C_1 and C_2. The frequency of the oscillator is determined by the resonance of the tuned circuit but this is affected by transistor parameters and load resistance. We shall now consider the effect of the

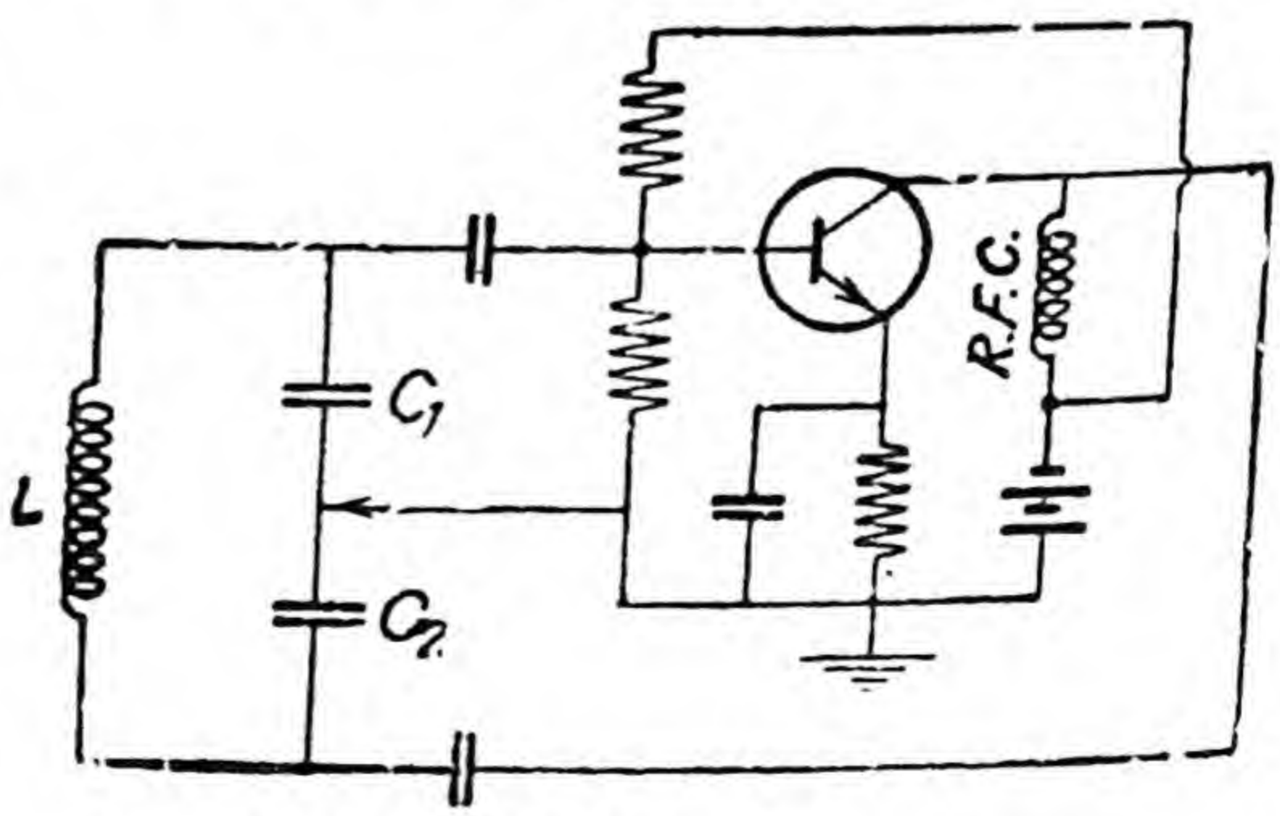


Fig. 4 (*a*) Colpitt's oscillator.

transistor parameters on the frequency of the oscillator.

Figure 4 (*b*) shows the equivalent circuit of Colpits oscillator.

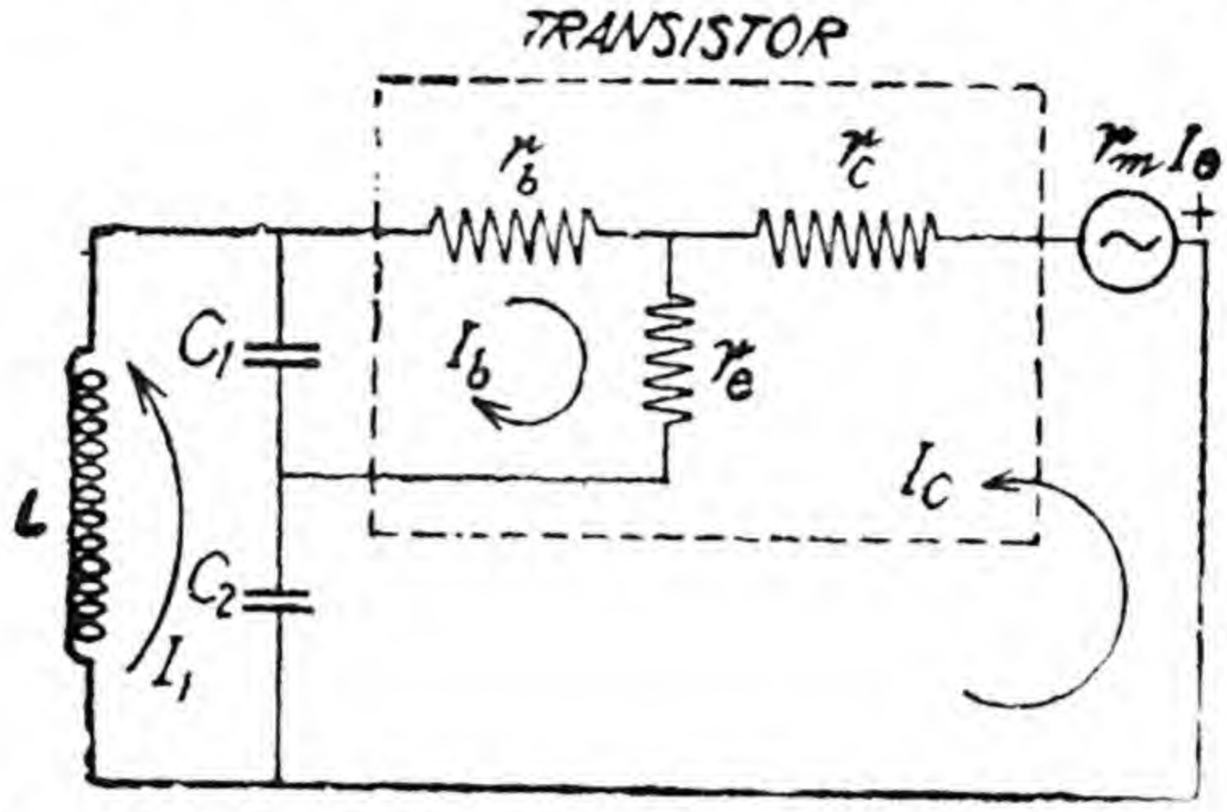


Fig. 4 (*b*) Equivalent circuit of Colpits oscillator.

Let $C_s=\dfrac{C_1C_2}{C_1+C_2}$, then the circuit equations may be written as

$$j\left(\omega L-\frac{1}{\omega C_s}\right)I_1-\frac{j}{\omega C_1}\ I_b+\frac{j}{\omega C_1}\ I_c=0, \qquad \ldots(1)$$

$$\frac{-j}{\omega C_1}\ I_1+\left(r_b+r_e-\frac{j}{\omega C_1}\right)I_b+r_eI_c=0, \qquad \ldots(2)$$

$$\frac{j}{\omega C_2}\ I_1+(r_e-r_m)\ I_b+\left(r_e+r_c-r_m-\frac{j}{\omega C_2}\right)I_c=0. \qquad \ldots(3)$$

For oscillations to occur, the determinant of the above equations must be zero. Equating it to zero, the real and imaginary terms of the determinant separately, we obtain the feed back condition and operating frequency. Hence,

$$\begin{vmatrix} j\left(\omega L-\dfrac{1}{\omega C_s}\right) & \dfrac{-j}{\omega C_1} & \dfrac{j}{\omega C_2} \\ -\dfrac{j}{\omega C_1} & \left(r_b+r_e-\dfrac{j}{\omega C_1}\right) & r_e \\ \dfrac{j}{\omega C_2} & (r_e-r_m) & \left(r_e+r_c-r_m-\dfrac{j}{\omega C_2}\right) \end{vmatrix} =0 \ldots(4)$$

$$j\left(\omega L-\frac{1}{\omega C_s}\right)\left\{\left(r_b+r_e-\frac{j}{\omega C_1}\right)\left(r_e+r_c-r_m-\frac{j}{\omega C_2}\right)-r_e\,(r_e-r_m)\right\}$$
$$-\frac{j}{\omega C_1}\left\{(r_e)\left(\frac{j}{\omega C_2}\right)-\left(\frac{-j}{\omega C_1}\right)\left(r_e+r_c-r_m-\frac{j}{\omega C_1}\right)\right\}$$
$$+\frac{j}{\omega C_2}\left\{\left(\frac{-j}{\omega C_1}\right)\ (r_e-r_m)-\frac{j}{\omega C_2}\left(r_b+r_e-\frac{j}{\omega C_1}\right)\right\}=0. \qquad \ldots(5)$$

Equating real part to zero, we get

$$\left\{\left(\omega L-\frac{1}{\omega C_s}\right)\left(\frac{r_b+r_e}{\omega C_2}+\frac{r_e+r_c-r_m}{\omega C_1}\right)\right\}+\left\{\frac{r_e}{\omega^2C_1C_2}+\frac{r_e+r_c-r_m}{\omega^2C_1^2}\right\}$$
$$+\left\{\frac{r_c-r_m}{\omega^2C_1C_2}+\frac{r_b+r_e}{\omega^2C_2^2}\right\}=0$$

$$\text{or } \left(\omega L-\frac{1}{\omega C_1}\right)\left(\frac{r_b+r_e}{\omega C_2}+\frac{r_e+r_c-r_m}{\omega C_1}\right)+\frac{r_e+r_c-r_m}{\omega^2 C_1^2}+\frac{r_b+r_e}{\omega^2 C_2^2}+\frac{2r_e-r_m}{\omega^2 C_1 C_2}=0. \quad \ldots(6)$$

The frequency at which oscillations start is very near to the resonant frequency of resonant circuit.

i.e. $\omega L-\frac{1}{\omega C_s}=0$. It may also be assumed that $r_e \ll r_m$.

The equation (6), then becomes

$$\frac{r_e+r_c-r_m}{\omega^2 C_1^2}+\frac{r_b+r_e}{\omega^2 C_2^2}+\frac{2r_e-r_m}{\omega^2 C_1 C_2}=0$$

$$\text{or} \quad (r_b+r_e)\left(\frac{C_1}{C_2}\right)^2-r_m\left(\frac{C_1}{C_2}\right)+(r_e-r_m)=0 \quad \ldots(7)$$

since $r_e \ll r_m$.

Equation (7) is quadratic, hence

$$\frac{C_1}{C_2}=\frac{r_m+\sqrt{\{(r_m^2-4\,(r_b+r_e)(r_e-r_m)\}}}{2\,(r_b+r_e)} \quad \ldots(8)$$

The second term in square root is small in comparison to r_m.

$$\therefore \quad \frac{C_1}{C_2} \geqslant \frac{r_m+r_m}{2(r_b+r_e)} \geqslant \frac{r_m}{(r_b+r_e)}. \quad \ldots(9)$$

Equating the imaginary part to zero, we get

$$\left(\omega L-\frac{1}{\omega C_s}\right)\left\{(r_b+r_e)\,(r_e+r_c-r_m)-\frac{1}{\omega^2 C_1 C_2}\right\}-\frac{1}{\omega^3 C_1^2 C_2}-\frac{1}{\omega^3 C_1 C_2^2}=0$$

$$\text{or } \left(\omega L-\frac{1}{\omega C_s}\right)\left(A-\frac{1}{\omega^2 C_1 C_2}\right)-\frac{1}{\omega^3 C_1 C_2}\left(\frac{1}{C_1}+\frac{1}{C_2}\right)=0, \quad \ldots(10)$$

where $A=(r_b+r_e)\,(r_e+r_c-r_m)$.

From equation (10), we get

$$f=\frac{1}{2\pi}\sqrt{\left(\frac{1}{LC_s}+\frac{1}{AC_1C_2}\right)}. \quad \ldots(11)$$

Equation (11) gives the frequency of oscillations.

In order to obtain a good frequency stability, crystal controlled oscillator is used. The crystal controlled oscillator is similar to colpitt's oscillator except that the inductance is replaced by a crystal. Figure 5 shows the circuit diagram of a crystal oscillator.

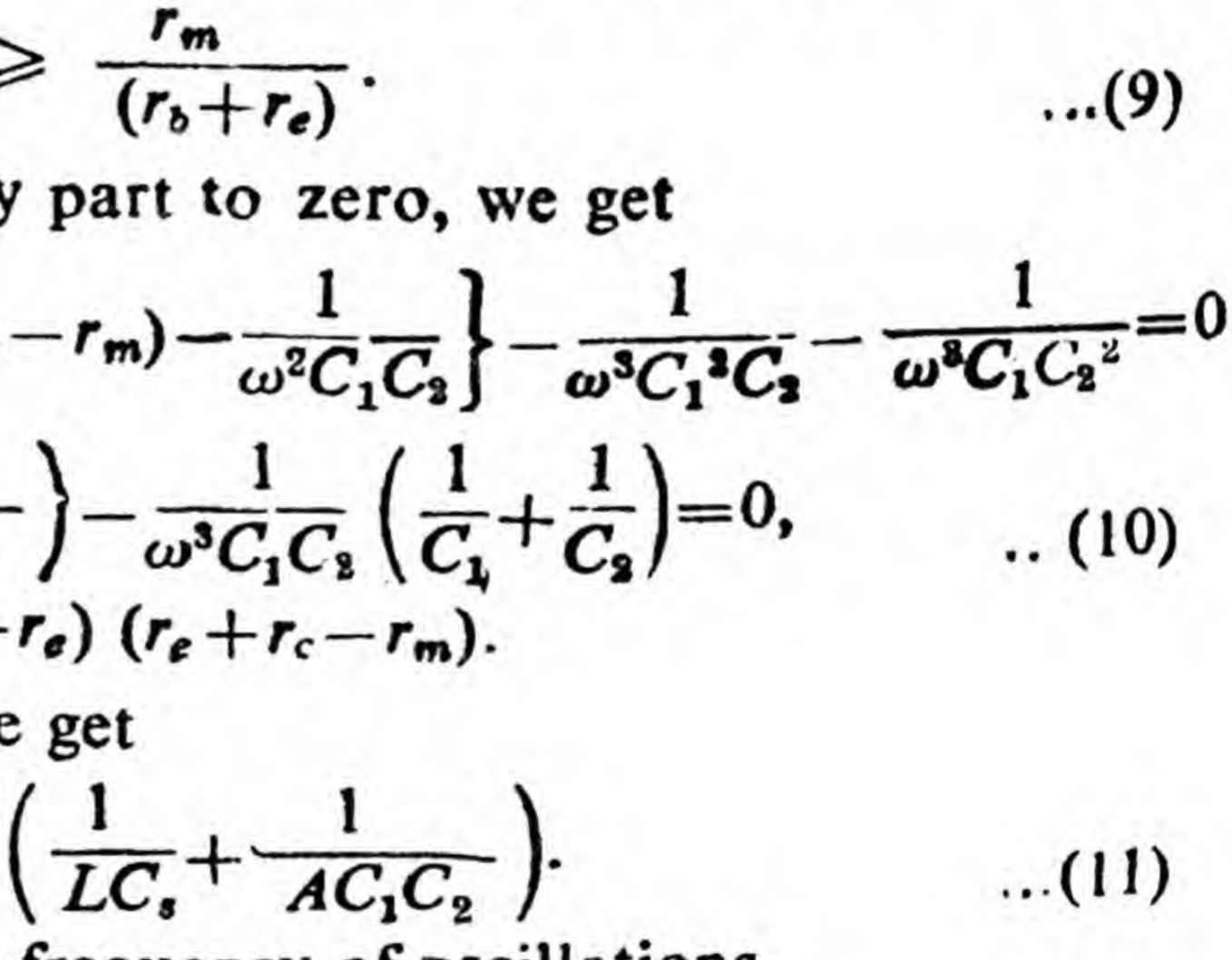


Fig. 5. Crystal controlled oscillator.

18·5. PHASE SHIFT OSCILLATORS :

18·5-1. RC OSCILLATOR :

The phase shift transistor oscillator is similar to vacuum tube phase shift oscillator. Figure 6 shows the circuit diagram of phase shift oscillator. To obtain a positive feed back essential for oscillations, the frequency determining circuit must introduce a phase change of 180°. This phase shift of 180° is obtained with three cascade sections C_1R_1, C_2R_2, C_3R_3 (eacn section consists of a series coupling capacitor and a shunt resistor) each shifting the signal by 60°. The phase shift comes about because R and C provides a current which leads the applied voltage by certain angle. The smaller is the capacitance more will the current leac the voltage for a given resistance. With a proper choice of R and C, a phase shift of 60° per section is achieved.

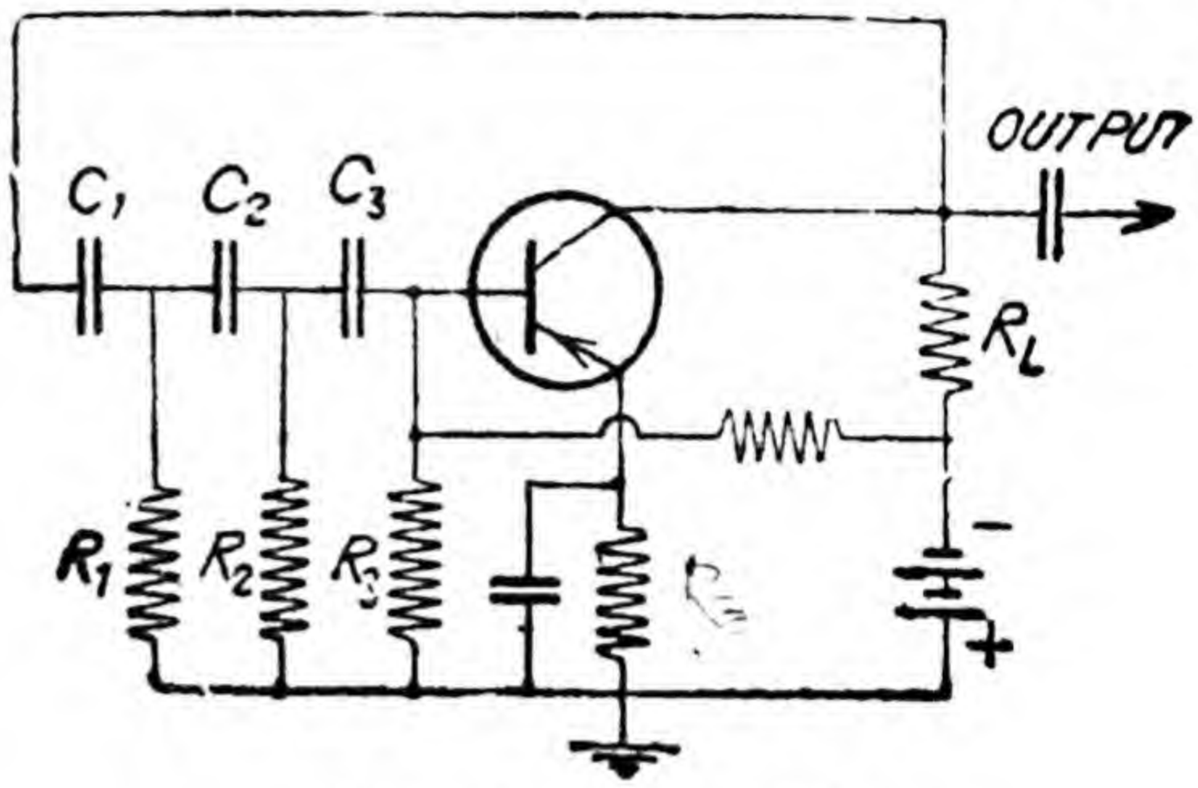


Fig. 6. *RC* oscillator.

The problem of a phase shift network for a transisor is somewhat complicated in comparison to vacuum tube because of low input impedance of the transistor. The last resistance in *RC* combination is not simply R_3 but the transistor input resistance in parallel with R_3.

Normally the frequency determining resistors are equal and similarly the frequcncy determining capacitors are equal, *i.e.*

$$R_1=R_2=R_3=R \text{ (say)}$$

and

$$C_1=C_2=C_3=C \text{ (say).}$$

In order to consider the frequency of oscillation and attenuation of the network we proceed as follows: Figure 7 shows the equivalent circuit of figure 6 in which I_1 is the signal current from the collector circuit and I_4 is the signal current into the base circuit.

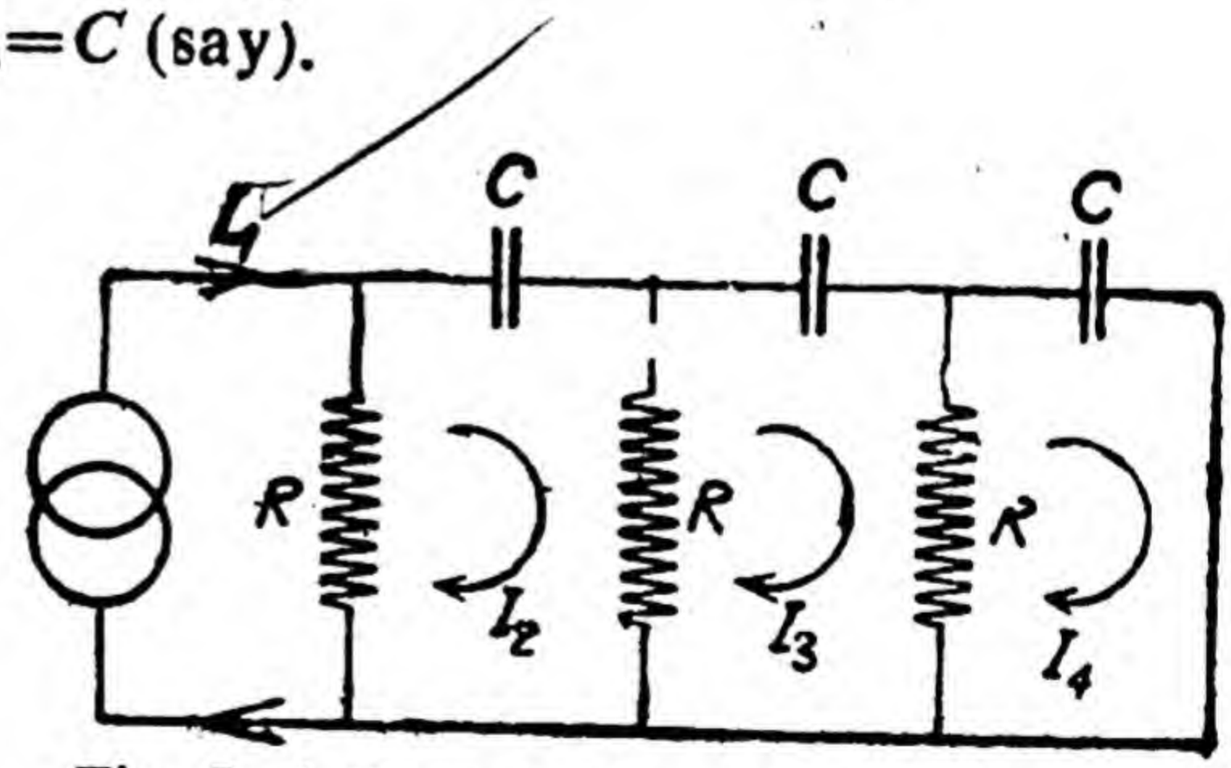


Fig. 7. A three section *RC* network.

Applying Kirchoff's law, we have

$$I_1R=I_2\left(2R-\frac{j}{\omega C}\right)-I_3R, \qquad \ldots(1)$$

$$I_2R=I_3\left(2R-\frac{j}{\omega C}\right)-I_4R \qquad \ldots(2$$

$$I_3R=I_4\left(R-\frac{j}{\omega C}\right). \qquad \ldots(3)$$

Eliminating I_2 from equations (1) and (2) *i.e.* substituting the value of I_2 from equation (2) into equation (1), we have

$$I_1R=\left\{\frac{I_3(2R-j/\omega C)}{R}-I_4\right\}\left(2R-\frac{j}{\omega C}\right)-I_3R$$

or
$$I_1R=\frac{\left(2R-\frac{j}{\omega C}\right)^2}{R}I_3-\left(2R-\frac{j}{\omega C}\right)I_4-I_3R. \qquad \ldots(4)$$

Substituting the value of I_3 from equation (3) in equation (4), we get

$$I_1R=\left\{\frac{\left(2R-\frac{j}{\omega C}\right)^2\left(R-\frac{j}{\omega C}\right)}{R^2}\right\}I_4-\left(2R-\frac{j}{\omega C}\right)I_4-I_4\left(R-\frac{j}{\omega C}\right)$$

$$\therefore \quad \frac{I_4}{I_1}=\frac{R^3}{\left(2R-\frac{j}{\omega C}\right)^2\left(R-\frac{j}{\omega C}\right)-\left(3R-\frac{2j}{\omega C}\right)R^2}$$

$$=\frac{R^3}{R^3-j\frac{6R^2}{\omega C}-\frac{5R}{\omega^2C^2}+\frac{j}{\omega^3C^3}}. \qquad \ldots(5)$$

For a phase shift of 180° between I_4 and I_1, the terms containing j should vanish and we have

$$\frac{j}{\omega^3C^3}-\frac{j\,6R^2}{\omega C}=0$$

or
$$\frac{1}{\omega^2C^2}=6R^2.$$

$$\therefore \quad \omega=\frac{1}{\sqrt{(6).RC}}.$$

The frequency of oscillations is given by

$$f=\frac{1}{2\pi\sqrt{(6).RC}}. \qquad \ldots(6)$$

At this frequency, from equation (5), we have

$$\frac{I_4}{I_1}=\frac{R^3}{R^3-\frac{5R}{C^2}\times 6R^2C^2}=-\frac{1}{29}. \qquad \ldots(7)$$

From equation (7) it is clear that transistor must give a current gain of atleast 29 to achieve oscillations. Thus a transistor with high value of feed back is selected to give oscillations. For sinusoidal output the transistor must not oscillate too strongly and the gain should be adjusted to give only a small amplitudes of oscillations.

A convenient value for the frequency determining resistor is 4·7 $C\Omega$ because this permits a collector current of 1 or 2 mA.

with a low supply voltage. The value of capacity required (say for oscillator 1 Kc/s) is calculated from equation (6), *i.e.*

$$C=\frac{1}{2\pi\sqrt{(6)}.fR}$$

$$=\frac{1}{2\times 3\cdot 14\times 2\cdot 45\times 10^{3}\times 4\cdot 7\times 10^{3}}\text{ Farad}$$

$=0\cdot 015\ \mu$F approximately.

18·5-2. WIEN BRIDGE OSCILLATOR :

This is also audio frequency oscillator. The advantage of this oscillator is that the frequency may be varied over a frequency range 1 : 10, whereas in *RC* oscillators the frequency cannot be varied, *i.e.* *RC* oscillator is a fixed frequency oscillator. In *RC* oscil. lator both frequency determining network and amplifier introduce a phase change in the signal and positive feed back is obtained: On the other hand, the oscillations may be obtained by arranging both network and amplifier to introduce zero phase shift and actually this is the principle of Wien-bridge oscillator.

The Wien-bridge network is shown in figure 8. The upper arm consists of resistance and capacitance in series, while the lower arm of equal resistance and capacitance in parallel. The network is supplied from a constant voltage source and is terminated in an infinite impedance. The phase shift and attenuation introduced by the network can be calculated as follows :

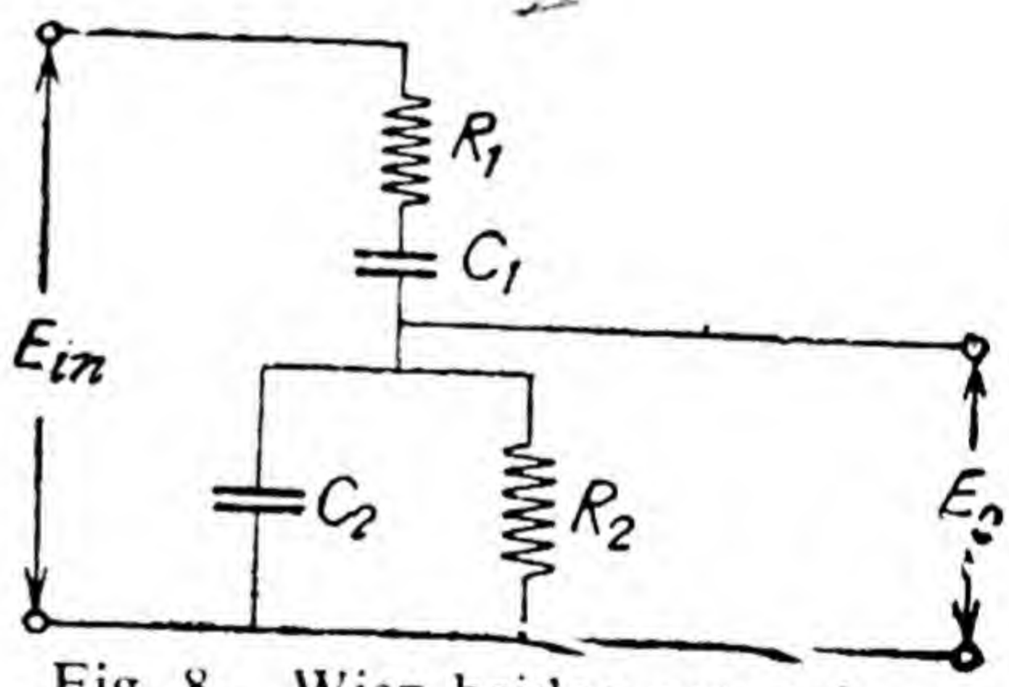


Fig. 8. Wien bridge network.

The impedance of parallel *RC* network is $\frac{-(R.j\,\omega C)}{R-(j/\omega C)}$ and *RC* series network is $\left(R-\frac{j}{\omega C}\right)$

$$\therefore\quad \frac{E_0}{E_{in}}=\frac{\text{impedance of parallel combination}}{\text{Total impedance}} \qquad \ldots(1)$$

$$=\frac{\frac{-(Rj/\omega C)}{R-(j/\omega C)}}{R-\frac{j}{\omega C}-\frac{(Rj/\omega C)}{R-(j/\omega C)}}$$

$$=\frac{-\frac{jR}{\omega C}}{\left(R-\frac{j}{\omega C}\right)^{2}-\frac{jR}{\omega C}}$$

$$= \frac{-\frac{jR}{\omega C}}{\left(K^2 - \frac{1}{\omega^2 C^2} - \frac{3jR}{\omega C}\right)} \quad \ldots(2)$$

There is zero phase shift if the terms in j vanish, *i.e.*,

$$R^2 - \frac{1}{\omega^2 C^2} = 0$$

or $$\omega = \frac{1}{RC}.$$

$$\therefore \quad f = \frac{1}{2\pi RC}. \quad \ldots(3)$$

Again from equation (2), we get

$$\frac{E_0}{E_{in}} = \frac{-jR/\omega C}{-3jR/\omega C} = \tfrac{1}{3}. \quad \ldots(4)$$

The maintaining amplifier thus requires a gain just exceeding 3 to sustain oscillations. Figure 9 gives the circuit diagram of the Wien-bridge oscillator. The amplifier section employs three transistors, two as common emitter stages connected in cathode to give the desired characteristic. The final stage is an emitter follower which gives unity voltage gain for feeding to the frequency determining network and the output terminals of the oscillators.

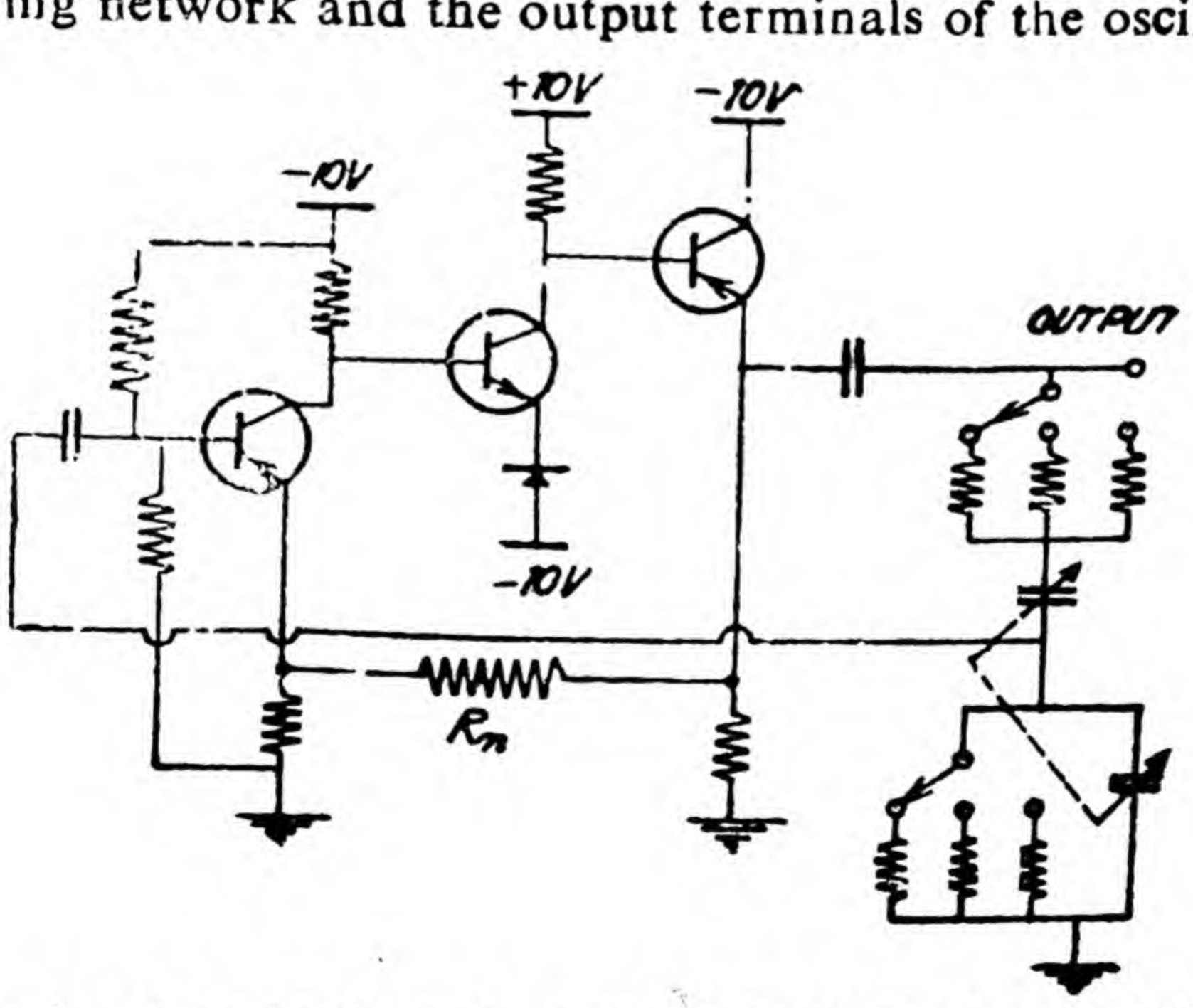


Fig. 9 Simplified circuit diagram for a Wien–bridge oscillator with range switching.

A resistors R_n with negative temperature coefficient is connected as shown in figure which stabilises the amplitude of the oscillalations. Any tendency for the output amplitude to increase, increases the negative feed back, and thus decreases the amplifier gain; consequently checks the rise in amplitude.

EXERCISES AND PROBLEMS

1. **Answer the following :**
   (i) **Is it correct that oscillator is an amplifier with infinite gain ?**
   (ii) **Why do you prefer CE configuration in oscillator circuit ?**
   (iii) **How are oscillations normally produced in most transistor oscillators ?**
2. **Draw the circuit diagrams of Hartley and Colpitt oscillators and obtain the expressions for the frequency of oscillations.**
3. **Explain with the aid of circuit diagram the working of a transistor R-C oscillator.**
4. **Discuss the operation of a phase shift oscillator.**

   A phase shift oscillator using a *PNP* transistor has the following circuit constants ?

   $V_{cc} = -10$ **volts** $\qquad R_1 = R_2 = R_3 = 3{\cdot}2\ K\Omega$

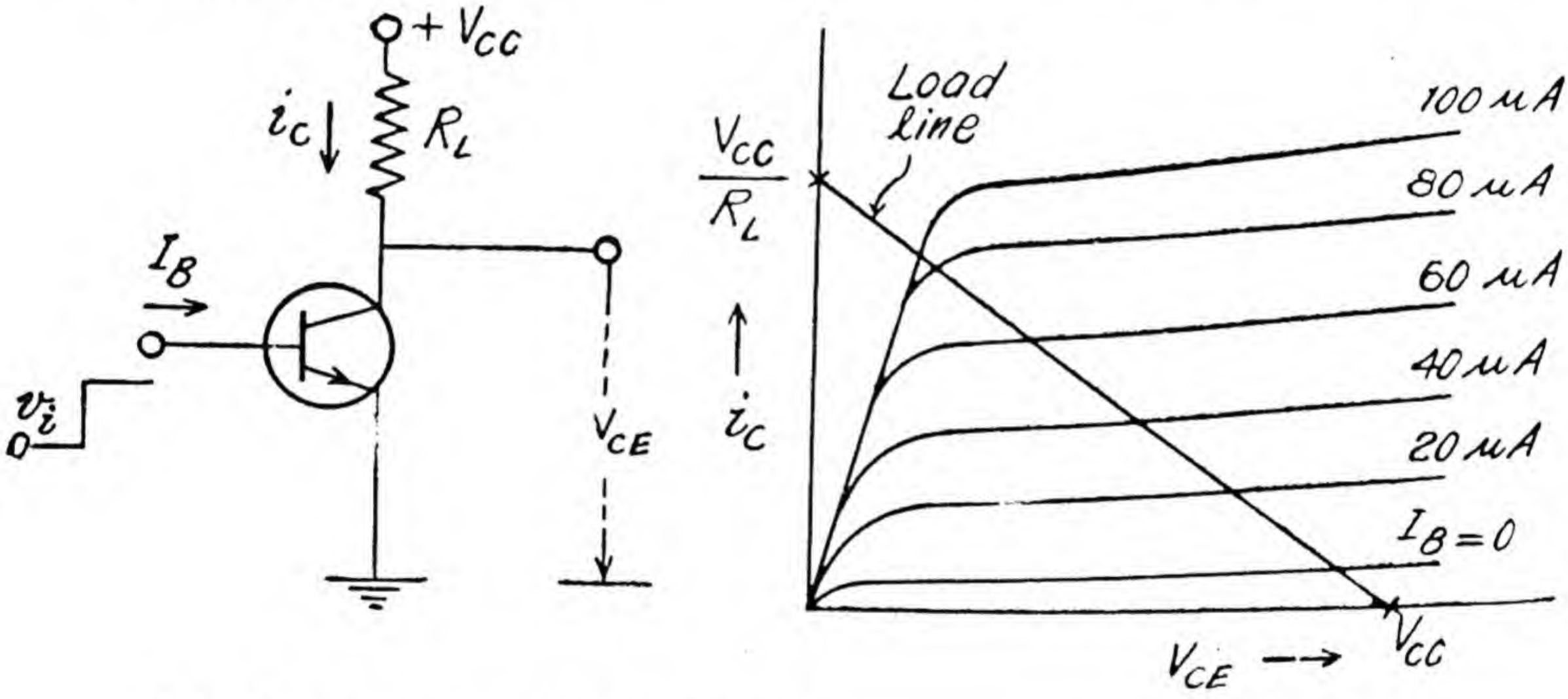
   $R_L = 10\ K\Omega \qquad C_1 = C_2 = C_3 = 0{\cdot}02\ \mu f.$

Calculate the current and frequency of oscillations.

19 MULTIVIBRATORS

19·1. SWITCHING CHARACTERISTICS OF A TRANSISTOR

An ideal switch is one which offers infinite resistance when open and zero resistance when closed. Consider the case of a transistor in common emitter configuration with a step input voltage v_i, applied at the base as shown in fig. (1 *a*). The load line is shown in fig. (1 *b*).



(a) Common emitter transistor with step input voltage

(b) Load line

Fig. 1. Common emitter transistor

In the absence of input voltage *i.e* , $v_i = 0$, the base current I_B is zero. Now no current flows through load resistance R_L except the leakage current I_{CEO} which, of course, is very small. In this state the voltage from collector to emitter is maximum ($\approx V_{CC}$) because there is negligible voltage drop across R_L when $I_C = 0$. The transistor is offering a fairly large resistance to the flow of current ($R_{off} = V_{CE}/I_{CEO}$) and hence the transistor is cut off. The leakage current I_{CEO} can be minimised by reverse biasing the base emitter junction with a proper bias. The leakage current now comes down to cut off current I_{CO}. The transistor offers very high resistance andone can think the transistor as being open.

Let us now consider that a step input voltage is applied between base and emitter. Now base current flows. The switch is in ON position. The base current necessary for it can be obtained from load line. Usually, more current than the minimum base

current to just cause saturation is used to switch the transistor. It should be remembered that if the transistor is in heavily saturated state, it will take a longer time to switch off when v_i is reduced to zero. In the ON position V_{CE} is almost zero and full V_{CC} appears across the load. The resistance of the transistor depends upon the collector current and is very small.

It is obvious from the output characteristics fig. (1 *b*) that as the base current increases, the voltage V_{CE} decreases. The decrease in voltage is appreciable till the base current reaches a certain value (80 μA). When the base current is further increased, the change in V_{CE} is marginal. Physically as the collector current increases, the base collector voltage decreases and reverses sign for a certain value of collector current Under this condition the minority carriers are injected from emitter and collector into the base and the transistor is saturated. The voltage across the collector and emitter is the difference between the base emitter and base collector voltages. This is represented by $V_{CE\ (sat)}$ and is very small. The ratio of $V_{CE\ (sat.)}$ to collector current represents the resistance offered by the transistor in saturation. This resistance is very small. *Thus when the transistor is in* OFF *state, it offers very large resistance conducting only few nano amperes of current and when in* ON *state, it offers a low resistance of few ohms.*

19·2. SWITCHING TIMES IN A TRANSISTOR

It is observed in a transistor that the application of a square wave input to the base-emitter junction is not followed directly by a change in the output voltage because of capacitive effects. There is some delay. The time interval between the instant of switching on and the instant the transistor goes into active state is called the delay time. After applying V_{BE}, the current I_C starts to rise after certain time *i.e.*, there is a delay. This delay of time is represented by t_d and is shown in fig. (2). After the transistor goes into active

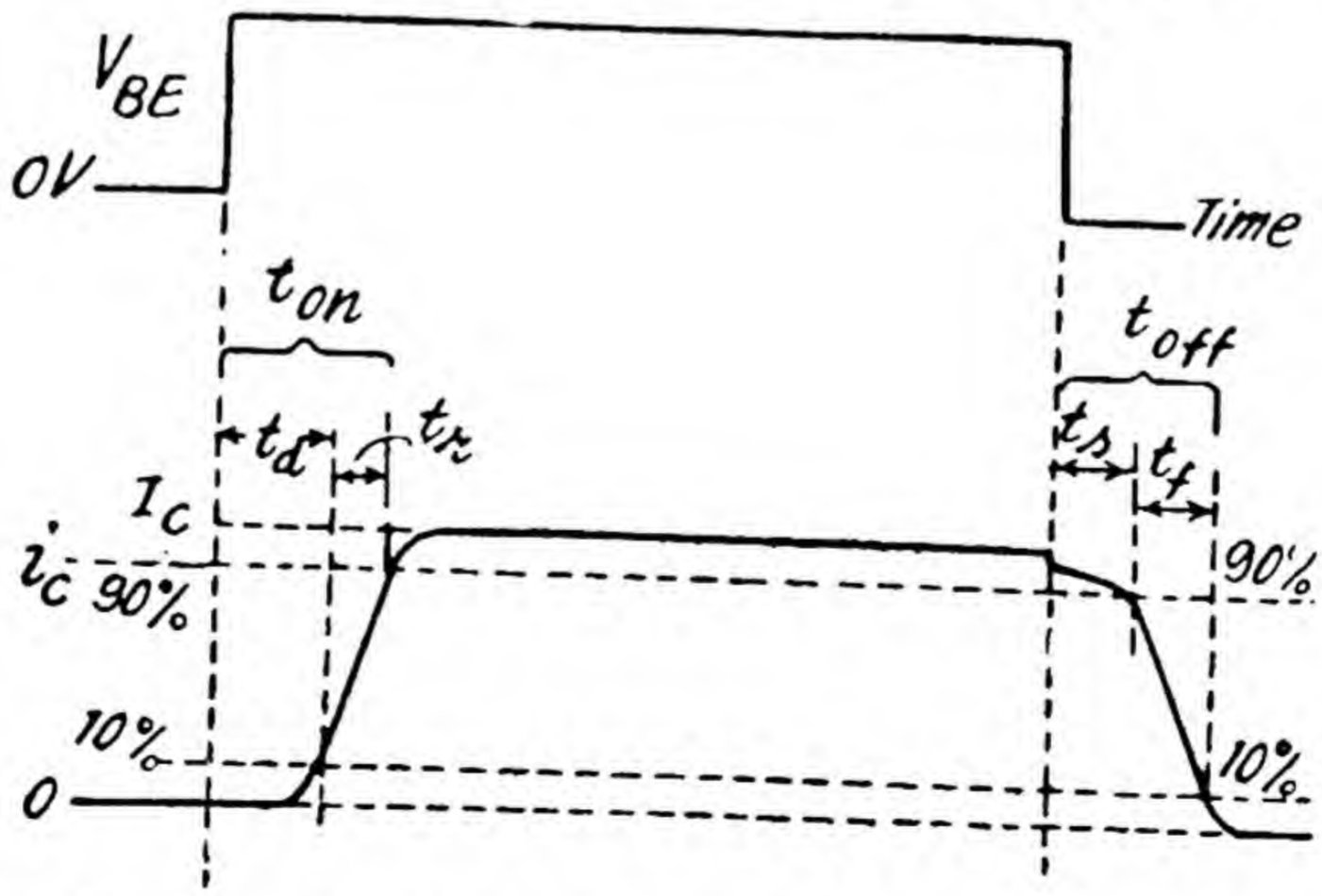


Fig. (2) Showing the input square wave, resulting output waveform and the time delay involved when the transistor is driven into heavy saturation.

state the output current increases exponentially with time due to the capacitance existing at the input terminals of the transistor. The time required for the output current to rise from 10% to 90% of its final value is called the *rise time*. This is denoted by t_r. The sum of the two times is known as *turn on time* t_{on} *i.e.*

$$t_{on}=t_d+t_r \qquad ...(1)$$

When V_{BE} goes to zero, there is a delay before i_o starts to change due to the charge stored in the transistor. The delay is referred to as charge storage time and is denoted by t_s. The fall time to reach 10% of the initial value is denoted by t_f. The sum of these two delays is termed as turn off time t_{off}, *i.e.*,

$$t_{off}=t_s+t_f \qquad ...(2)$$

It should be remembered that rise and fall times depend upon the transistor's high frequency characteristics.

19·3. MULTIVIBRATORS

A multivibrator is a relaxation oscillator consisting of two stage resistance coupled amplifier with the output of each stage coupled regeneratively to each other. The relative condition of each of the two stages *i.e.*, conduction in the first stage and cut off in the second stage or vice versa is termed as *state* of the multivibrator. The condition in which the multivibrator may remain indefinitely until the circuit is triggered by some external signal is known as *stable state*. Multivibrators performs a variety of useful functions like generation of square waveforms, counting, frequency division, generation of time delays etc. The multivibrators can be classified in the following three categories :

(1) *Astable or free running multivibrator*

(2) *Monostable multivibrator*

(3) *Bistable multivibrator.*

(1) Astable multivibrator—Astable multivibrator is a multivibrator in which neither stage is a stable state. There are two quasi-stable (temporary) states. The circuit changes state continuously from one quasi stable state to another state at regular time intervals without any triggering. This generates continuous square waveform without any external signal.

(2) Monostable multivibrator—Monostable multivibrator is also known as one shot multivibrator or univibrator. This has only one stable state. When a pulse is applied to the input circuit, the circuit state is changed abruptly to unstable state for a predetermined time after which the circuit returns to its original stable state automatically *i.e.*, by itself. The circuit enables production of pulses of variable width at the required moment.

(3) Bistable multivibrator—Bistable multivibrator is also known as flip-flop, binary and scalar of two circuits. It has two stable states. Bistable circuit may stay in a particular state indefinitely and hence can be used to store binary bit of information.

When an external trigger is applied, the state of binary is changed At the occurrence of each triggering pulse, the circuit state changes abruptly from one state to another This is used as a digital memory device.

19·4. ASTABLE MULTIVIBRATOR

A free-running or astable multivibrator has no stable states, but the circuit has two quasi-stable states 'transistor first ON and transistor second OFF' and 'transistor first OFF and transistor second ON', and it makes periodic transitions between these states without any external triggering. Fig. (3) shows the circuit diagram of an astable multivibrator and fig. (4) gives the waveforms. In fig. (3), R_{L1} and R_{L2} are the collector resistances for transistors T_1 and T_2 respectively. C_1 and C_2 are coupling capacitors, R_{B_1} and R_{B_2} provide ON state base current to the two transistors respectively during saturation region. In symmetrical multivibrator

$$R_{L_1}=R_{L_2},\ R_{B_1}=R_{B_2} \text{ and } C_1=C_2$$

i.e., the components in one half of the circuit are considered to be equal to their components in the other half.

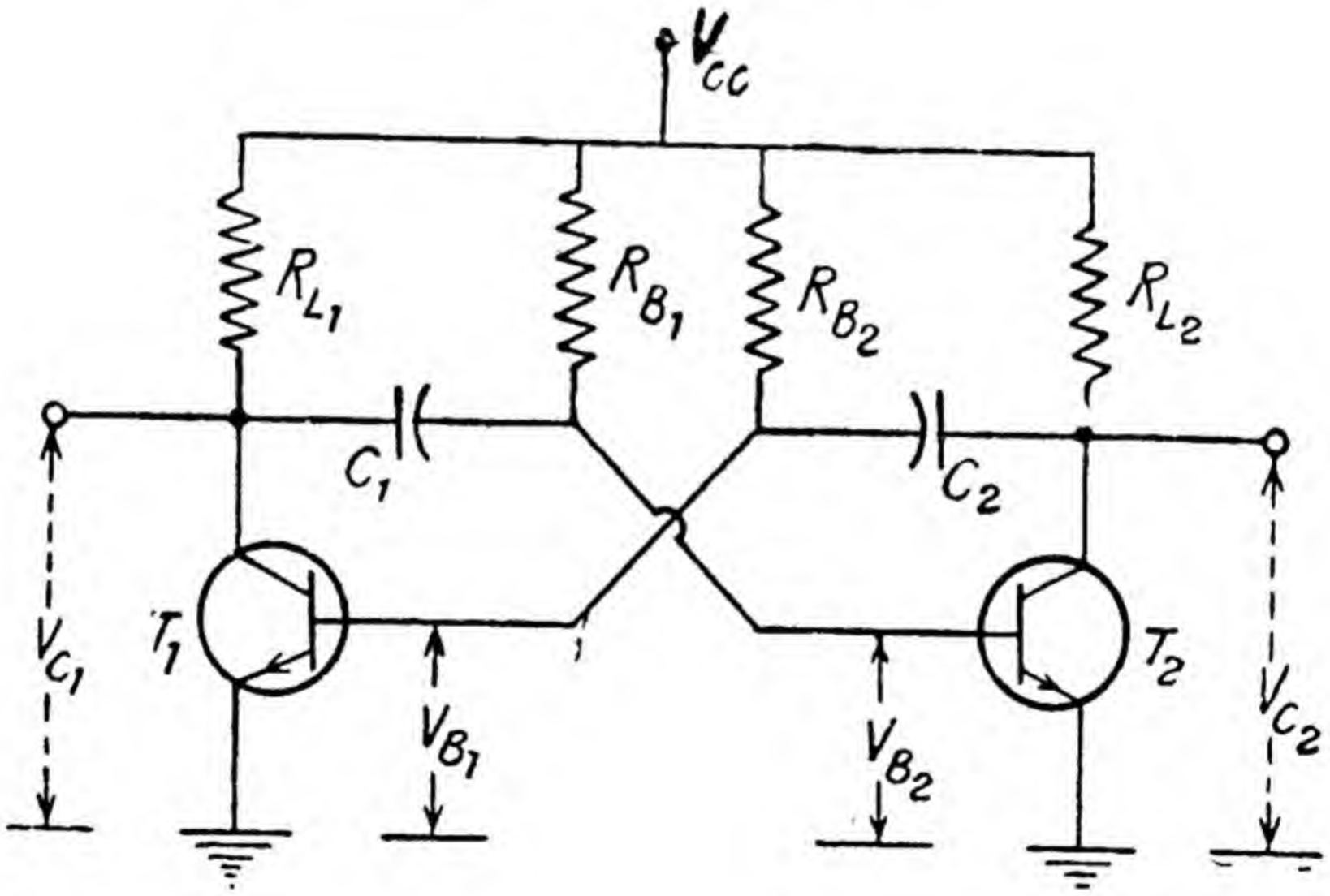


Fig. (3) Astable multivibrator.

Operation : The operation of astable multivibrator is as follows :

First of all we shall explain the starting situation *i.e.*, how the transistor T_1 goes into saturation and T_2 goes to cut off region. When the power is applied, let due to some circuit imbalance, the transistor T_1 conducts more than transistor T_2 *i.e.* current flowing through transistor T_1 is more than the current flowing in transistor T_2. Due to this fact the rate of fall of V_{C_1} is more than that of V_{C_2}. The changes of collector voltage of T_1 will make the

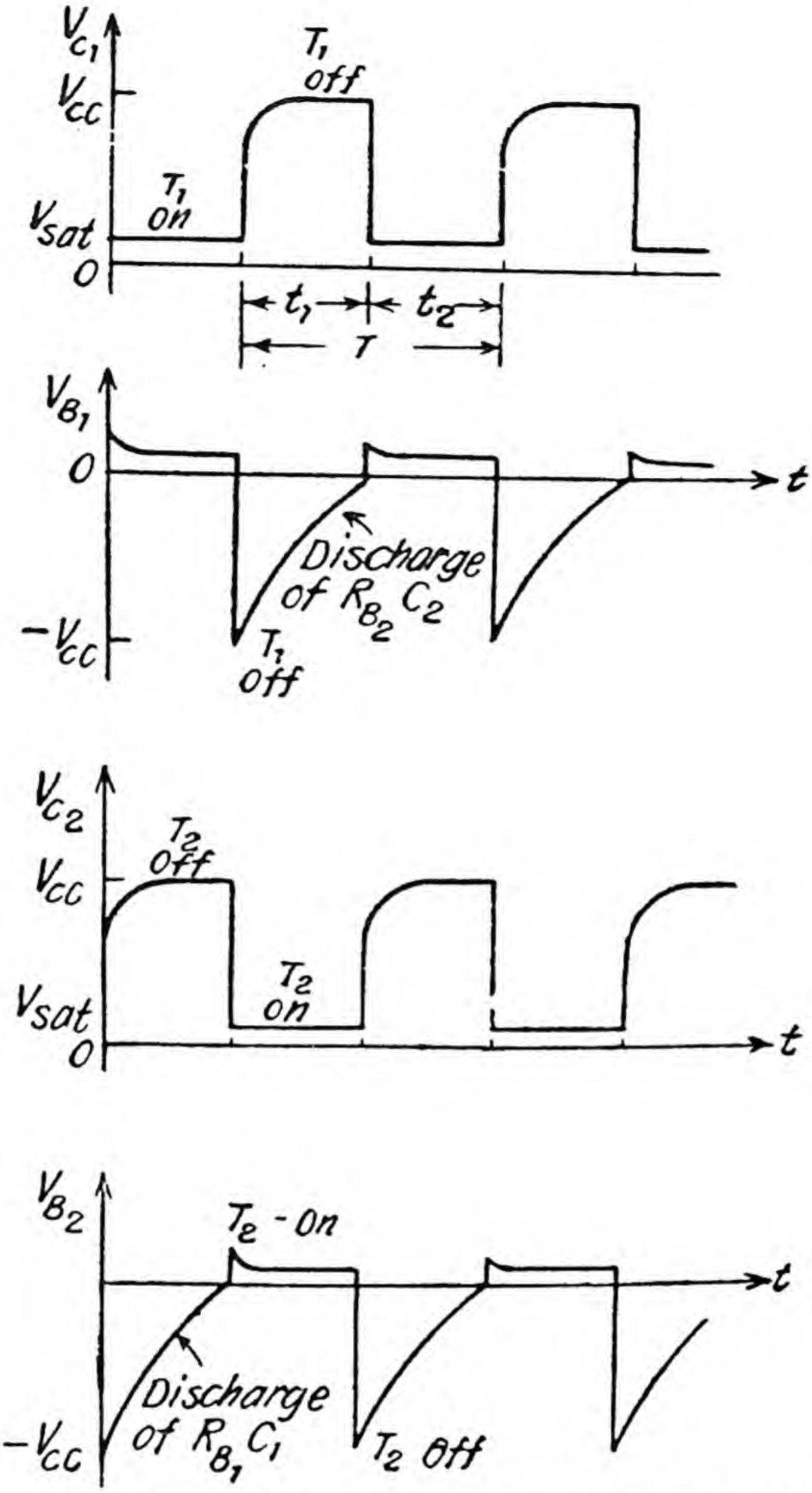


Fig. (4) Waveforms at the collector and base of transistors T_1 and T_2 in astable multivibrator.

base of transistor T_2 negative. This decreases the conduction of T_2 *i.e.*, collector voltage V_{C_2} increases and approaches towards V_{CC}. The increase in V_{C_2} will be transferred to the base of transistor T_1 through capacitor C_2. So the conduction of T_1 further increases. This further reduces the collector voltage V_{C_1} which in turn makes the base of transistor T_2 more negative thereby further reducing its conduction. Thus due to the positive feed back action, the transistor T_1 comes in the saturated region and transistor T_2 comes in cut off region. The circuit settles down momentarily with transistor T_1 in saturation and transistor T_2 in cut off. This is only a quasi-stable stable. In this state V_{C_1} is approximately zero while V_{C_2} is approximately equal to V_{CC}. This permits the charging of

C_2 through R_{L_2} and the base emitter junction of transistor T_1 comes to a potential approximately equal to V_{CC}. The charging path is shown in fig. (5a). Meanwhile the condenser C_1 which had previously charged to V_{CC} (when T_2 was conducting) discharges through T_1 and R_{B_1} as shown in fig. (5b). The initial pulse of discharge current from C_1 through R_{B_1} makes the base of T_2 sud-

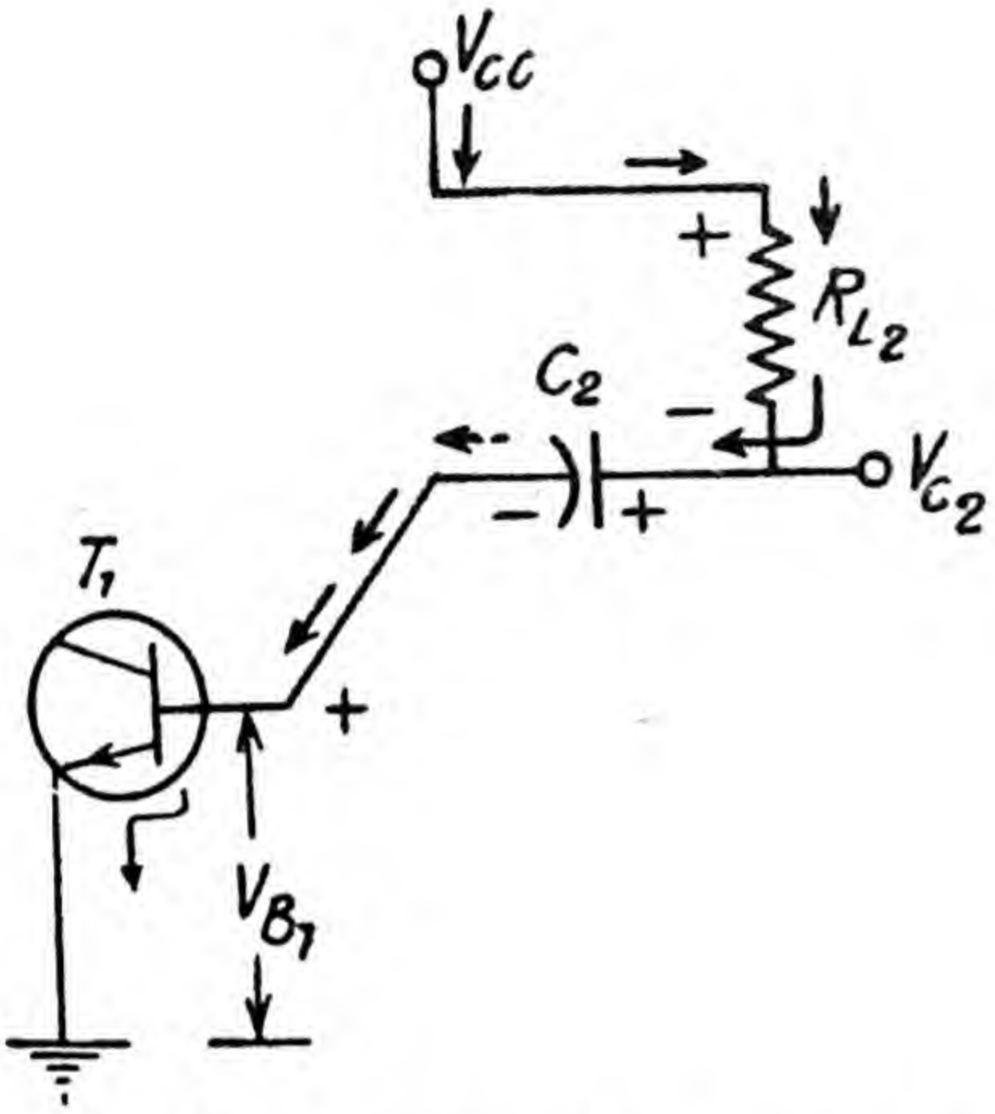


Fig. (5a) Charging of C_2 through T_1 and R_{L_2}.

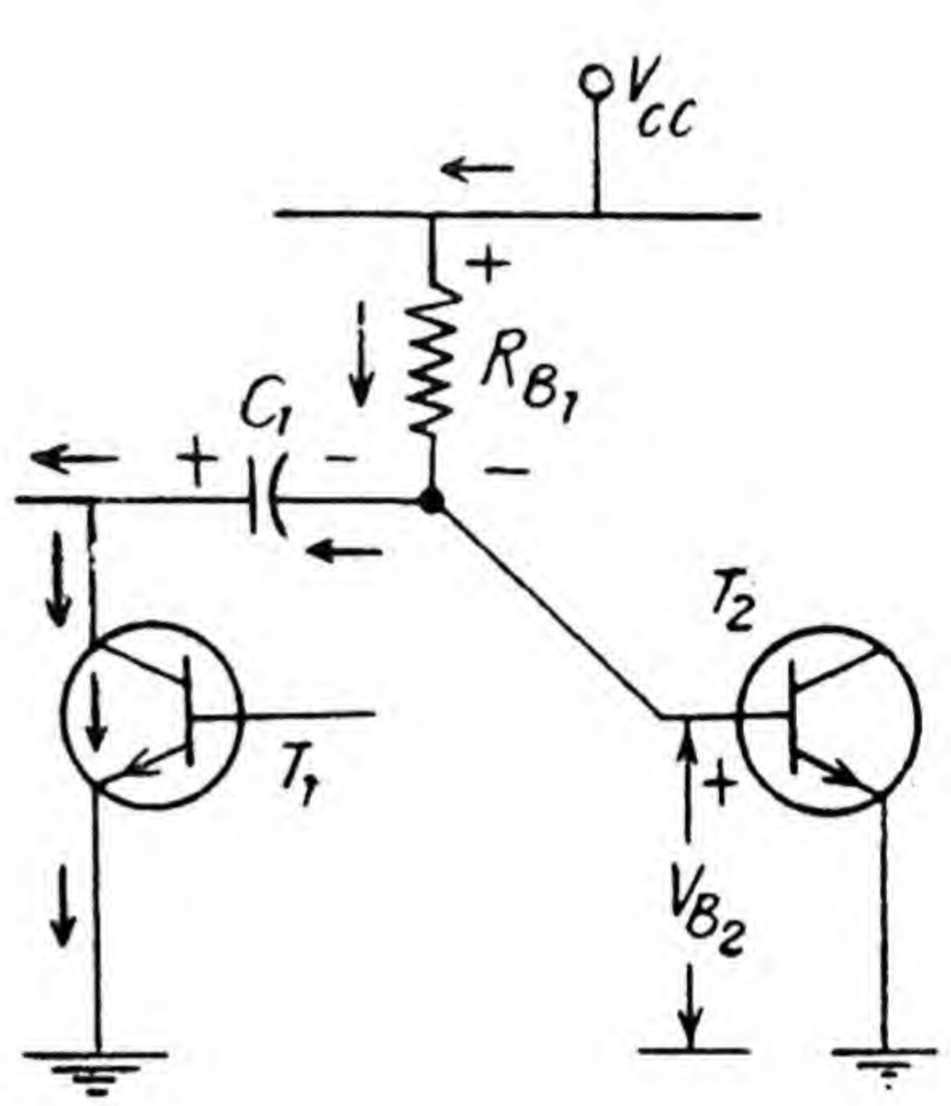


Fig. (5b) Discharging path of C_1 through T_1 and R_{B_1}.

denly very negative approximately equal to $-V_{CC}$ as shown in waveform of V_{B_2}. This causes transistor T_2 to be cut off as assumed. The transistor T_1 is kept in the conducting state by base current provided from V_{cc} through R_{B_2}. so V_{B_1} is slightly positive. The charging current of C_2 through R_{L_2} has now ceased. The length of the time for which T_2 is held off is determined by the time constant of discharge of C_1 through T_1 and R_{B_1}.

Now T_2 is no longer cut off but starts conducting. The collector voltage of T_2 begins to fall. The falling collector voltage is communicated to the base of transistor T_1 *via* capacitor C_2. This causes the reduction of base voltage of transistor T_1 and consequently there is an increase in collector voltage which is communicated to base of T_2 *via* capacitor C_1. The conduction of T_2 increases. The regenerative process continues till the transistor T_2 is in saturated region and T_1 is in cut off region. Under these conditions $V_{C_1} = V_{CC}$ and $V_{C_2} = 0$. The condenser C_1 charges through R_{L_1} and T_2 causing an increase in the positive voltage at base of T_2 as shown in fig. (6a). Meanwhile C_2 discharges through T_2 and R_{B_2} as shown fig. (6b). The interval in which transistor T_1 is held in non-conducting state depends upon R_{B_2} C_2 time constant.

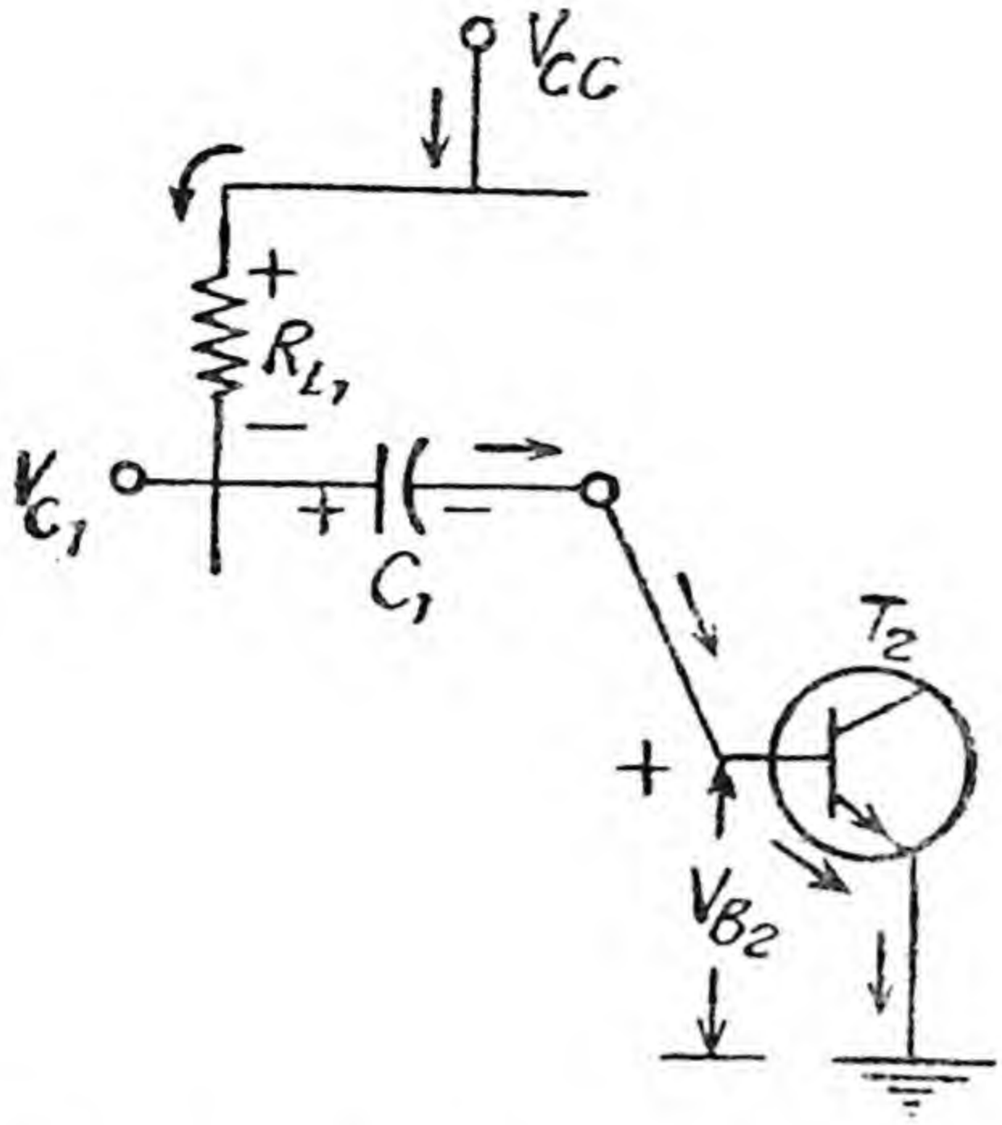


Fig. (6a). Charging path of C_1 through T_2 and R_{L_1}.

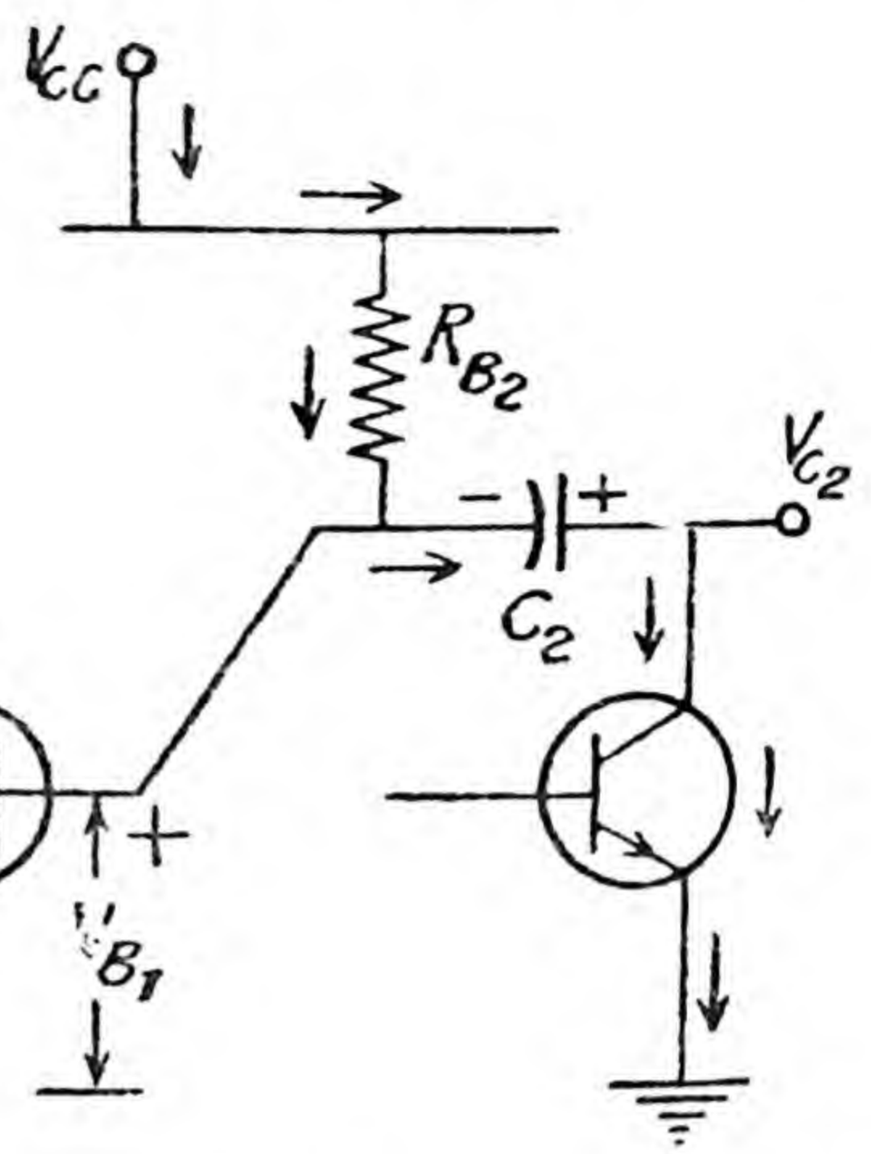


Fig. (6b) The discharging path of C_2 through T_2 and R_{B_2}.

Here it should be remembered that turning on of one transistor, say T_2 and its falling voltage at the collector permits the discharge of a capacitor which drives the transistor T_1 into cut off. The rising voltage at the collector of T_1 'feeds back' to the base of T_2 tending to turn it *on* more. The process is regenerative or cumulative.

Frequency of oscillations : The frequency of oscillations can be calculated in the following way :

During the discharge of condenser C_2, the value of V_{B_2} is given by

$$V_{B_2}=V_{CC}-i_{C_1} R_{B_1} \qquad ...(1)$$

The initial discharge value of current i_{C_1} s given by

$$i_{C_1}=\frac{V_{CC}+V_{CC}}{R_{B_1}},$$

because the capacitor charges upto V_{CC} approximately. The current decays exponentially with a time constant $R_{B_1}C_1$

$$\therefore \qquad V_{B_2}=V_{CC}-2V_{CC} \text{ exp. } (-t/R_{B_1}.C_1) \qquad ..(2)$$

Suppose $V_{B_2}=0$, when the transistor is switched, then

$$0=V_{CC}\,[1-2 \text{ exp. } (-t/R_{B_1}.C_1)]$$

or $$\text{exp. } (t/R_{B_1}.C_1)=2 \qquad (\because \quad V_{CC}\neq 0)$$

or $$t=R_{B_1}.C_1 \log 2=0{\cdot}694\ R_{B_1}\ C_1 \qquad ...(3)$$

Thus the *off* time (t_2) for transistor T_2 or the *on* time of transistor T_1 is given by

$$t_2=0\ 694\ R_{B_1}.C_1 \qquad ...(4)$$

Similarly, the *off* time (t_1) for transistar T_1 or the *on* time of transistor T_2 is given by

$$t_1 = 0{\cdot}694\ R_{B_2}.C_2 \qquad \ldots(5)$$

The total time period $T = 0{\cdot}694\ (R_{B_1}\ C_1 + R_{B_2}\ C_2)$

when $R_{B_1} = R_{B_2} = R$ and $C_1 = C_2 = C$, then

Total time period $T = 1{\cdot}39\ RC$ ...(6)

The frequency of free running multivibrator is given by

$$f = \frac{1}{\text{Total time period } (T)}$$

$$= \frac{1}{1{\cdot}39\ RC}.$$

The frequency stability of the circuit is not good as it is not only the function of the product of RC but also depends on load resistances, supply voltages and circuit parameters. In order to stabilise the frequency, synchronising signals are injected which terminate the unstable periods earlier than would occur naturally.

19·5. MONOSTABLE MULTIVIBRATOR :

A monostable or a one-shot multivibrator or a univibrator has one stable state *i.e.*, permanent and the other as temporary *i e.*, quasi-stable. When an external trigger is applied to the input, the multivibrator changes the state from stable state to quasi-stable state. It stays in quasi-stable state for a pre-determined length of time and after which the circuit returns to its original stable state automatically. The circuit produces the pulses of variable width at the required moment.

A generalized circuit of a monostable multivibrator using two active devices A_1 and A_2 (tubes or transistors) is shown in fig. (7).

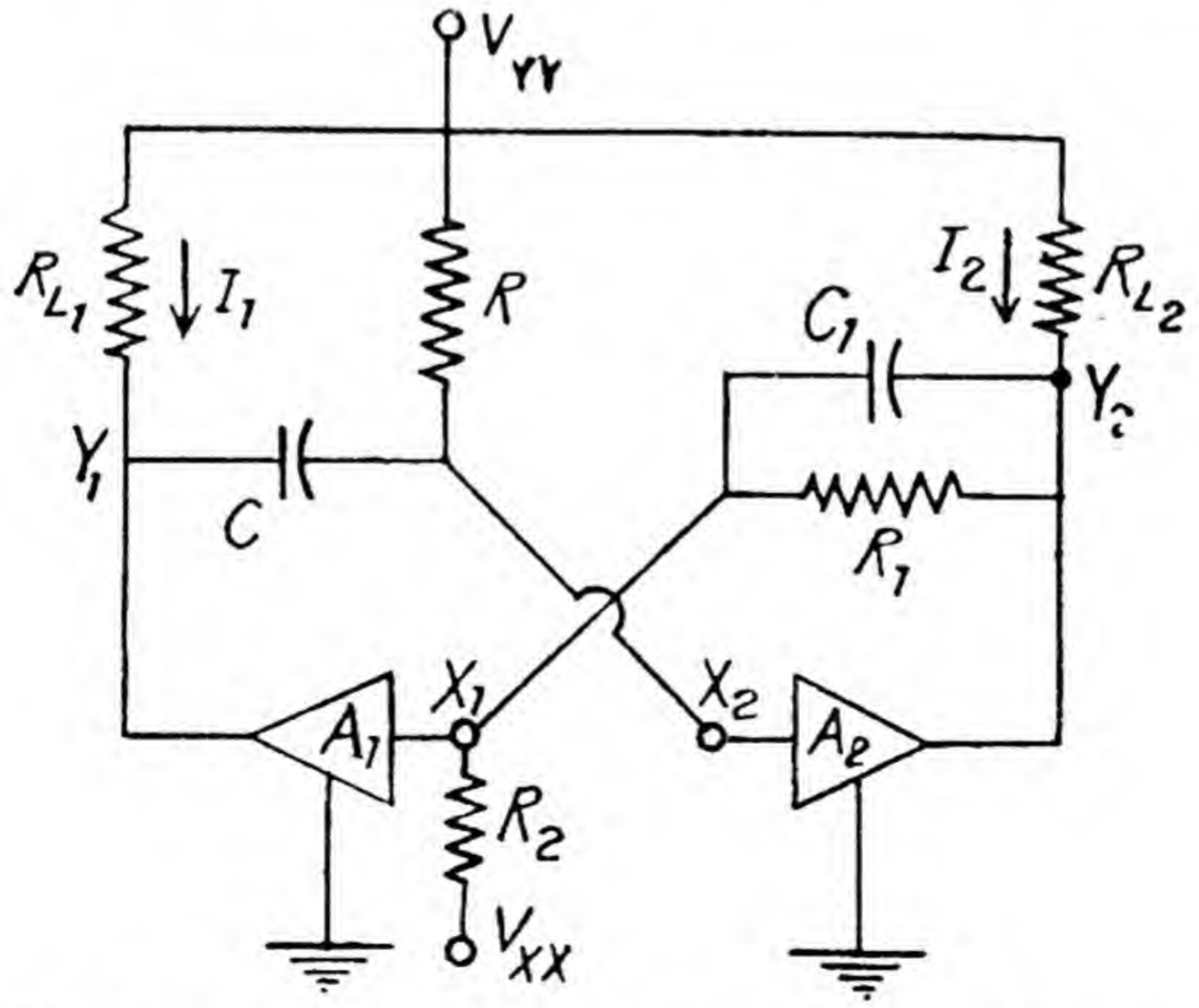


Fig. 7. A generalized circuit configuration of a monostable multivibrator.

Suppose that a single trigger is applied to X_2 and a regenerative action takes place which derives A_2 to comp letely cut off. The voltage at Y_2 rises approximately to V_{YY}. Now A_1 comes in conduction. The device may be driven into saturation or it may operate within its active region. In either case a current I_1 exists in R_{L_1} which causes a voltage drop $I_1 R_{L_1}$ at Y_1 and thereby decreasing the voltage X_2 by the same amount because the voltage across condenser C cannot change instantaneously. The multivibrator is now in its quasi-stable state. The circuit will remain in this state only for a finite time because X_2 is connected to V_{YY} through R. Now X_2 will rise in voltage. When the voltage at X_2 is equal to the cut in voltage of A_2, a regenerative action takes place. This turns A_1 off and eventually returning the multivibrator to its initial stable state. A collector-coupled monostable circuit using *NPN* transistors is shown in fig. 8. The triggering signal is applied at $t=0$ and the reverse transition occurs at $t=T$.

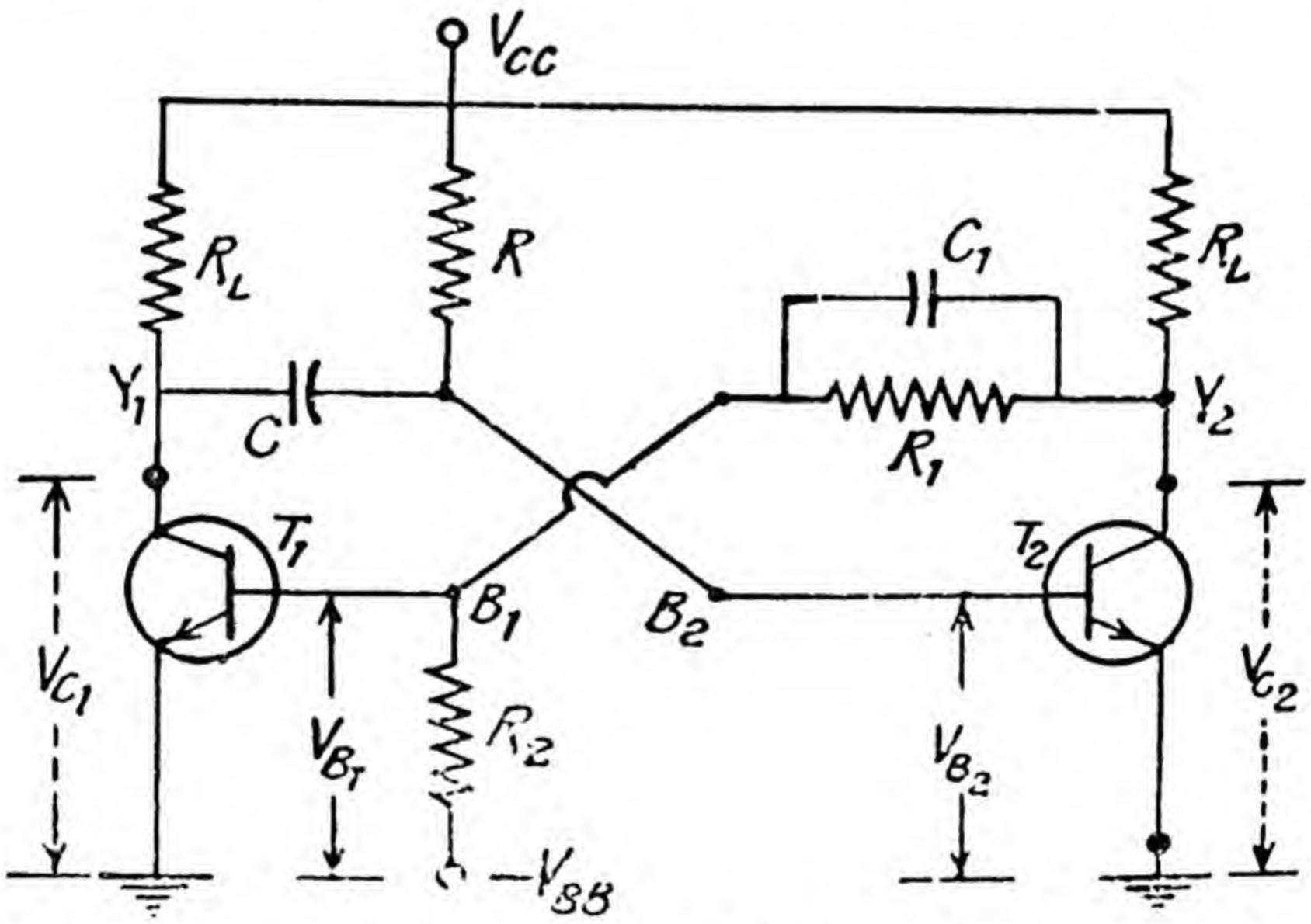


Fig. 8. The collector coupled *NPN* transistor monostable multivibrator.

The stable state : For $t < 0$, the circuit is stable with transistor T_2 in saturation and transistor T_1 OFF. As transistor T_2 is in saturation, the different voltages and currents are expressed as

$$V_{C_2}=V_{CE(sat)}, \qquad V_{B_2}=V_{BE(sat)}, \qquad V_{C_1}=V_{CC}$$

$$I_{C_2}=\frac{V_{CC}-V_{CE(sat)}}{R_L} \text{ and } I_{B_2}=\frac{V_{CC}-V_{BE(sat)}}{R}$$

The base voltage V_{B_1} of transistor T_1 is given by

$$V_{B_1}=\frac{V_{CE(sat)}\cdot R_2}{R_1+R_2}-\frac{V_{BB}\ R_1}{R_1+R_2} \qquad ..(1)$$

The value of V_{BB} is selected in such a way that V_{B_1} has a negative value so that transistor T_1 is OFF.

The quasi-stable state : Let a negative trigger (short duration pulse) is applied at Y_1 at $t=0$. The change in potential at Y_1 is immediately transmitted to B_2 through capacitor C. Now the transistor T_2 is switched off. As the transistor T_2 switches off, the potential at Y_2 rises to almost the value V_{CC} from its stable value $V_{CE(sat)}$. This increases the potential at B_1 and the transistor T_1 comes into saturation. The potential at Y_1 falls by $V_{CC}-V_{CE(sat)}$. This change is communicated to B_2. All above changes take place instantaneously due to regenerative feed back.

Immediately after the trigger *i.e*, at constant $t=0^+$, the initial value of V_{B_2} is given by

$$(V_{B_2})_i = V_{BE(sat)} - (V_{CC} - V_{CE(sat)})$$

or
$$(V_{B_2})_i = V_{BE(sat)} + V_{CE(sat)} - V_{CC} \quad \ldots(2)$$

At $t=0^+$, transistor T_2 is cut off and the potential at B_2 rises exponentially with time due to charging of capacitor C. The potential (V_{B_2}) starts rising towards $+V_{CC}$ [say $(V_{B_2})_f$] with a time constant RC. The instantaneous base voltage $(V_{B_2})_t$ can be expressed as

$$(V_{B_2})_t = (V_{B_2})_f + [(V_{B_2})_i - (V_{B_2})_f] \text{ exp. } (-t/RC)$$

or
$$(V_{B_2})_t = V_{CC} + [V_{BE(sat)} + V_{CE(sat)} - 2\,V_{CC}] \text{ exp. } (-t/RC) \quad \ldots(3)$$

The potential at the base of transistor T_2 does not rise to the maximum value because as soon as the potential rises to cut in voltage, T_2 is switched on due to regenerative feed back. The potential at the point Y_2 falls to its stable value $V_{CE(sat)}$.

Gate width : The gate width of the output pulse is defined as the time interval between T_2 switching off and switching on. In order to calculate the gate width, we make use of the fact that when $(V_{B_2})_t = Y_y$ (cut in voltage) at $t=T$, the transistor T_2 is switched on. From eq. (3), we have

$$V_y = V_{CC} + [V_{BE(sat)} + V_{CE(sat)} - 2V_{CC}] \text{ exp. } (-T/RC)$$

Solving we get

$$T = RC \log \frac{2V_{CC} - V_{BE(sat)} - V_{CE(sat)}}{V_{CC} - V_y}$$

or
$$T = RC \log 2 + RC \log \frac{V_{CC} - \{(V_{BE(sat)} + V_{CE(sat)})/2\}}{V_{CC} - V_y}$$

At room temperature, $V_{BE(sat)} + V_{CE(sat)} \approx 2Y_y$

$$\therefore \quad T = RC \log 2 + RC \log 1 = RC \log 2$$

or
$$T \approx 0{\cdot}69\ RC \quad \ldots(4)$$

As soon as T_2 begins to conduct, a regenerative action makes T_1 OFF and T_2 ON and thereafter, the circuit remains in stable state till another external trigger is applied.

Waveforms for t > T. At $t=T^+$, transistor T_2 conducts and transistor T_1 is cut off. In order to find the potential at B_2, we draw the base circuit of transistor T_2 as shown in fig. (9). At the instant when transistor T_1 is cut off, the instantaneous voltage appearing in the circuit is $V_{CC}-V_\gamma$ as V_γ is the potential at B_2. The instantaneous voltage drives an instantaneous current I'_B, given by

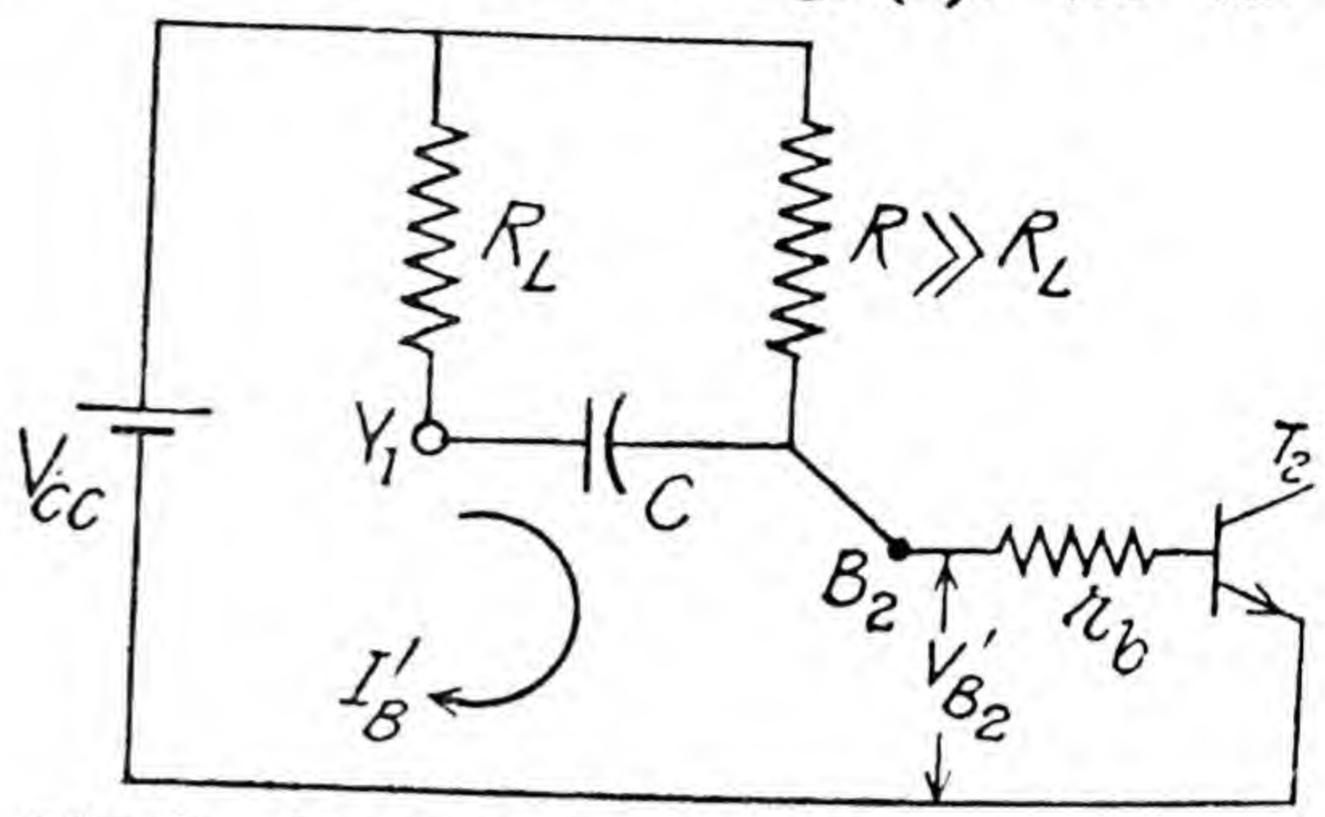


Fig. 9. Equivalent circuit for calculating the over shoot at the base B_2 of T_2 at $t=T^+$. T_1 is off and T_2 is driven heavily into saturation.

$$I'_B = \frac{V_{CC}-V_\gamma}{R_L+r_b} \quad \ldots(5)$$

The current in R may usually be neglected compared with I'_B. The path of I'_B is shown in fig. (9). Here r_b is the base spreading resistance which is considered in series with the junction. The potential at Y_1 now becomes

$$V_{C_1}=V_{CC}-R_L I'_B \quad \ldots(1)$$

The change in potential δ at Y_1 is given by

$$\delta=V_{CC}-R_L I'_B-V_{CE(sat)} \quad \ldots(7)$$

As the capacitor is considered as shorted, the change in potential at B_2 is also the same *i.e.*,

$$V_{CC}-R_L I'_B-V_{CE(sat)}$$

The potential at B_2 immediately after the transient is

$$V'_{B_2}=V_\gamma+V_{CC}-R_L\ I'_B-V_{CE(sat)}$$

or

$$V'_{B_2}=V_{CC}-V_\gamma-R_L\ I'_B+V_{BE(sat)} \quad \ldots(8)$$

$$\because \quad V_{CE(sat)}+V_{BE(sat)}=2V_\gamma$$

From eq. (5), $V_{CC}-V_\gamma=R_L\ I'_B+r_b\ I'_B$

Substituting this value in eq. (8), we get

$$V'_{B_2}=r_b\ I'_B+V_{BE(sat)} \quad \ldots(9)$$

This shows that the transistor is driven deep into saturation. The overshoot at the base of transistor T_2 has a magnitude

$$V_{CC}-R_L\ I'_B-V_{CE(sat)} \quad \ldots(10)$$

The voltage at the point B_2 falls off exponentially to the steady value $V_{BE(sat)}$ with a time constant $C\ (R_L+r_b)$. The poten-

tial at Y_1 suddenly rises to a value $(V_{CC}-R_L)\ I_B'$ and thereafter increases exponentially to a maximum value V_{CC} with a time constant $C(R_L+r_b)$.

The waveforms are shown in fig. (10).

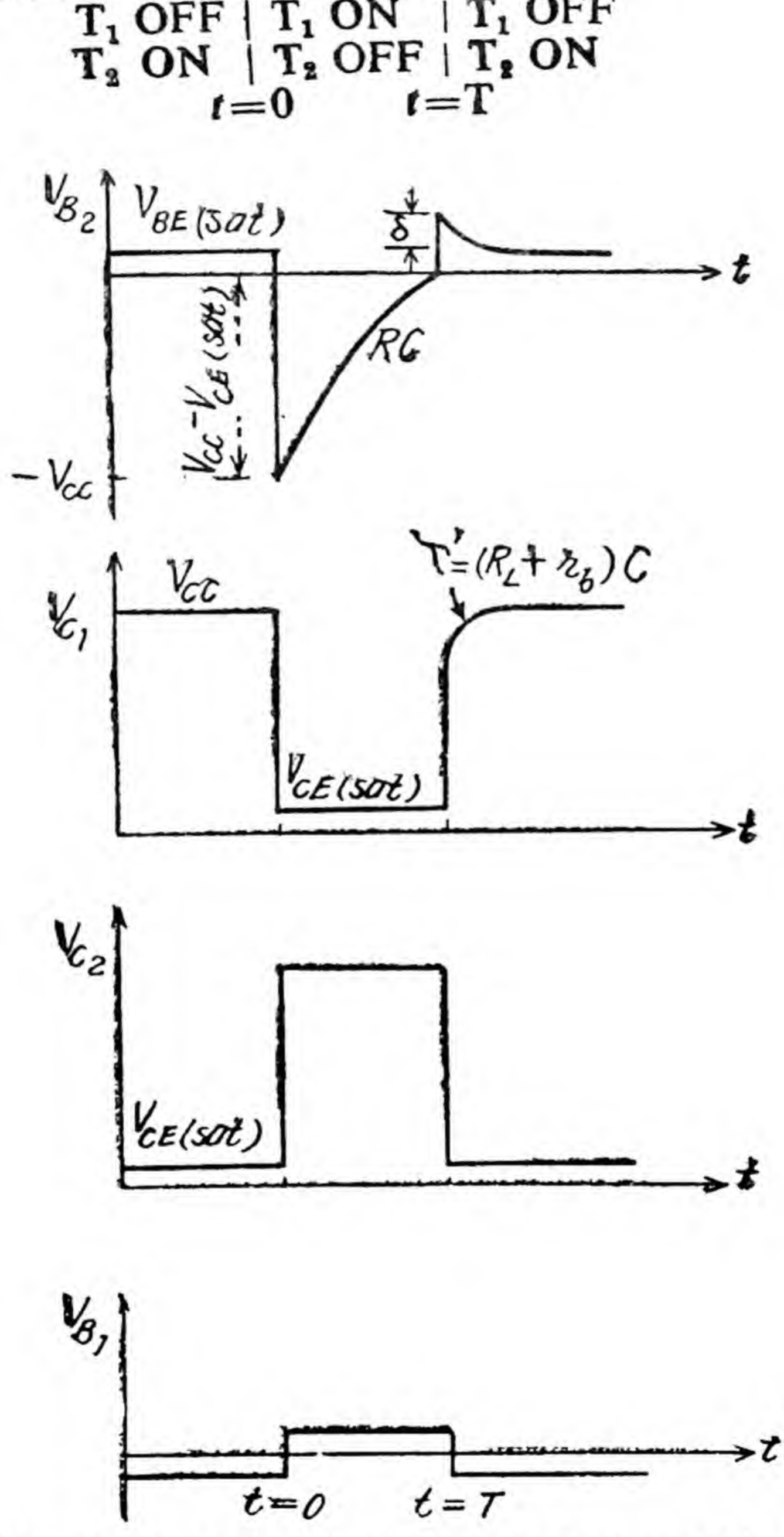


Fig. 10. Waveforms of the collector coupled monostable multivibrator.

19·6. BISTABLE MULTIVIBRATOR :

A bistable multivibrator is also known as *flip-flop* and has two stable states. The circuit can stay in any one of the states indefinitely and makes an abrupt transition from one state to other state when an appropriate trigger or excitation is applied. A bistable multivibrator is used for the performance of many digital operations such as counting and storing the binary information. This can also be used in pulse generation. The bistable circuits use two active devices which function as amplifiers. The amplifiers are coupled in such a way that when one device is heavily con-

ducting or in saturation, the other device is in OFF state. When the device in OFF state is switched ON by external means, the circuit goes in another stable state *i.e.*, the heavily conducting device is now switched OFF.

Generalized circuit : The generalized circuit of a bistable multivibrator is shown in fig. (11). A_1 and A_2 are the active devices which may be tubes or transistors. As shown in the figure, the

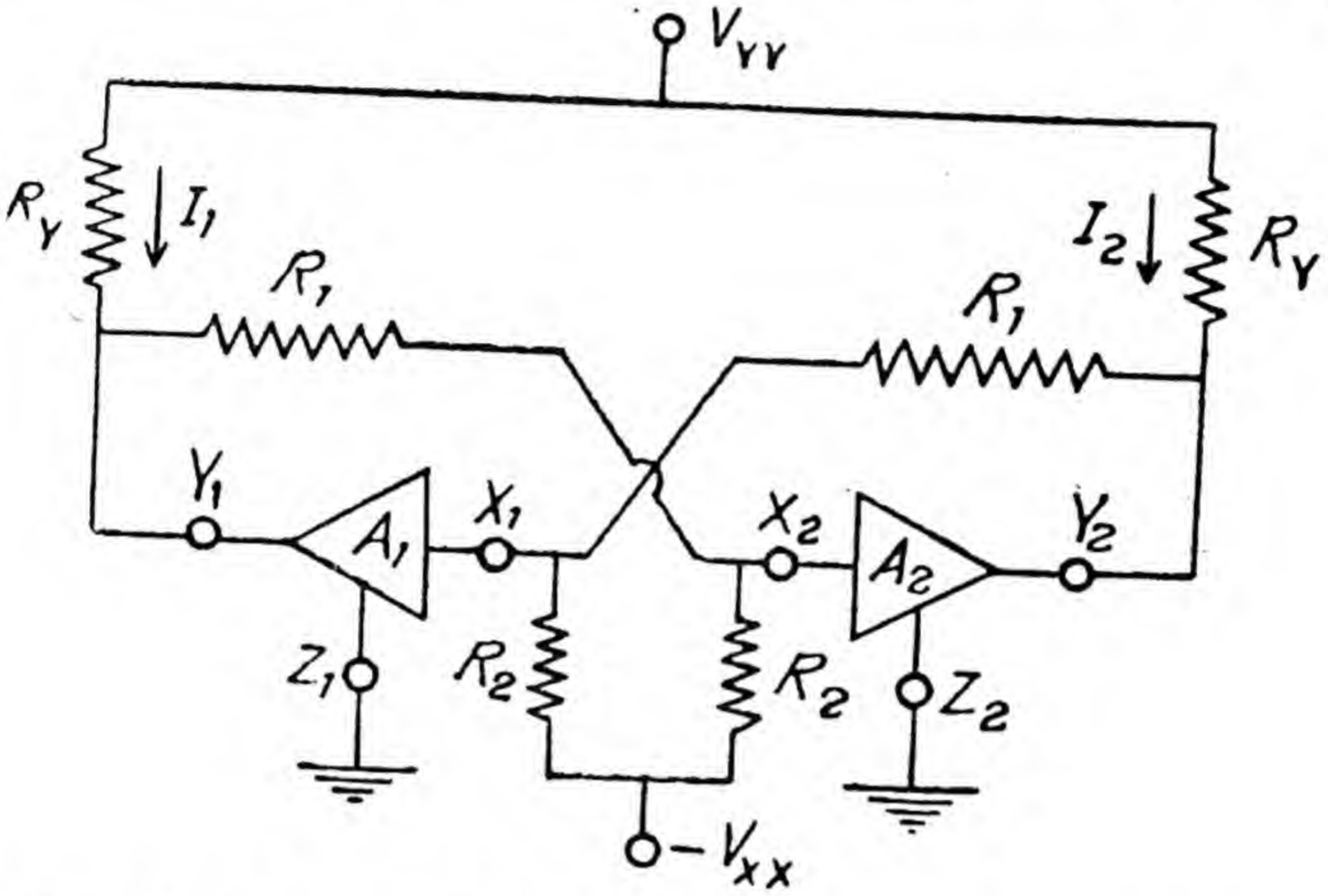


Fig. 11 A generalized circuit of bistable multivibrator.

output of one amplifier is fed back to the second amplifier. Each amplifier produces a phase change of 180° and hence the amplifiers are provided with positive feed back at all frequencies. Suppose I_1 and I_2 be the currents through the devices at any instant as shown in fig. (11). Let there be a momentary decrease in current I_1, then the potential at Y_1 will increase. The will increase the potential at X_2 (input of A_2). The increase in the input voltage of A_2 causes a lowering of potential at Y_2. Hence the potential of X_1 is decreased. As a consequence, the current I_1 is also decreased. This regenerative process continues till A_2 is driven into saturation or heavily conducting state and A_1 is switched OFF.

Transistor (Flip-Flop) : A fixed-bias *NPN* transistor bistable multivibrator is shown in fig. (12*a*). The need for the negative power supply may be eliminated by using a common emitter resistor R_e to provide self-bias. Fig. (12*b*) shows the circuit of a self-biased *NPN* transistor bistable multivibrator.

In fig. (12*a*) two *PNP* transistor amplifiers are crossed coupled so that the voltage appearing across the load resistor of one transistor is fed to the base of the other transistor. The multivibrator is called collector coupled binary as the voltage appearing across the collector resistance of one transistor serves as the input to the other transistor. As explained earlier, the circuit stablises with

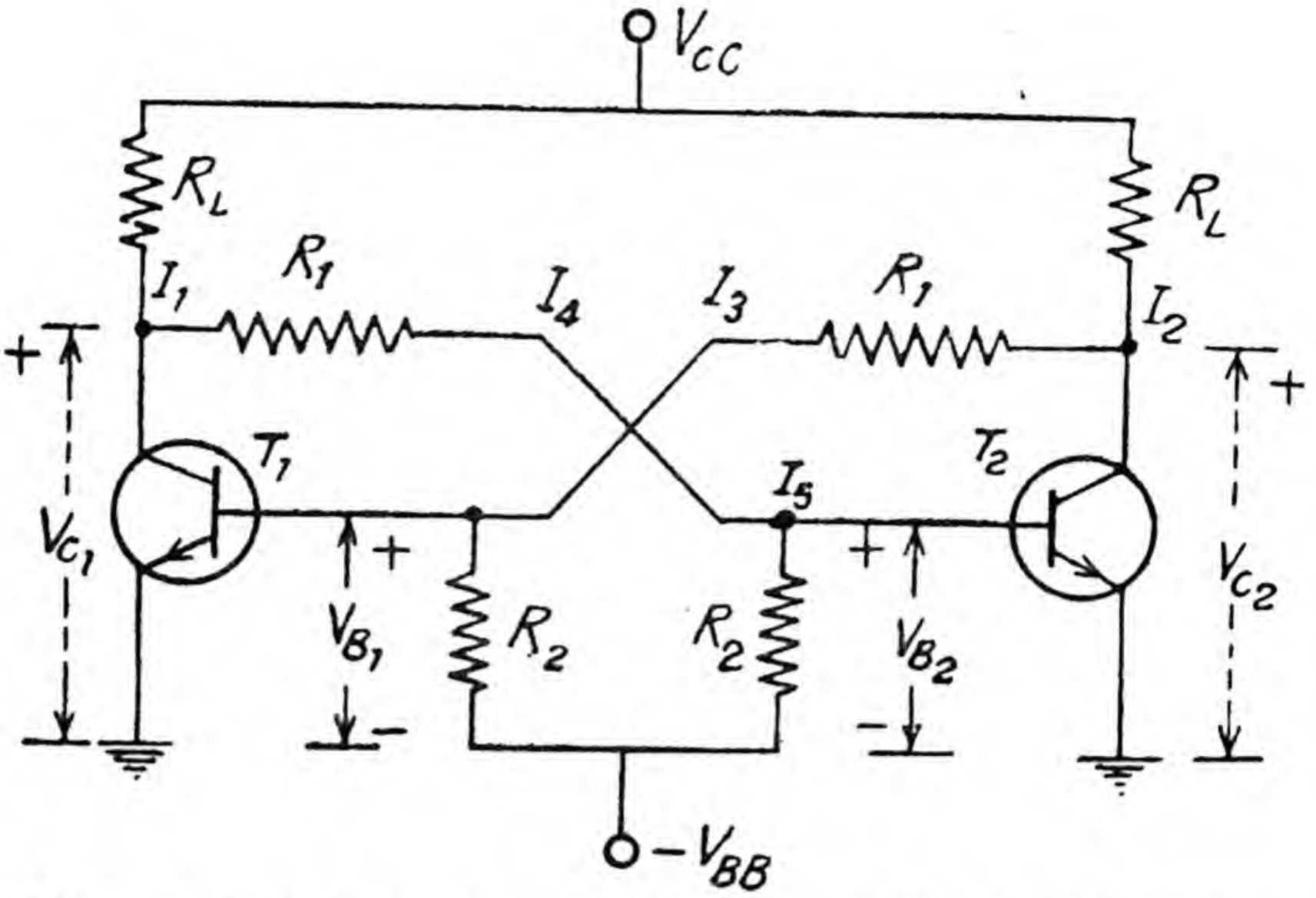


Fig. 12*a*. A fixed bias *NPN* transistor bistable multivibrator.

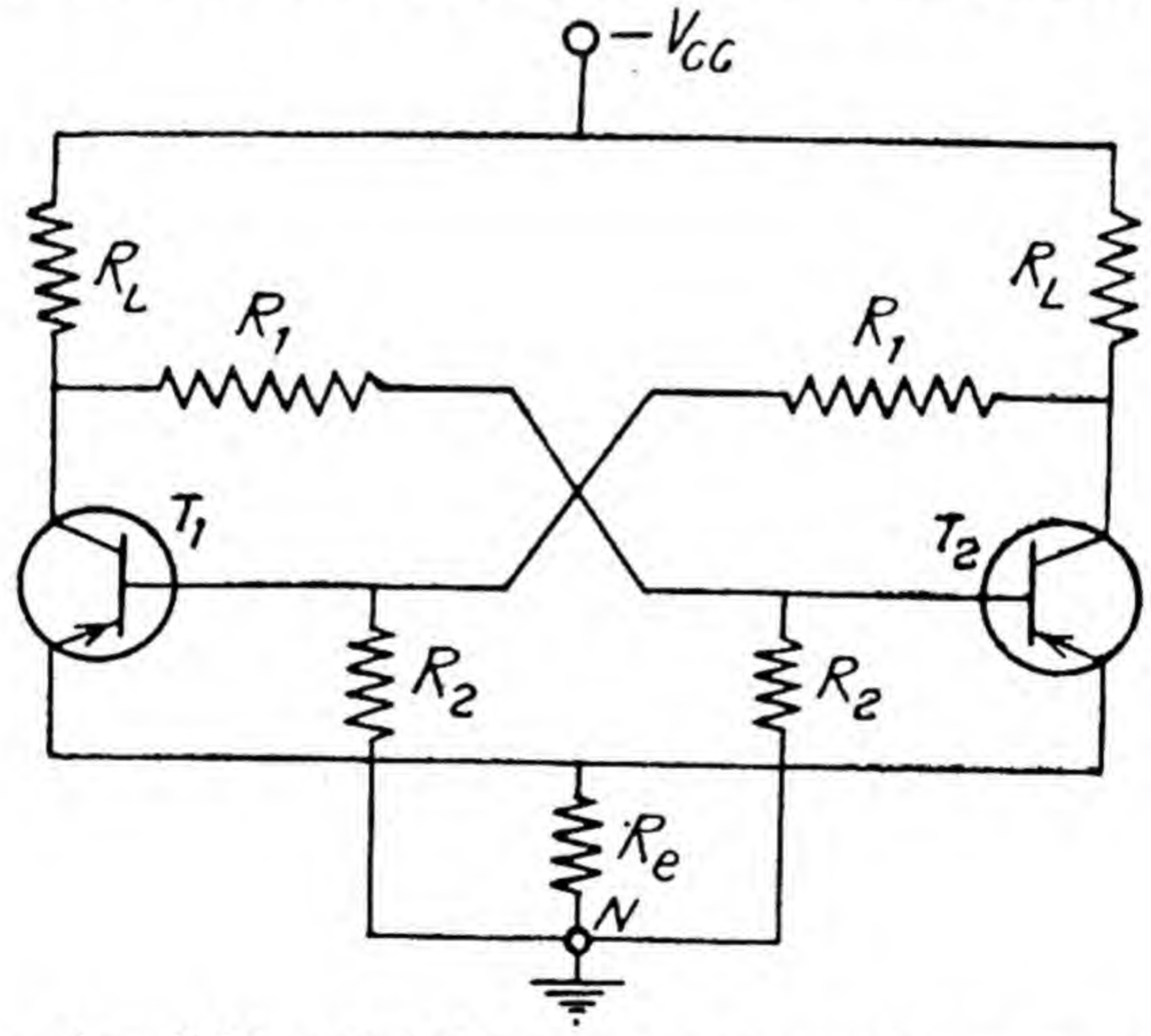


Fig 12*b*. A self-biased *PNP* transistor bistable multivibrator.

one transistor in OFF state and the other transistor in saturation. The two stable states of the bistable multivibrator are :

(i) transistor T_1 OFF and transistor T_2 ON

(ii) transistor T_2 OFF and transistor T_1 ON.

Now we shall calculate different currents and voltages when transistor T_1 is OFF and transistor T_2 is ON. When transistor T_2 is in saturation, then the collector potential is $V_{CE(sat)}$ and the base potential is $V_{BE(sat)}$. The current I_2 flowing through the load resistor R_L in series with collector of transistor T_2 is given by

$$I_2 = \frac{V_{CC} - V_{CE(sat)}}{R_L} \quad \ldots(1)$$

Again

$$I_3 = \frac{V_{CE(sat)} - (-V_{BB})}{R_1 + R_2} \quad \ldots(2)$$

The collector current $I_{C_2}=I_2-I_3$

or $$I_{C_2}=\frac{\{V_{CC}-V_{CE(sat)}\}}{R_L}-\frac{\{V_{CE(sat)}-(-V_{BB})\}}{R_1+R_2} \quad ...(3)$$

As the transistor T_1 is in OFF state, the current-flowing through this is negligibly small. The current I_1 is given by

$$I_1=\frac{V_{CC}-V_{BE(sat)}}{R_1+R_2} \quad ..(4)$$

$$I_4=I_1 \qquad (\because \text{ collector current}=0) \quad ..(5)$$

The current I_5 flowing through R_2 connected to the base of transistor T_2 is given by

$$I_5=\frac{V_{BE(sat)}-(-V_{BB})}{R_2} \quad ...(6)$$

Now the base current of transistor T_2 is given by

$$I_{B_2}=I_4-I_5$$

or $$I_{B_2}=\frac{\{V_{CC}-V_{BE(sat)}\}}{R_1+R_2}-\frac{\{V_{BE(sat)}-(-V_{BB})\}}{R_2} \quad ...(7)$$

The base potential V_{B_1} of transistor T_1 can be calculated as

$$V_{B_1}=\frac{V_{CE(sat)}\cdot R_2}{R_1+R_2}-\frac{V_{BB}R_1}{R_1+R_2} \quad ...(8)$$

The value of V_{B1} is negative as transistor T_1 is cut OFF.

When the transistors change state, the potential of the collector changes by an amount (voltage swing) given by

$$V_{C_1}-V_{CE(sat)} \quad ...(9)$$

Commutating capacitors. We know that a multivibrator remains in one of its stable states till a triggering signal such as a pulse is not applied from the external source. The time interval during

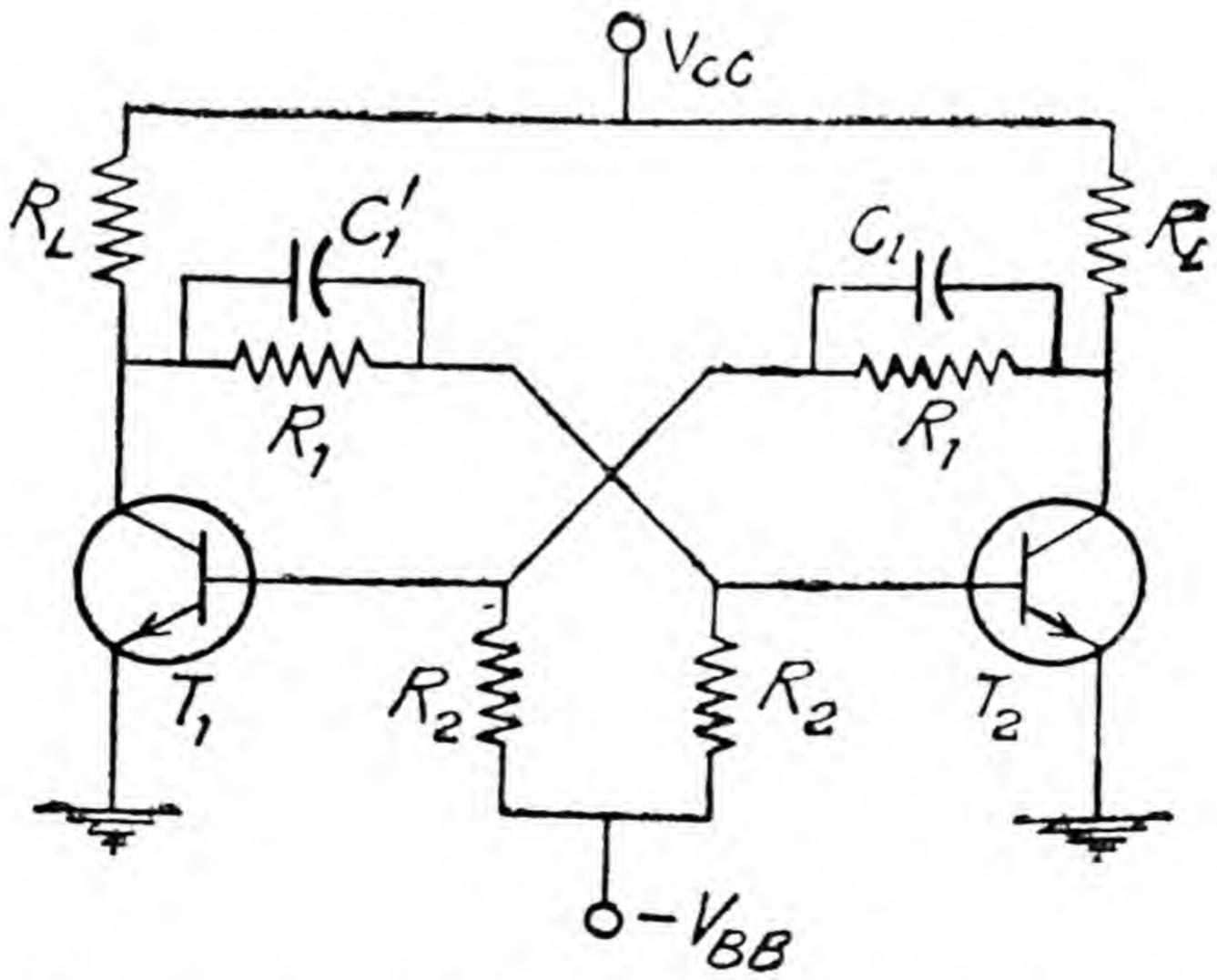


Fig. 13. Fixed bias *NPN* transistor multivibrator including commutating capacitors C_1 and C_1'.

which the conduction transfers from one transistor to another transistor is defined as *transition time*. The transition time may be reduced by introducing small capacitors in shunt with the coupling resistor R_1 as shown in fig. (13). These capacitors are known as *commutating capacitors*. They are also known as speed up capacitors. Now an additional time is required for the purpose of completing the recharging of capacitors across resistor R_1 after the transfer of conduction. This time is called as *settling time*. The sum of transition time and settling time is called as *resolution time*. It is observed that for very small values of commutating capacitors, the transition time is lengthened. On the the other hand for very large values of commutating capacitors, the settling time is increased. Hence a compromise is made for selecting the values of these capacitors.

The approximate time constant with which commutating capacitor C_1' is recharged is given by

$$\tau_1 \cong \left(\frac{R_1 R_2}{R_1 + R_2}\right) C_1'$$

and for commutating capacitor C_1, we have

$$\tau_2 \cong \left(\frac{R_1 R_2}{R_1 + R_2}\right) C_1$$

The maximum frequency of operation is given by

$$f_{max} = \frac{1}{2\tau} = \frac{1}{2(\tau_1 + \tau_2)}$$

Triggering of bistable multivibrator. To change the stable state of the bistable multvibrator it is necessary to apply an appropriate pulse in the circuit. The pluse brings both the transistors into active region and the resulting regenerative feed back results in the change of state. Triggering are of two types :

(*i*) *asymmetrical triggering, and*

(*ii*) *symmetrical triggering.*

(*i*) *Asymmetrical triggering* – In asymmetrical triggering, two trigger pulses which are derived from separate sources are connected to the two transistors T_1 and T_2 individually and sequentially they change the state of the bistable binary. Fig. (14) shows the circuit diagram of an asymmetrical triggered bistable multivibrator. In this circuit a negative triggering is used. C_3–R_5 and C_4–R_6 network is the familiar differentiating network and diode permits the transmission of negative polarity pulses only. When T_1 is ON, D_1 is back-biased by the voltage drop across R_{L1} and will not transmit a triggering pulse unless the pulse is negative and has amplitude larger than the drop across R_{L1}. Thus when

T_1 is ON, the set pulse is blocked. However, drop across D_2 is zero and hence the reset pulse will pass through D_2 to the base of T_1 switching it OFF and hence will initiate regenerative state

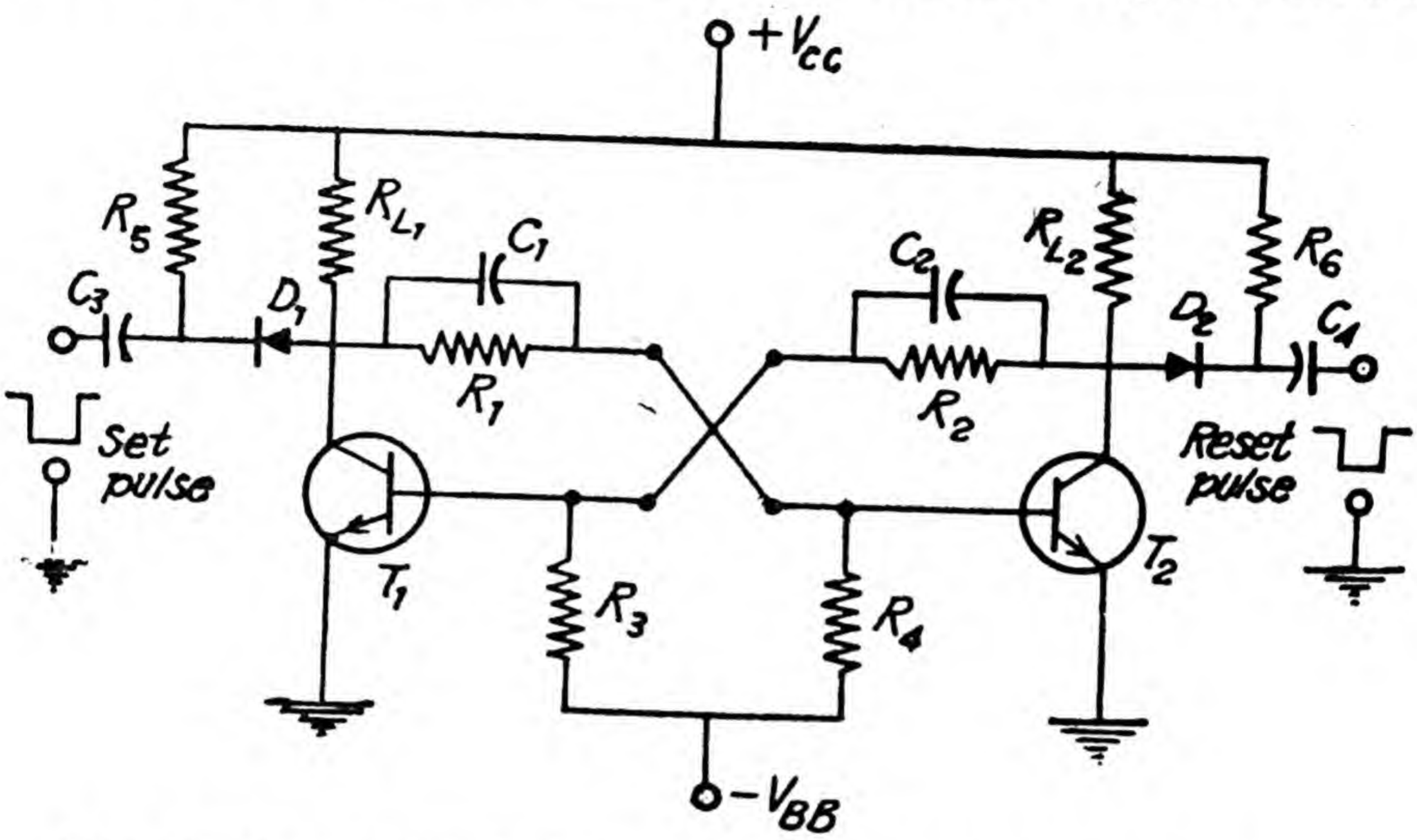


Fig. 14. Asymmetrical triggered bistable multivibrator using negative pulse triggering basic.

transition action. On the other hand, when T_1 is OFF, the drop across D_1 is zero and the negative trigger is transmitted to the base of transistor T_2 switching it OFF and transistor T_1 is switched ON. So in this case set input will be effective while reset input will be ineffective. The asymmetrical triggering is widely used in computer logic circuits.

(*ii*) *Symmetrical triggering*—In symmetrical triggering, the trigger pulses are applied only at one input and after differentiating

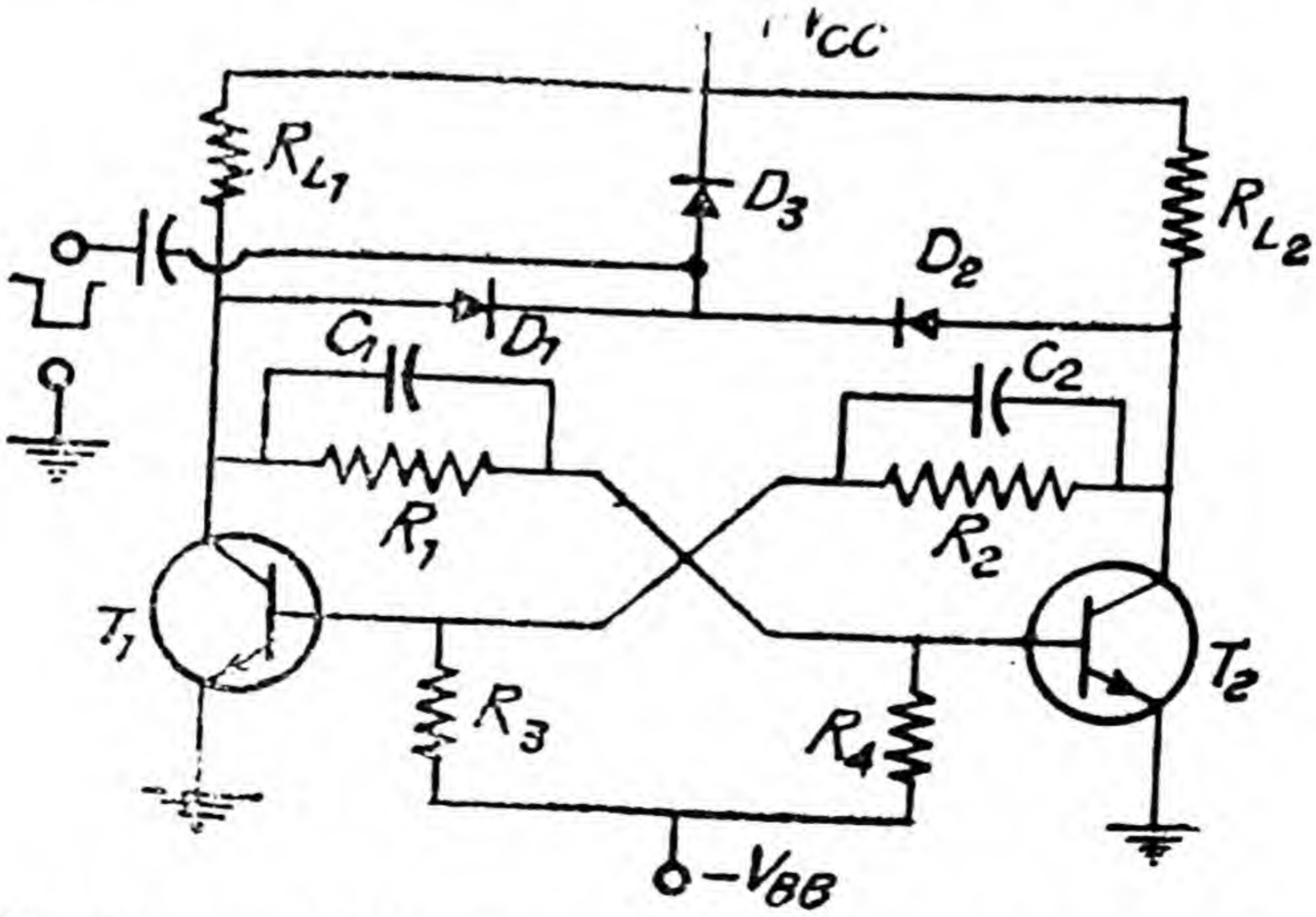


Fig. 15. Symmetrical triggering of bistable multivibrator.

they are directed (steered) to appropriate transistor sequentially. Fig. (15) shows the circuit diagram of symmetrical collector triggering in which pulse steering diodes D_1 and D_2 are used.

Suppose T_1 is ON and T_2 is OFF. In this case the voltage across R_{L1} back-biases D_1 so that it will not transmit the negative trigger while the drop across D_2 is zero (T_2, OFF) so it transmits the negative triggering pulse to the base of T_1 through R_2–C_2 network. Now a transition will take place. Let T_2 is OFF and T_1 is ON. In this case D_2 is reversed biased while D_1 is at zero volts. The next negative triggering pulse passes through D_1 to the base of T_2 through R_1–C_1 combination. Again a transition will take place. D_3 serves as a high resistance to minimize the loading on the trigger source. This triggerin g is used in counter circuits.

EXERCISES

1. Describe the switching characteristic of a transistor.
2. A rectangular pulse of voltage is applied to the base of a transistor driving it from off state to saturation. Discuss the changes in output potential. Explain the various times involved in the switching process.
3. Discuss the waveforms generated by an astable multivibrator.
4. Draw a neat circuit diagram of transistor monostable multivibrator and discuss its working.
5. Derive an expresison for the gate-widths of a transistor monostable multivibrator.
6. What is a bistable circuit ? Discuss the operation of a transistor bistable multivibrator.
7. Explain the symmetrical and asymmetrical triggering in case of bistable transistor multivibrator.

20 PULSE GENERATION AND SCALING VACUUM TUBE CIRCUITS

Recent developments in electronics require the extensive use of pulse type waves. Among them are radar, televison, telemetering, computers and a host of other applications. Since these various types of waves (such as nonsinusoidal, square and sawtooh waves, pulse waves) have assumed an important position in electronics, it is desirable to consider the specific aspects of pulse techniques under different separate heads *e.g.* pulse generators, wave shaping circuits and switching circuits. We shall discuss only vacuum tube circuits. For transistor circuits, readers are refered to the chapter on 'Multivibrators' and chapter on 'Voltage and current sweep generators'.

20·0. PULSE GENERATORS

In most pulse generators, the action is controlled by the charge and discharge of a capacitor through a resistor. During the operation, the tube either conducts heavily or is cut-off. Such circuits are called as relaxation oscillators. The most commonly employed relaxation oscillators are the multivibrator, blocking oscillator and sawtooth wave generators.

20·0-1. FREE RUNNING MULTIVIBRATORS (Astable Multivibrator).

In oscillators we have described in detail the operation of astable multivibrator and also in chapter 19. We shall not take up the discussion again except to open it for following few remarks :

(*i*) It is worthwhile to mention that the frequency of oscillation in mvltivibrator is readily controlled by an injected voltage and the wave, that is generated, is very rich in harmonics.

(*ii*) The free running multivirators do not have any stable condition of operation. Both tubes conduct alternately for short periods. These two short periods of steady operation for some times referred to as *quasi stable* states. These astable multivibrators are good enough so far as we are interested in the production of recurrent pulse waves such as square waves, rectangular waves, but in certain applications this recurrent feature is not desired and we become interested in that the multivibrator be in one stable condition of operation-one tube conducting and the other being held beyond cut-off, unless disturbed by a pulse of voltage from an external source. Such an arrangement is called one shot multivibrator or monostable multivibrator

20·0-2 A. MONOSTABLE MULTIVIBRATOR :

A stable state is the condition in which the multivibrator may remain indefinitely until the circuit is triggered by some external

source. As mentioned earlier, a monostable multivibrator also called *one shot*, *driven* and *triggered* multivibrator, has only permanent stable state, where in one tube is conducting and the other is held beyond cutoff. When a pulse of voltage is applied at the input of the circuit, conditions are reversed, *i.e.*, the tube formerly in conduction is driven into cutoff and the previously cutoff tube goes into conduction; but this state is quasi stable and the circuit returns spontaneously to its original stable state and remains in this state until another pulse is applied.

Operation :

The circuit is shown in fig. 1 (*a*). We shall describe the operation in two steps :

(i) In the absence of triggering pulse :

When any triggering pulse is not applied, the steady state exists with tube V_1 cut off and V_2 conducting. This is due to the reason that a high negative bias is applied to the tube V_1 through resistor R_1 while there is no current through R_2 and grid of V_2 is at ground potential enabling V_2 to conduct.

Since the tube V_1 is not conducting, its plate voltage is maximum to which the capacitor C_2 is charged through resistor R_3 and low grid to cathode resistance of V_2. Though a small positive voltage from the plate of V_2 is applied to the grid of V_1 through resistor R_5 yet it is much smaller as compared to negative bias applied and therefore V_1 is held well beyond cut-off.

(ii) Application of triggering pulse :

Suppose at time t_1 (see fig. 1 *b*) a positive pulse of sufficient amplitude (greater than grid bias to tube V_1) is applied to the grid of V_1 making the grid positive so that V_1 conducts. This increase in e_{c1} is shown in fig. 1 (*b*). Due to the flow of plate current in V_1, its plate potential e_{p1} falls and (since the charge on capacitor C_2 can not change instantaneously) a negative voltage e_{c2} to the grid cathode circuit of V_2 across R_2 is applied. Capacitor C_2 which was already charged to the maximum plate voltage almost equal to E_{bb}, the supply voltage, starts to discharge and develop more negative bias to tube V_2 and drives the latter beyond cut-off. Thus plate voltage of tube V_2 rises as shown by the waveform e_{p2}. This state in which V_2 is cut-off while V_1 is conducting is called *quasi-stable* state.

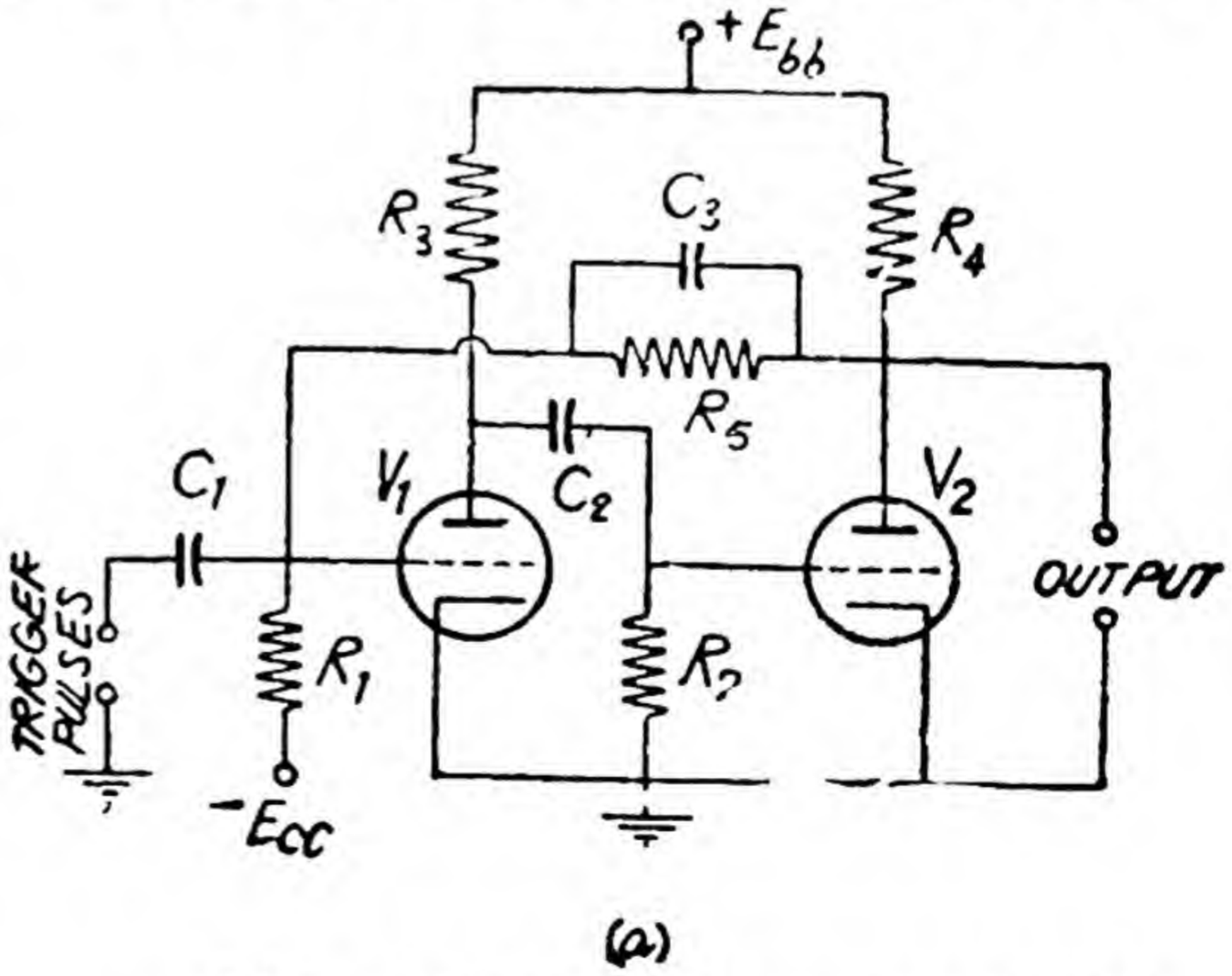


Fig. 1. (a) Monostable multivibrator.

After sometime current due to the discharge of C_2 becomes so small as to be insufficient to bias the grid of V_2 beyond cut-off and therefore V_2 starts conducting at time t_2 This decreases the plate voltage e_{p2} of V_2 and consequently, plate voltage of V_1 increases. Capacitor C_2 charges quickly, e_{c2}, going slightly positive and then dropping exponentially to zero as C_2 assumes maximum charge. T his further increases the plate current of V_2 and bias the tube V_1 beyond cut-off. Thus the original stable state is again restored.

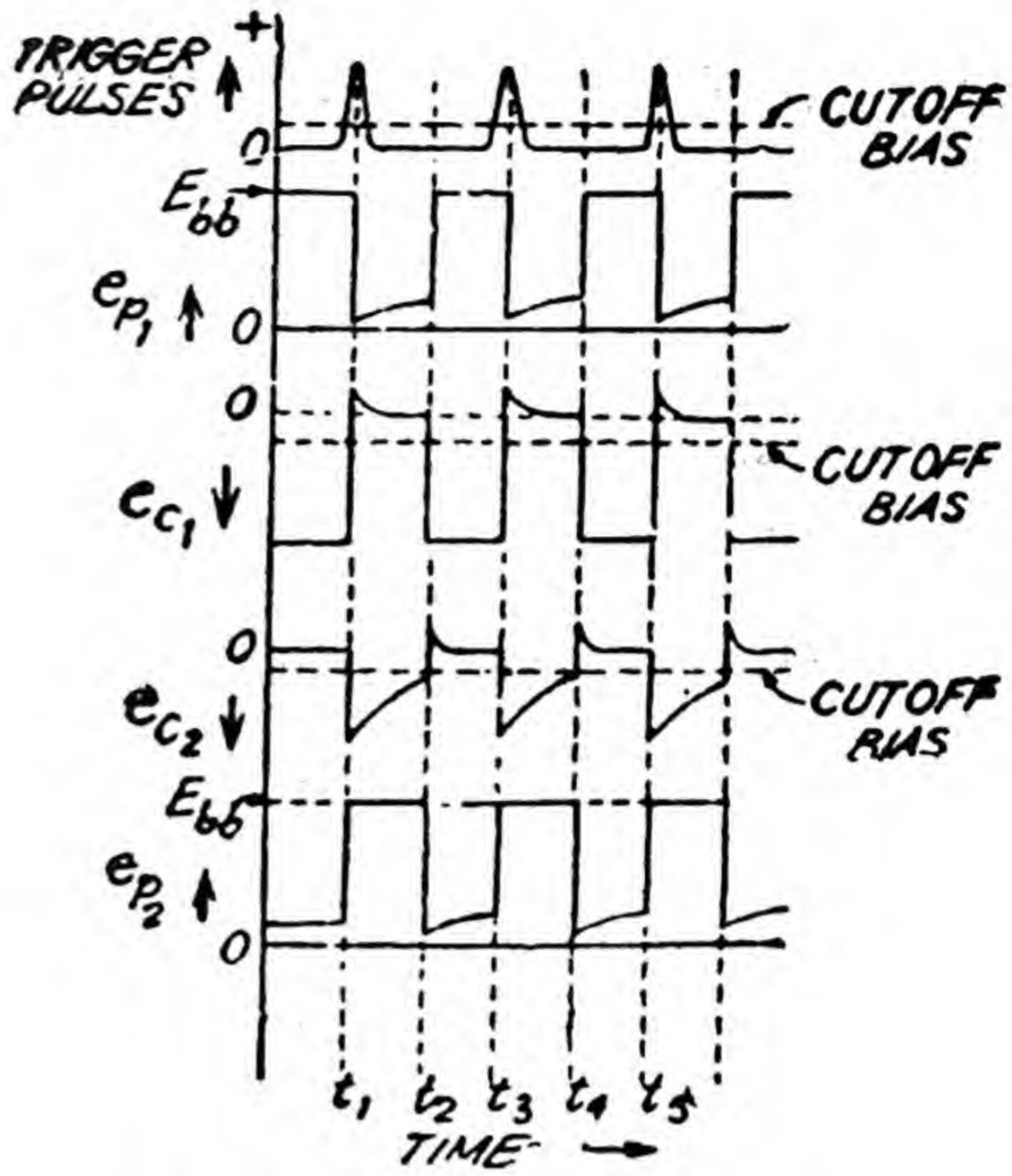


Fig. 1. (b). Waveforms.

Output Voltage : The output voltage is taken from the plate of V_2 and is normally low. We note from the waveform e_{p2} that when trigger pulse is applied. e_{p2} shoots up to the maximum value an d remains at this value for a period $(t_2 - t_1)$ for which tube V_2 does not conduct and then comes down to minimum value when tube V_2 starts conducting and remains at this value for a period $(t_2 - t_2)$ when the next trigger pulse is due. Therefore *a single trigger pulse produces one complete cycle.*

We note from fig. 1 (*b*) that **output pulse duration depends upon :**

(*i*) *Maximum negative voltage of* e_{c_2} **:** greater the value of e_{c_2} more time will be required to raise the bias above cutoff value for the tube V_2 to conduct and hence increasing the duration $(t_2 - t_1)$.

The maximum value of e_{c2}, in turn, depends upon the tube characteristics and on load resistor R_3, *i.e.* greater the value of R_3 and the increase in plate current, more will be e_{c2}.

(*ii*) *Larger time constant* R_2C_2 : larger the time constant R_2C_2, more time will be required for the condenser C_2 to discharge *i.e.* it would delay the conduction of tube V_2 and hence amounts to the increase in pulse duration. A monostable multivibrator is often used to delay the application of a pulse to a circuit by a time equal to multivibrator duration time. It is also used for pulse equilisation.

20·0–2B. BISTABLE MULTIVIBRATOR : Flip Flop Circuit :

In the bistable multivibrator, there are two stable conditions of operation *e g.* V_1 can be in heavy conduction with V_2 in cutoff, or V_2 can be in conduction with V_1 cutoff. Either of these states is stable, *i.e.*, no reversal of any of the operation states unless triggered by an external signal. Thus in bistable multivibrator an output pulse is obtained only if a triggering pulse is applied to the input. A full cycle of output is produced for every two triggering pulses properly applied and of correct polarity and magnitude. This circuit is also called *Eccles—Jordon trigger circuit* after its inventors, *flip flop* circuit due to its circuit action, *trigger* circuit because it controls other circuit actions and *binary* because it is used in a two number (zero and one) or binary counting system.

Operation :

The circuit is shown in fig. 2. We shall describe the operation in three steps :

(i) **In the absence of any trigger pulse :** When plate voltage is applied, each tube conducts. One tube always conducts more than the other, however slightly, due to such factors as inherent tube noise and unequally distributed emission along the length of cathode. This unbalance of plate currents starts multivibrator action.

Suppose tube V_1 conducts more than V_2. Then plate voltage of V_1 decreases. This decrease in plate voltage of V_1 is applied to grid of V_2 through R_3R_6 voltage divider and, in its effect, reduces the plate current of V_2, increasing thereby the plate voltage of V_2. This increase in the plate voltage of V_2 is applied to the grid of tube V_1 through the potential divider R_4R_5 making the grid more positive and further reduces the plate voltage of V_1. This decrease in plate voltage of V_1, as described earlier, makes the grid of V_2 more negative. The action is cumulative and finally the tube V_2 is driven to cutoff and now further decrease in grid voltage can not increase the plate voltage of V_2. Thus steady state is reached *i.e.*, the current will remain in this mode of operation for infinite time without any reversal unless triggered by some appropriate external pulse.

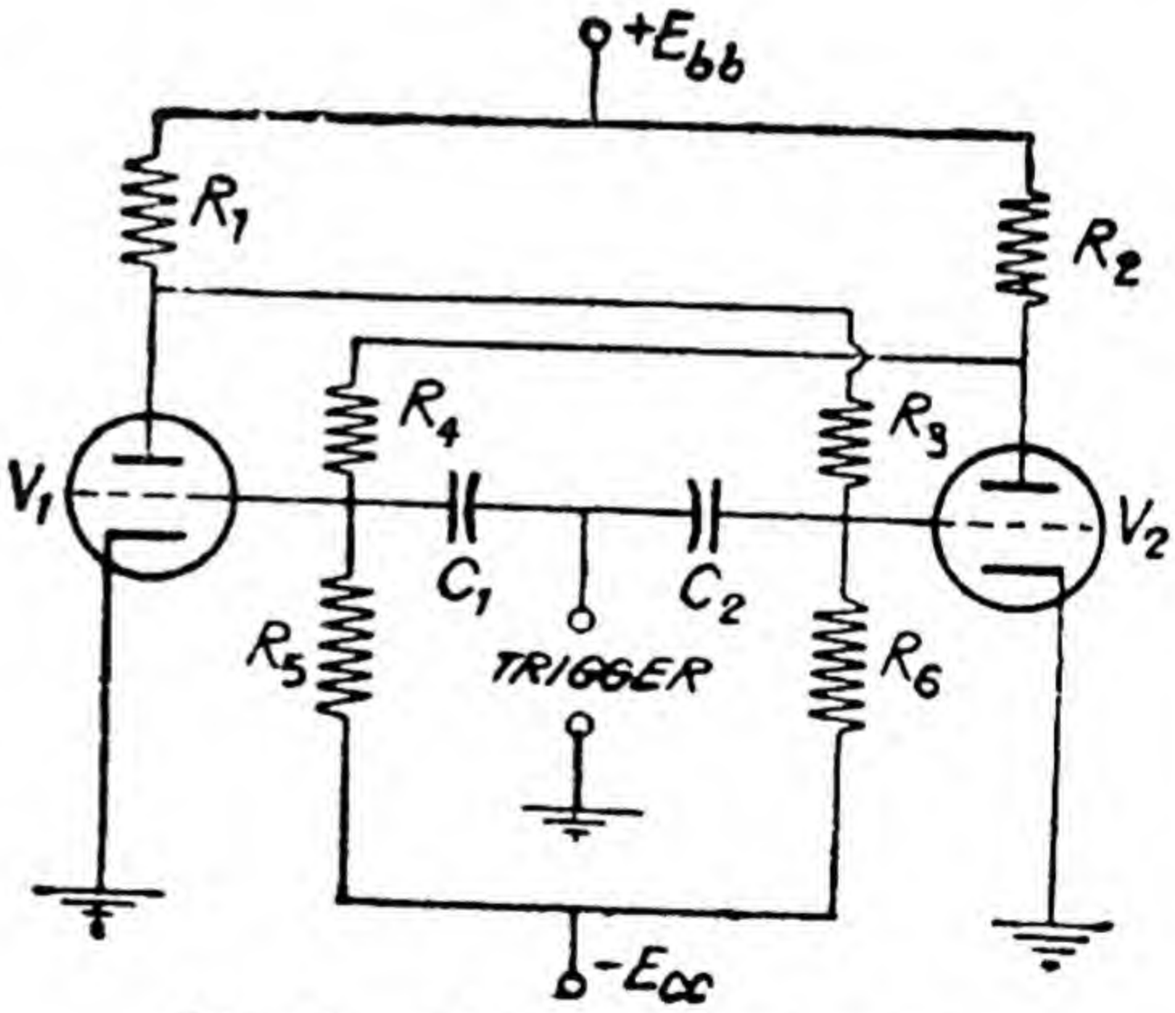


Fig 2. Bistable multivibrator.

(ii) **Application of first trigger pulse :** When a sharp narrow positive pulse is applied at the trigger input, tube V_1 remains unaffected as it is already conducting but if the pulse is of sufficient

magnitude the grid of V_2 will be raised above cutoff and the tube V_2 starts conduction.

Due to the conduction of V_2, its plate voltage decreases which makes the grid of V_1 less positive and thereby increases plate voltage of V_1. This increase in the plate voltage of V_1 makes the grid of V_2 more positive, increasing, therefore, the plate current of V_2. This action further reduces the plate voltage of V_2, making the grid of V_1 to approach to cutoff voltage. The action is cumulative and tube V_1 is finally cutoff and V_2 continues to conduct for an indefinite period. This state of operation is stable and reversal is possible only when some external triggering pulse is input to the circuit.

(iii) Application of second trigger pulse : Application of second triggering positive pulse will cause V_1 to conduct and V_2 to be cutoff. We note that, by the application of second pulse, we arrive at the initial stable state of operation, *i.e.* a full cycle of output is produced for every two triggering pulses properly applied and of correct polarity and amplitude.

A negative triggering pulse also causes the reversal of steady state. Suppose V_1 is conducting and V_2 is cutoff and a negative triggering pulse is applied. Then V_2 already cutoff will not be affected, but the grid voltage of V_1 would decrease, thereby increasing plate voltage of V_1. This increase in the plate voltage of V_1 will rise the grid voltage of V_2. If the pulse is of sufficient amplitude so as to drive, through the above indirect action, the grid of V_2 above cutoff tube V_3 would start conducting and a reversal of steady state V_2 conducting and V_1 cutoff will result.

Self biased binary : The need for a negative supply, E_{cc} as in fig. 2. may be eliminated by using a common cathode resistor R_K to provide self bias as shown in fig. 3.

The drop across R_K is normally the same for two stable states. When the cathode current of one tube decreases, that of other increases, so that the voltage variation across R_K is small and therefore, there is no necessity of bypass capacitor.

Condensers C_1 and C_2 are called *speed up* capacitors (each of value 50 $\mu\mu f$). As these capacitors cannot charge or discharge instantaneously, any change in the plate voltage of one tube is *rapidly* transmitted

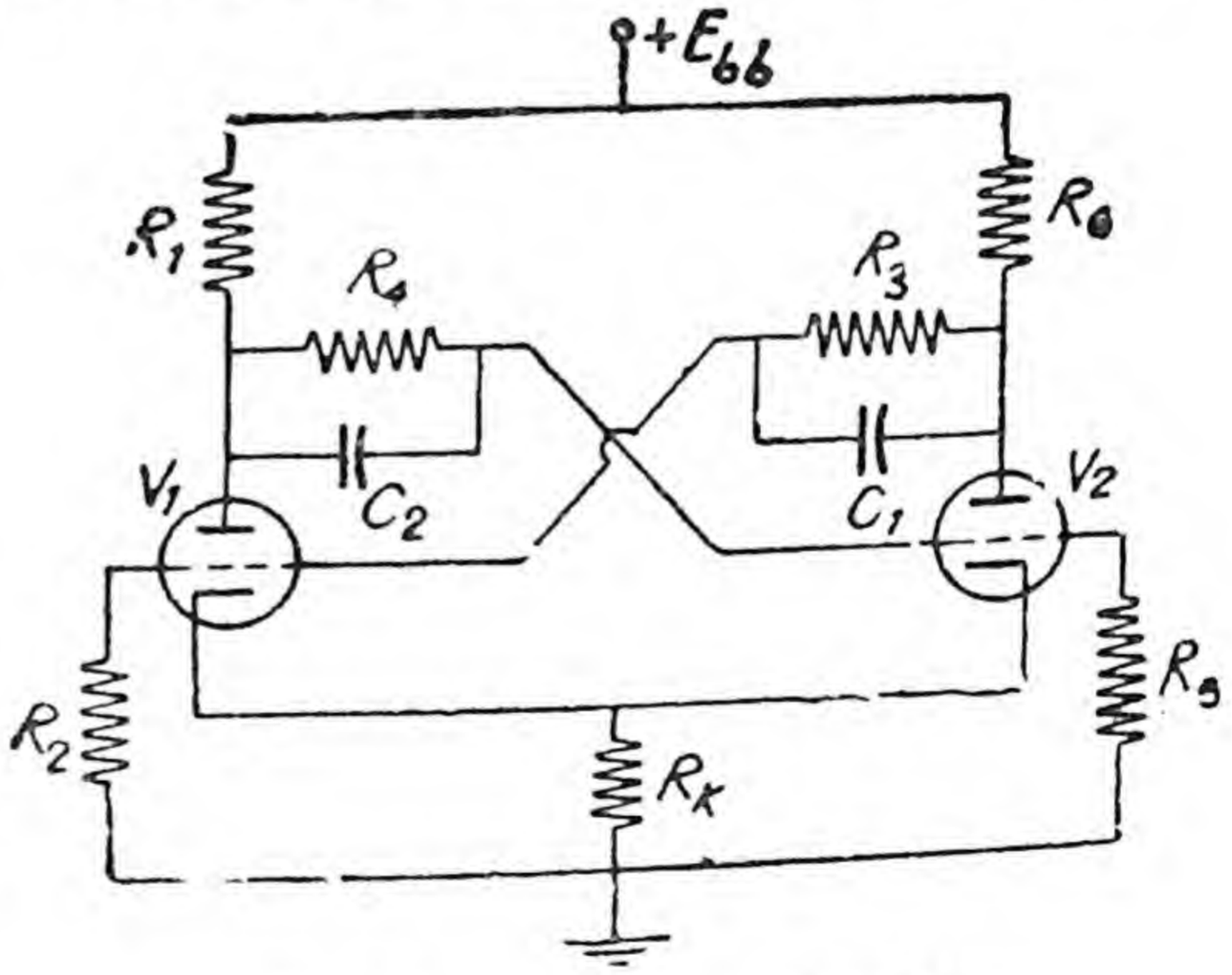


Fig. 3. Self biased binary.

to the grid of the other tube. This, therefore, enables to binary the trigger pulses of *short duration*.

Application of Bistable Multivibrator : Scale of 2 counter : A very important application of the flip flop is the basic scale of two circuit, shown in figure 4. This circuit produces an outgoing pulse for every second input pulse

A neon lamp is introduced in the plate grid circuit of valve V_1. When V_1 does not conduct, its plate voltage approaches E_{bb} value and neon lamp glows, while when V_1 conducts voltage across the lamp falls and it is extinguished. Suppose initially V_1 is cutoff and V_2 is conducting. When a positive trigger pulse is applied to the tube V_1 causing the plate current of V_1 to flow, then this increased current will drive the cathode of V_2 positive. The effect is to reduce the plate current of V_2 or to increase the plate voltage of V_2 which inturn, reduces the bias on tube V_1 driving its grid to positive voltage. This will continue till the tube V_2 is cutoff and V_1 is conducting. In this steady state neon lamp does not glow.

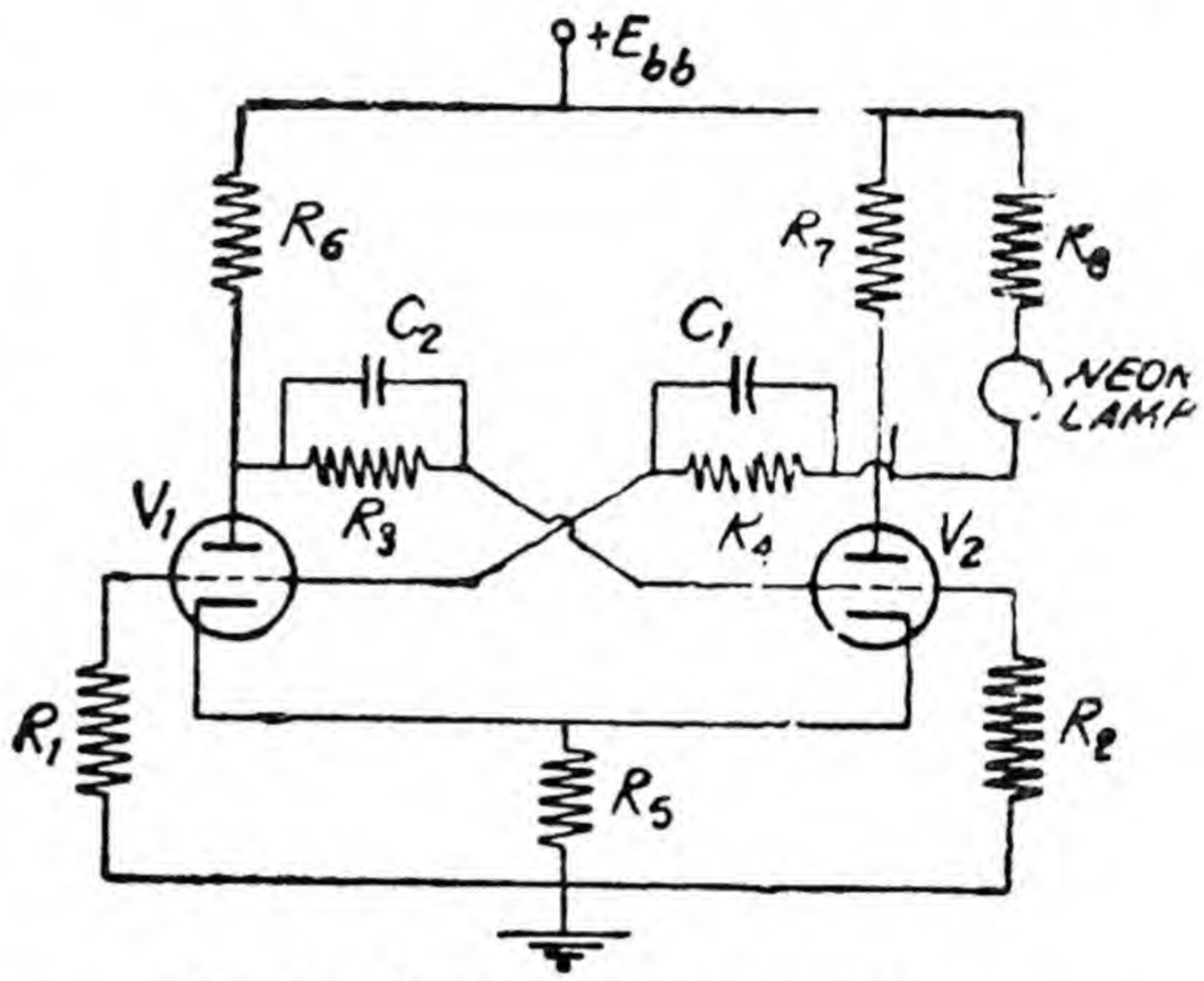


Fig 4. Scale of two counter.

Next, when the second positive trigger pulse is applied to the circuit, then steady states, as explained in binary operation, will be reversed *i.e.* we come to initial state in which V_1 is cutoff and V_2 is conducting. This will result in the glow of neon lamp. Thus for each second input pulse, neon lamp glows or in other words we achieve a scale of two counting.

If the two such flip flop circuits are connecting in cascade, the first circuit would register every *second* pulse, while the second every *fourth* pulse. Or in general, if several such flip flops are arranged in cascade, each subsequent flip flop will record second, fourth, eighth, etc. pulse. Thus it is obvious that 'flip flop' is the heart of scaling circuits.

20·0-3. BLOCKING OSCILLATOR :

Blocking oscillator is used for the generation of short pulses of duration about 0·05 to 0·25μ sec. The basic circuit is shown in fig 5 and consists of a triode and a pulse transformer. The transformer primary is connected in the plate circuit and secondary in

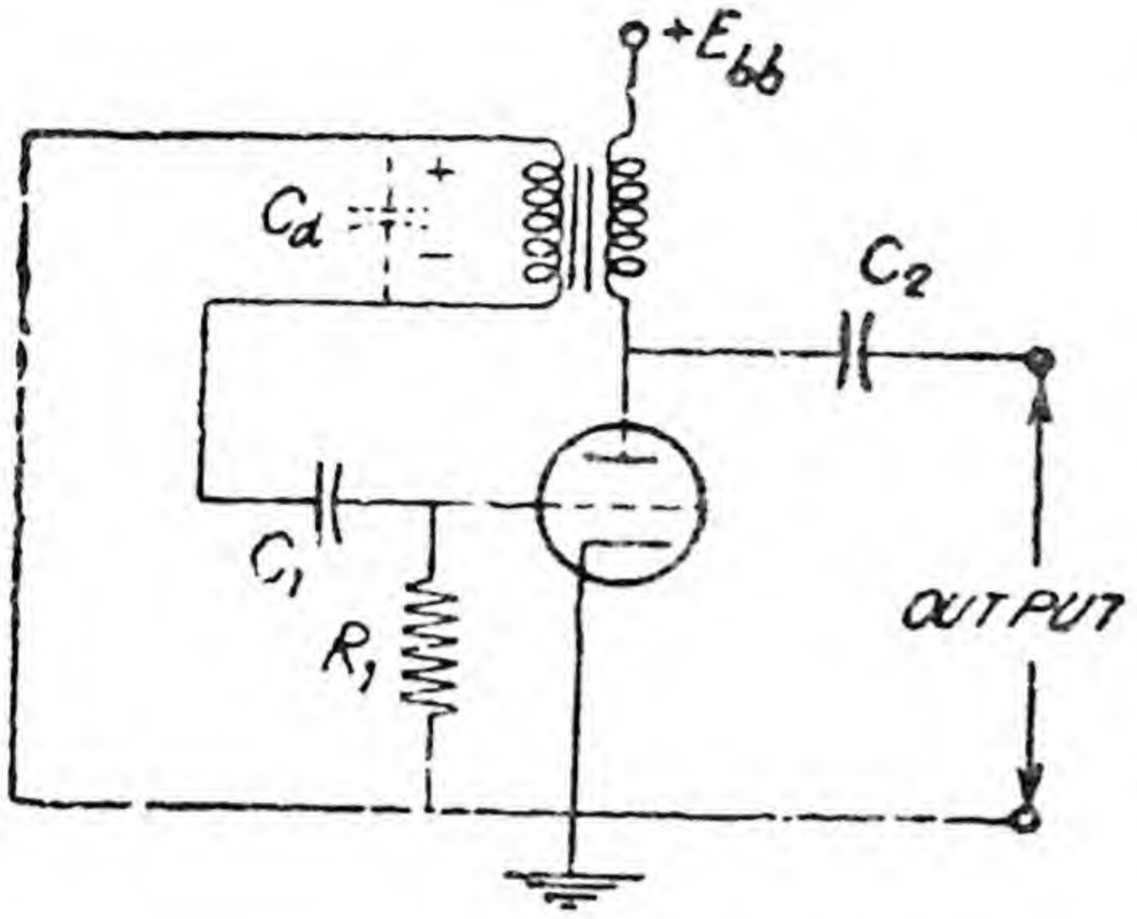


Fig. 5. Basic Blocking Oscillator.

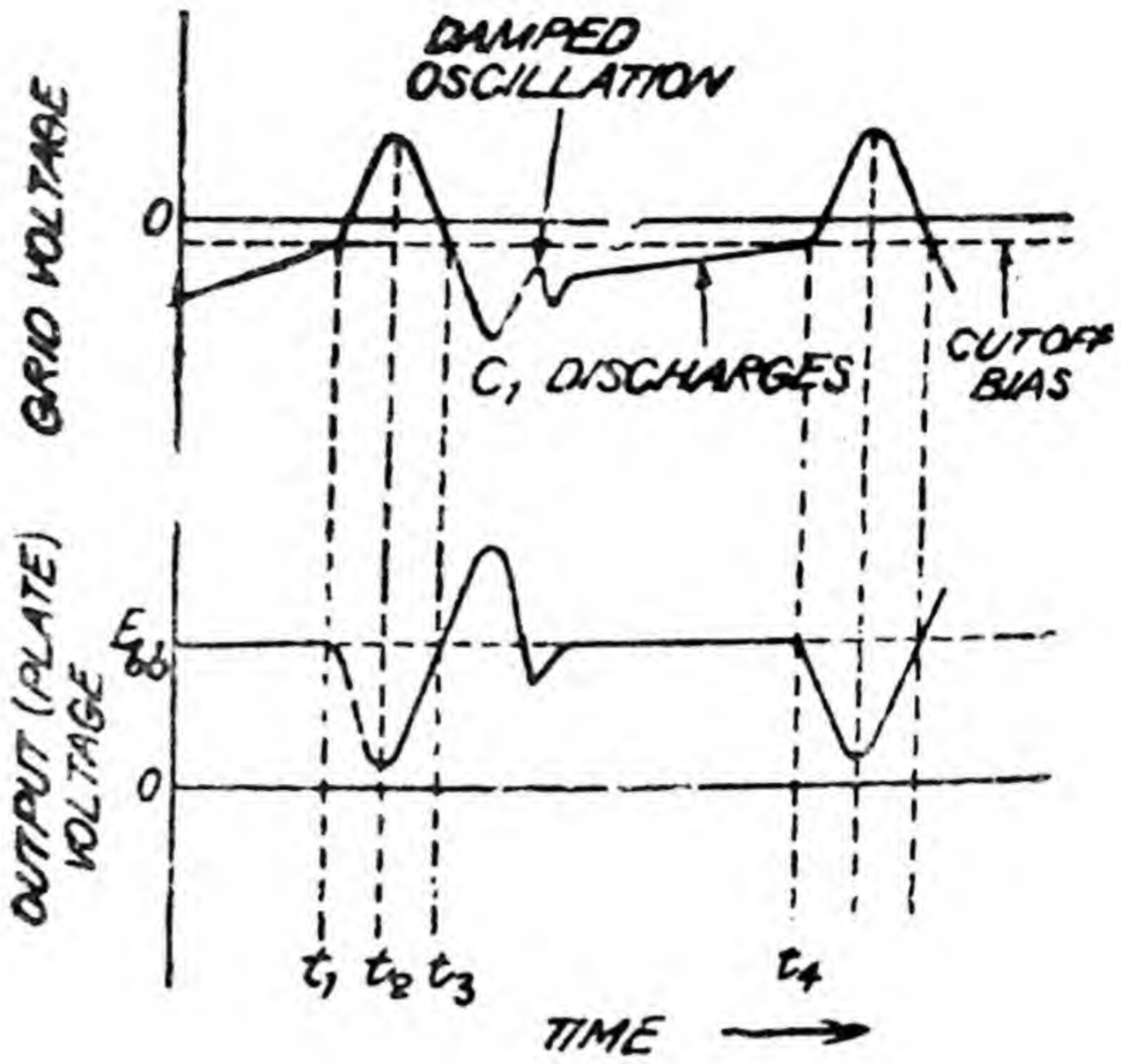


Fig. 6. Blocking Oscillator Waveforms.

the grid circuit. Winding are such as to provide polarity inversion, *i.e.*, when a plate voltage decreases due to the rise in plate current, positive voltage is coupled to the grid and since the grid now becomes more positive, plate current would further increase, decreasing thereby the plate voltage. The action will continue till the plate current reaches saturation so that the further increase in grid voltage will not affect the magnitude of plate voltage.

The distributed capacitance C_d forms a resonant circuit with secondary of transformer. This capacitor is charged by the voltage developed across the secondary windings. When the voltage decreases, the resonant circuit oscillates and the grid then swings, from a positive voltage to a highly negative voltage. During the positive half cyc'e, capacitor C_1 charges to the value of the positive grid voltage through the low grid to cathode resistance of the tube. On the negative half cycle, since the capacitor cannot discharge instantaneously the full negative charge is applied to the grid and drives the tube beyond cut-off. Since the tube is not conducting, its grid to cathode resistance will be high and therefore, the condenser will discharge through resistor R_1.

Operation : (*i*) *At instant* t_1 :

Suppose the operating potentials are applied at instant t_1. The plate current will increase and a positive voltage is coupled to the grid. Therefore grid voltage goes on increasing while the plate voltage continuously decreases. A stage comes when the plate current reaches saturation. As there is no further increase in plate current no more positive voltage will be coupled to the grid current. and grid acquires the maximum voltage. This occurs at instant t_2. Consider C_1 charges to the value of maximum grid voltage through low grid to cathode resistance.

(*ii*) *At instant* t_2 : For a short moment, after reaching at saturation value, the plate cur.ent does not change and C_1 begins to discharge through grid resistor R_1. This makes the grid negative decreasing the plate current. As the plate voltage increases, negative voltage will be coupled to the grid circuit *i.e.*, grid becomes more and more negative. Further, condenser C_1 also discharges with the passage of time. The net result is that the grid voltage goes down from its maximum positive voltage to zero and then in negative direction.

(*iii*) *At instant* t_3 : Oscillations continue in the tuned circuit formed by the transformer windings. To damp these oscillations quickly, transformer is designed with low Q value. Moreover as the tube is cut-off, energy supply to the oscillatory circuit is also stopped with the result that the oscillations damp quickly.

The duration for which the tube is cut-off depends primarily on time constant R_1C_1 because when C_1 discharges sufficiently to reach cut-off, conduction starts. The instant t_4 is reached.

(*iv*) *At instant* t_4 : Tube starts conduction and whole previous cycle is repeated.

Application : (*i*) Blocking oscilla'ors are commonly used as frequency dividers

(*ii*) They are also used as counters.

iii) They are also used as a low impedance switch used to discharge a capacitor quickly.

(*iv*) They can produce very sharp pulses at different repetition frequencies and especially at very low repetition frequencies.

20·0-4. SAW TOOTH WAVE GENERATOR :

In cathode ray tubes, we require a voltage that varies linearly with time to display the information. This voltage, called a sweep, causes the electron beam to move horizontally across the screen with a known velocity. The selection of the form of the sweep is important so as to ease the display of information on the screen. It is usually desired that the spot moves across the screen as linear function of time and fly back atonce to the starting point at the end of each sweep. Such a time base requires a wave of sawtooth form with linear rise time and abrupt return to zero as shown in fig. 7.

We know that potential across a condenser is

$$e_c = \frac{1}{C}\int_0^t i\,dt,$$

where initial charge has been taken as zero If the charging current to capac tor is supplied by pentode

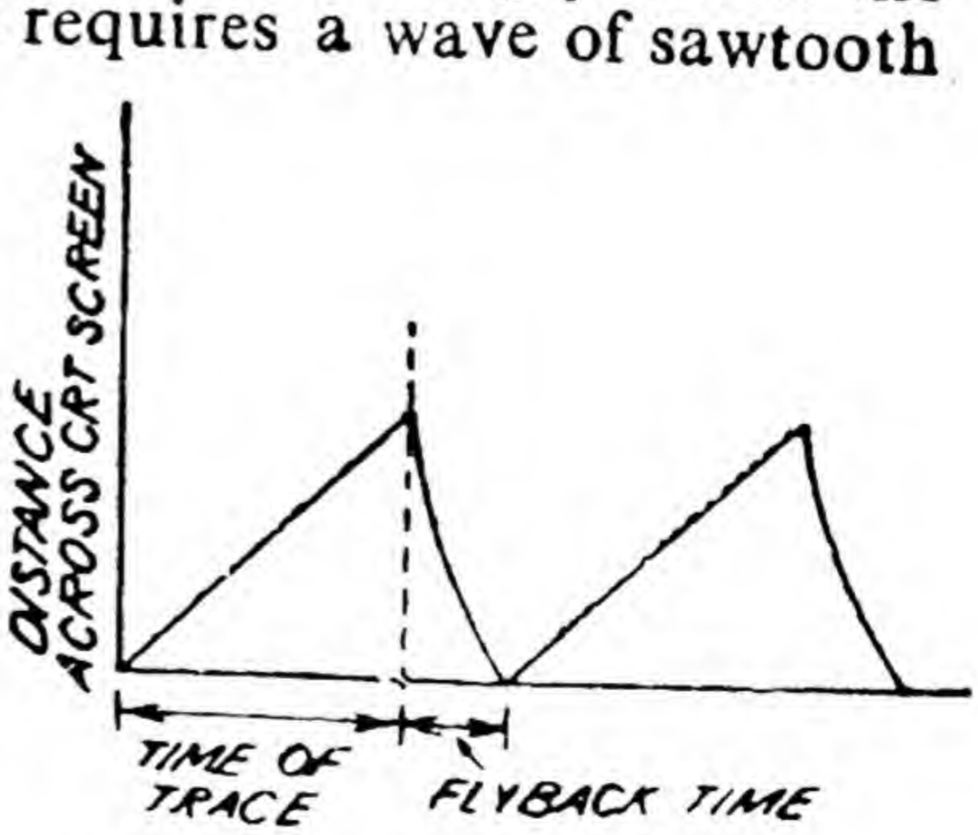


Fig. 7. Sawtooth Sweep.

then, since it will be independent of changes in the plate voltage (when high plate voltage is used) over a wide range, we can treat it as constant so that

$$e_o = \frac{1}{C} \; i \int_0^t dt = \frac{it}{C}$$

or $e_o \propto t$ *i.e.*, potential across condenser rises linearly with time. Means are also provided to discharge the condenser rapidly *e.g.*, a glow discharge tube is connected across the condenser.

(i) Neon time base circuit : The circuit is called simple *RC* sweep generator. We note that sweep voltage curve is not linear and this circuit is not used in commercial cathode ray tubes. Moreover, the device suffers from the disadvantage that striking and and extinction potentials of a neon tube are not particularly stable and because sweep frequency is the function of both these potentials, time base is distorted, resulting in the distortion of information presented on cathode ray tube screen.

(ii) Thyratron Sweep Generators : To synchronise the input pattern over a wide frequency range, neon lamp (gas diode) is not suited. For these reasons, thyratron tubes are used rather than gas diodes.

A thyratron time base generator with improved linearity is shown in figure 8. As explained earlier, plate current of pentode is constant, and therefore charging of a capacitor by pentode plate current will be linear.

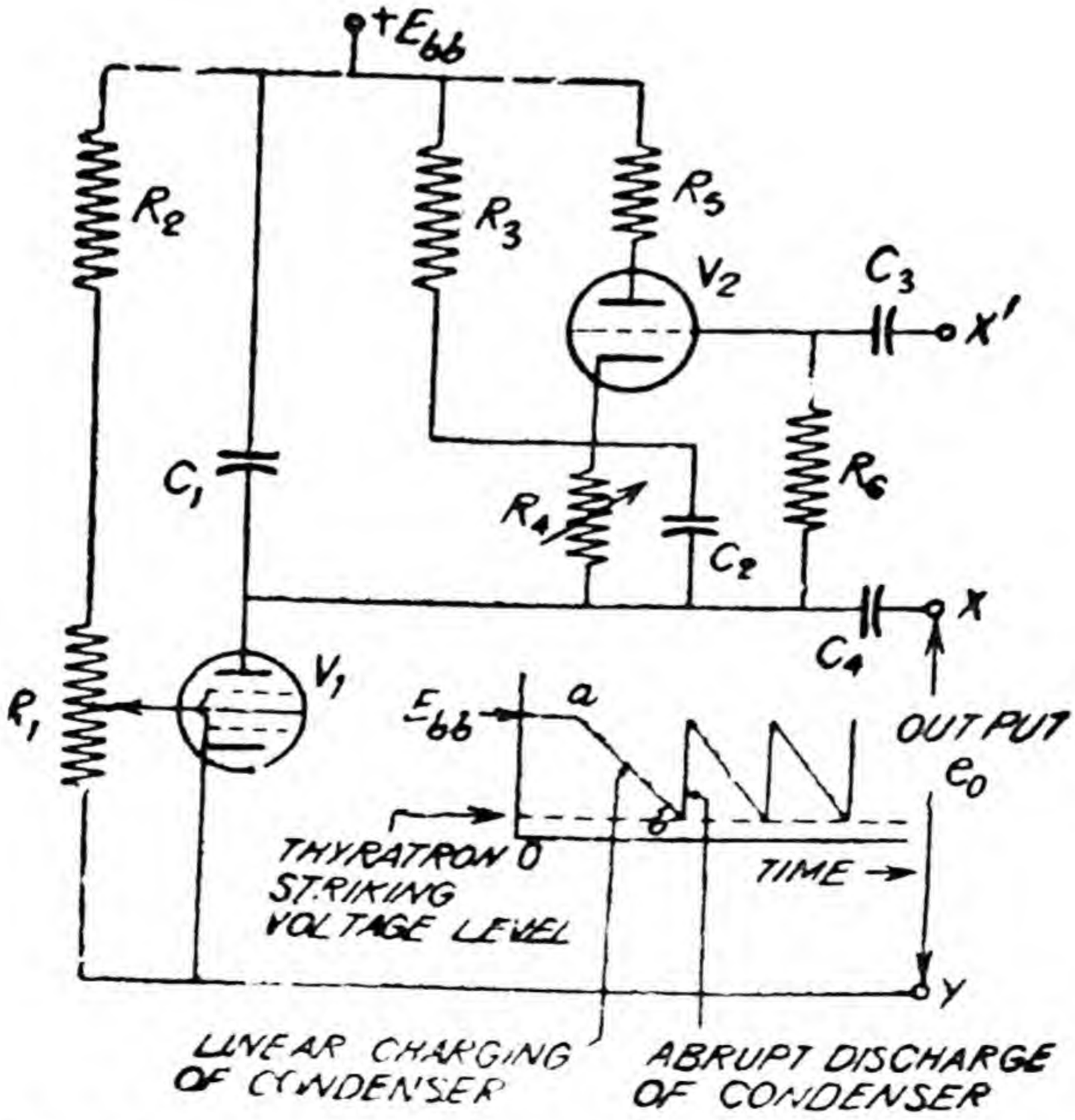


Fig. 8. Constant current pentode sweep generator.

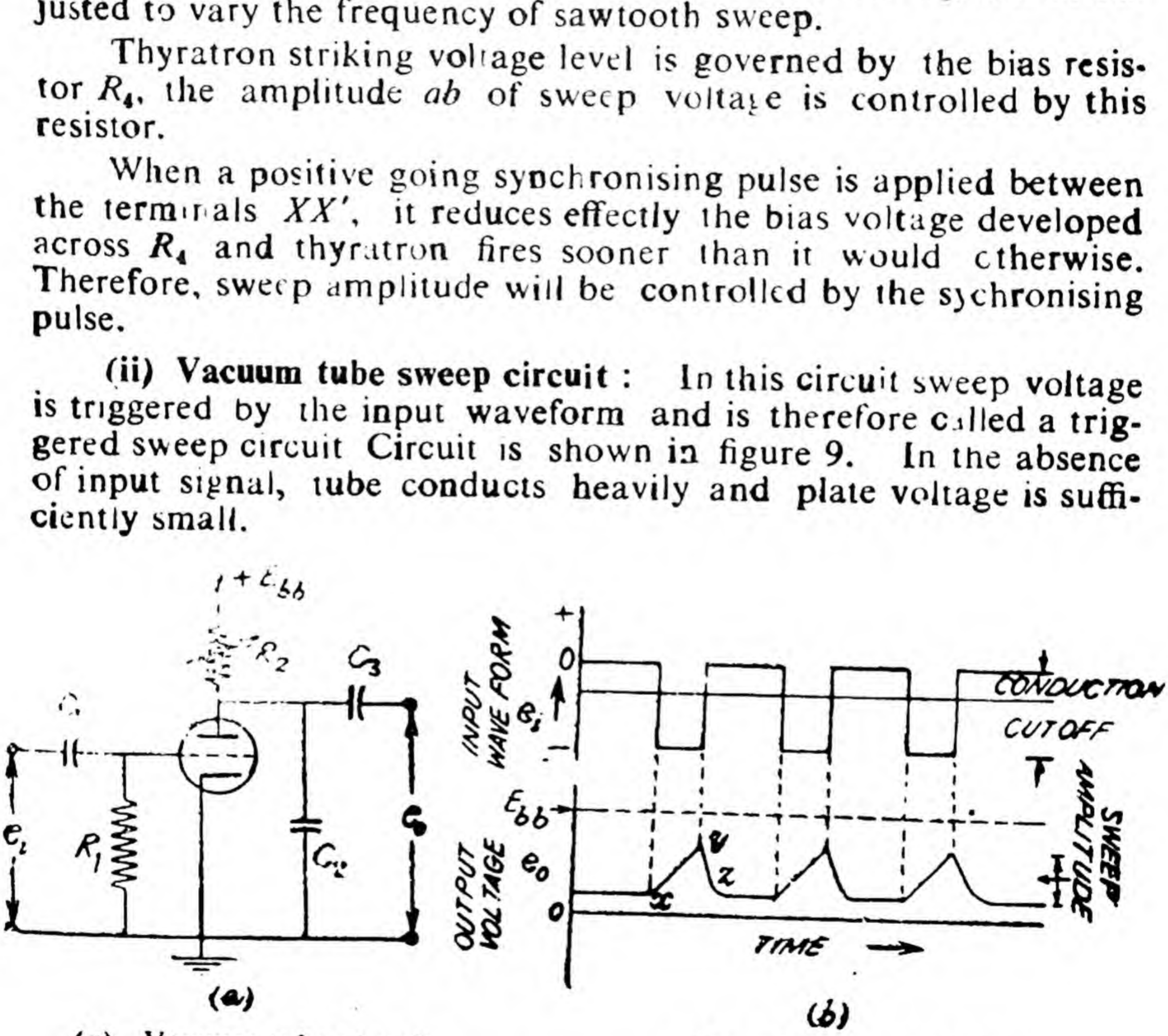
Operation. Capacitor C_1 is initially uncharged and the output voltage between the terminals XY is equal to the plate supply voltage E_{bb}. The capacitor then charges linearly and rapidly through pentode V_1, causing a similar decrease in the output voltage. When the voltage across the condenser reaches the striking voltage of thyratron V_2; the latter conducts because it is connected across the condenser and prevents the further increase in the voltage across C_1. The capacitor C_1, then discharges through V_2 at a quick rate determined by the value of R_5 The output voltage again reaches a value E_{bb}. The cycle then again starts.

Variable resistor R_1 which controls the plate current of pentode V_1 and hence the charging rate of condenser C_1, can be adjusted to vary the frequency of sawtooth sweep.

Thyratron striking voltage level is governed by the bias resistor R_4, the amplitude ab of sweep voltage is controlled by this resistor.

When a positive going synchronising pulse is applied between the terminals XX', it reduces effectly the bias voltage developed across R_4 and thyratron fires sooner than it would otherwise. Therefore, sweep amplitude will be controlled by the sychronising pulse.

(ii) Vacuum tube sweep circuit : In this circuit sweep voltage is triggered by the input waveform and is therefore called a triggered sweep circuit Circuit is shown in figure 9. In the absence of input signal, tube conducts heavily and plate voltage is sufficiently small.



(*a*) Vacuum tube circuit. (*b*) Sweep waveform.

Fig. 9. Vacuum tube circuit and waveform.

When input voltage e_i is applied : The plate current decreases as the input goes down to negative and the plate potential increases. The negative going input then drives the tube beyond cut-off and plate current stops and plate voltage increases at a rate determined by time constant R_2C_2 (portion xy in fig. 9 *b*) to reach E_{bb} value. Then e_i returns to zero, removing the negative bias and

the tube conducts heavily, rapidly discharging the capacitor C_2 (portion *yz*), The output voltage then again reaches the small initial value.

It is obvious that generation of sawtooth waveform is subject to the application of input waveform or in otherwords frequency of the output waveform is controlled by the input pulse. Pulse duration of voltage e_i also controls the amplitude of e_0 for a given time constant R_2C_2. Thus circuit generates a sweep waveform acurate as to starting time and duration.

20·0-5. CIRCUIT TO IMPROVE THE LINEARITY OF THE VACUUM TUBE SWEEPS :

We shall discuss here two circuits-one the *bootstrap* circuit and the other *Miller integrating* circuit to improve the linearity of vacuum tube sweeps Refer to chapter 22 for more detailed study involving transistorized circuits

(i) Bootstrap Circuit : In art. 20·0-4 we have explained that sweep is linear when the charging current is constant. Bootstrap circuit employs a form of feed back through capacitor C_2 to linearise *RC* sweep voltage and is shown in fig. 10.

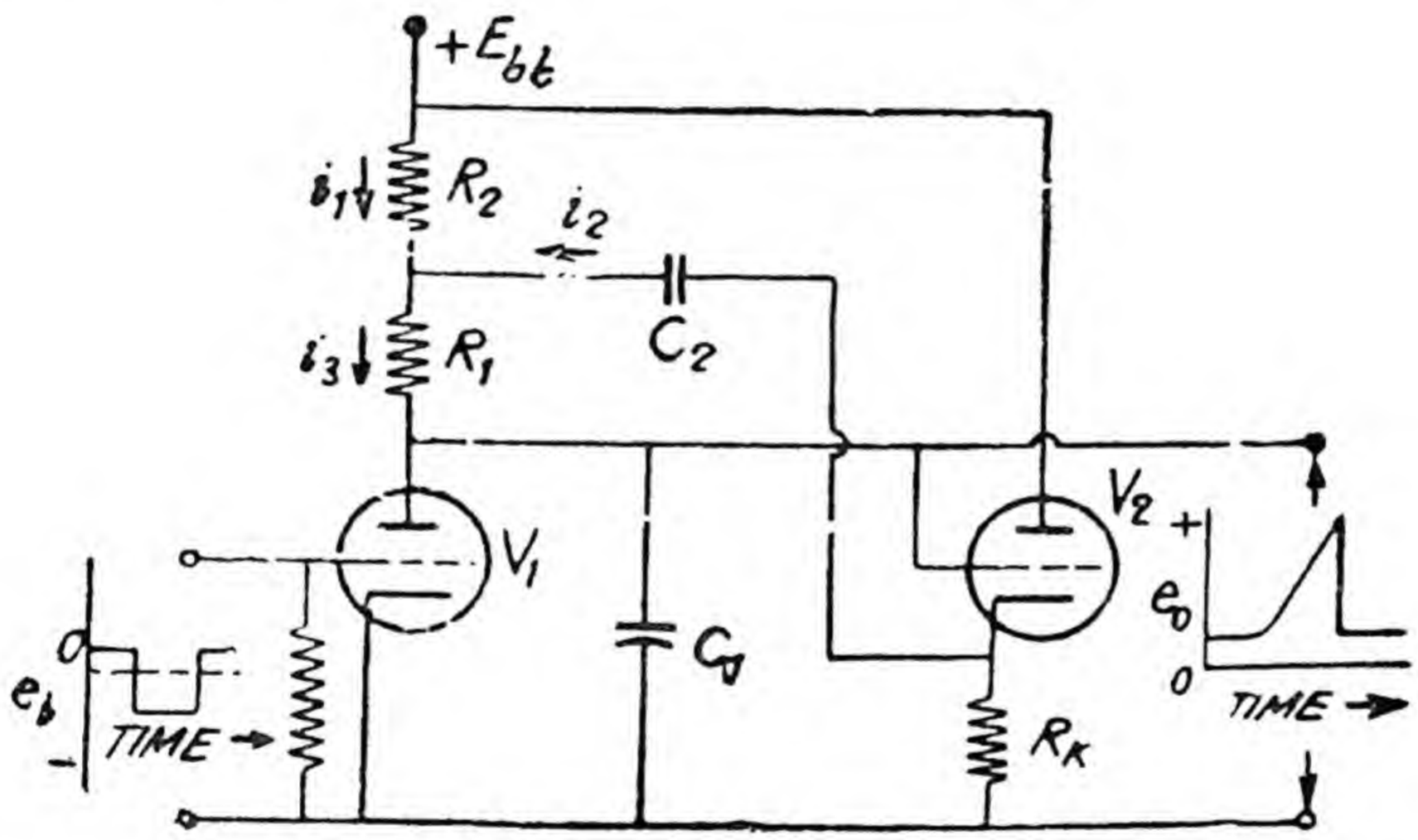


Fig. 10. Boot trap Circuit.

As shown in the circuit output is taken across the condenser C_1. Therefore for the sweep to be linear, it is essential that current, i cha ging this condenser C_1 must be constant.

Tube V_1 is normally conducting due to which its plate voltage is low and therefore the voltage across C_1 is also low. The plate of V_1 is connected to the grid of cathode follower V_2. The plate current through V_2 produces a voltage drop across R_K making the cathode positive with respect to the grid. Feed back is applied from the cathode of V_2 to junction of R_1 and R_2 in the plate circuit of V_1 through the capacitor C_2.

When the tube V_1 is triggered by a negative going square wave it is driven beyond cut off. This rises the plate voltage of V_1 and charges the condenser C_1 through R_1 and R_2. This rise in

voltage across C_1 also rises the grid voltage of V_2 and being a cathode follower, its cathode potential also rises. This rise of cathode potential is applied to the junction of R_1 and R_2 through C_2, *i.e.* voltage at this junction increases as the sweep voltage (across C_1) increases (and hence the name bootstrap).

If the gain of cathode follower is unity, then amplitude of voltage applied to the grid should be equal to the output voltage of cathode follower which is the voltage applied to the junction of R_1 and R_2. *This means the upper end of R_1 should rise in potential almost as rapidly as the lower end* which is possible only when a constant current flows through R_1 *i.e.* i_3 is constant. Since i_3 charges the condenser C_1, sweep voltage is linear.

Basis of bootstrap sweep can now be explained as follows :

Suppose there is an RC circuit with e_i and e_0 as input and output voltages as shown in fig. 11 (*a*). From fig. 11 (*b*) we note that charging current is i_3 playing the part of i of fig. 11 (*a*).

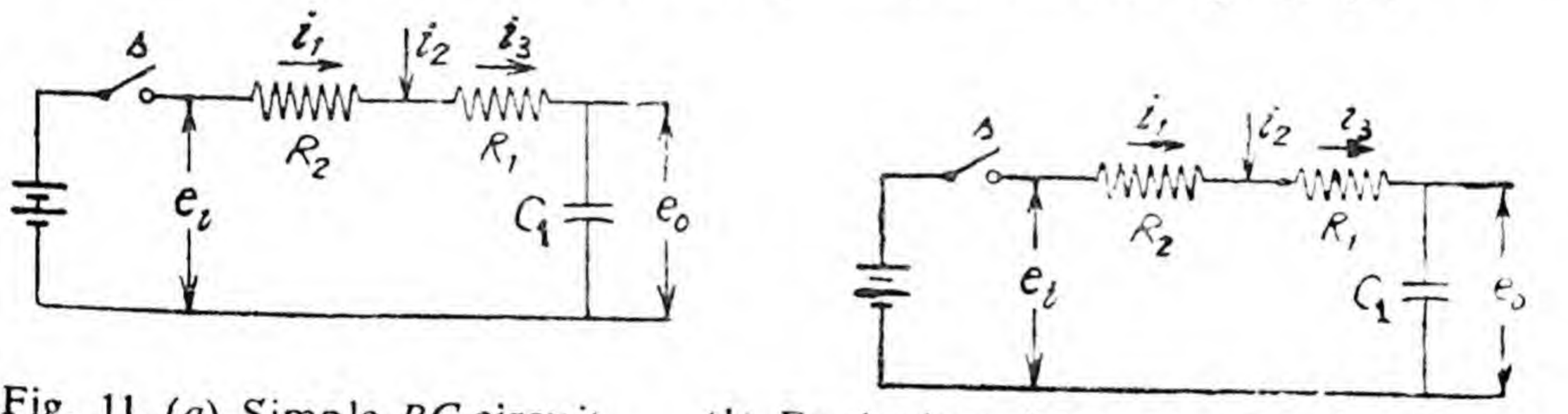


Fig. 11. (*a*) Simple RC circuit. (*b*) Equivalent bootstrap harging circuit.

From fig. 11 (*a*),

$$e_i = iR + \frac{1}{C}\int_0^t i\,dt + V_0,$$

where V_0 is the potential on condenser C at $t=0$ and if E_0 is the charging voltage and I_0 the current at $t=0$, then

$$V_0 = E_0 - I_0R,$$

so that

$$e_i = iR + \frac{1}{C}\int_0^t i\,dt + E_0 - I_0R. \qquad (1)$$

For linear sweep, charging current must be constant so that current should always be equal to its initial value, *i.e.*

$$i = I_0,$$

which when substituted in equation (1) gives

$$e_i = \frac{1}{C}\int_0^t i\,dt + E_0$$

$$= e_0,$$

that is, if the charging is to be constant and sweep linear, it is essential that charging voltage should vary by the same amount as the output voltage. For this to accomplish, we employ a cathode follower because in cathode follower, output voltage follows the input voltage and gain is also approximately equal to unity.

In bootstrap circuit we have noted that output voltage across C_1 is fed back to the junction of R_1 and R_2 and we reason that due to feedback by cathode follower, the output potential is equal to the voltage at the junction leading to the constant charging current condition.

(ii) Miller Integrating Sweep :

Miller integrating sweep circuit is shown in fig. 12.

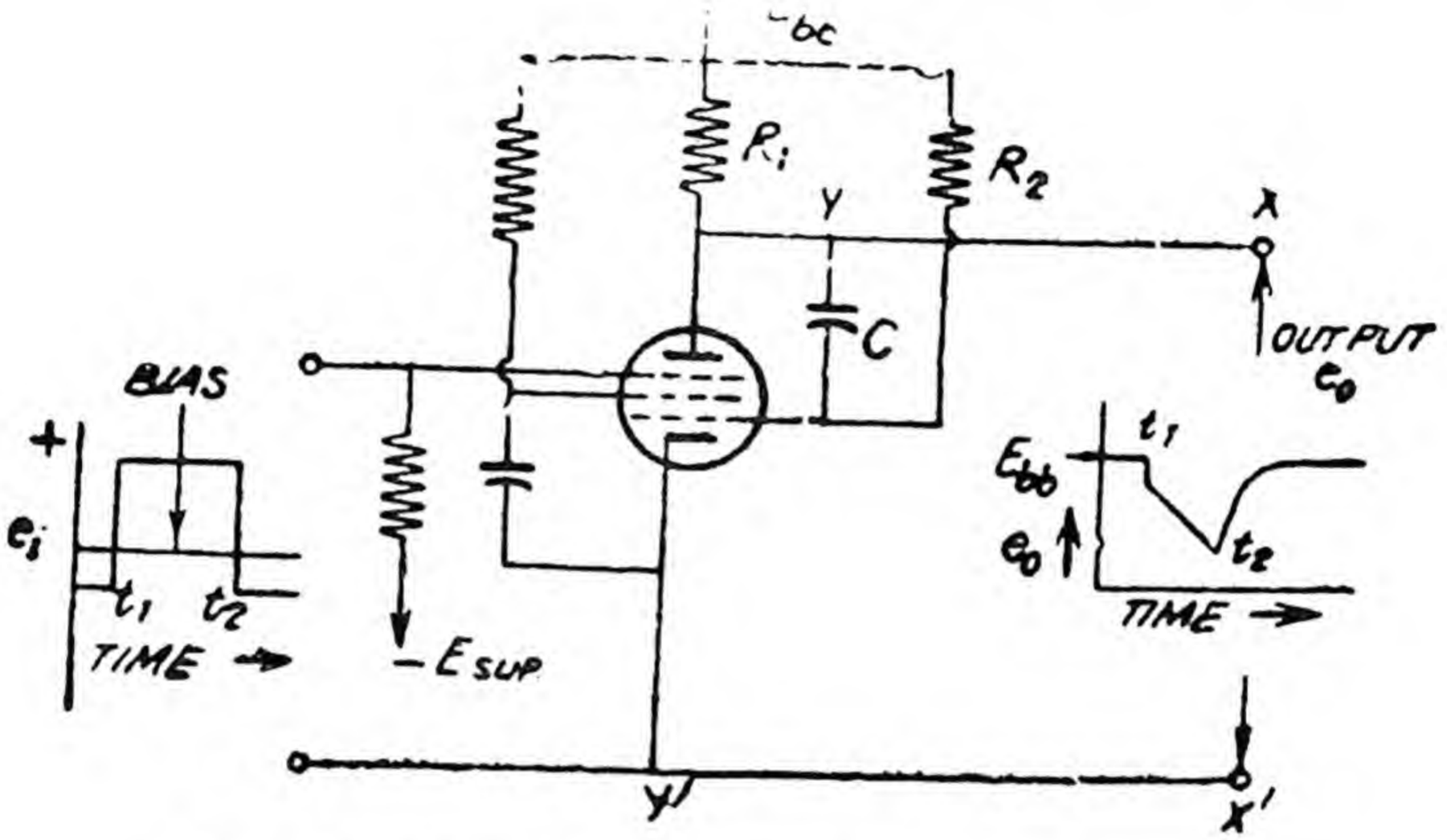


Fig. 22. Miller type sweep generator.

Suppressor grid is biased negatively beyond cutoff so that plate voltage is equal to E_{bb}, and sweep capacitor C is charged to almost this value. As control grid is connected through a high resistance to E_{bb}, it will be slightly positive with respect to cathode. Screen grid is also positive, allowing the flow of screen current and therefore screen potential is also low.

When a positive going square wave is applied to the suppressor the plate current flows, the plate potential starts to fall but it *can not fall rapidly because a rapid fall will allow the voltage across C to drive the grid negative, decresaing the grate current and plate potential to rise*. So only a linear fall of potential, as we shall show mathematically, is resulted. This fall is shown in fig. 12 by a straight line between the instants t_1 and t_2. At t_2, which marks the end of input pulse, suppressor goes negative. When next pulse comes, the cycle of operation is repeated. We note that Miller sweep is also a triggered sweep because a sweep is generated only when some input pulse is applied.

If e_0 be the voltage rise from cathode to grid then grid current will be

$$i_c = \frac{E_{bb} - e_c}{R_2}, \qquad \ldots(1)$$

which is quite small since R_2 is sufficiently high. Further from path $XYY'X'$, we find

$$e_c - e_0 = \frac{1}{C}\int i_c\, dt,$$

$$\frac{de_c}{dt} - \frac{de_0}{dt} = \frac{i_c}{C}$$

$$= \frac{E_{bb} - e_c}{CR_2}, \text{ from equation (1)}$$

or $$pe_c - pe_0 = \frac{E_{bb} - e_c}{CR_2}, \text{ where } p \text{ is } \frac{d}{dt}.$$

Introducing the gain A, where

$$A = \frac{e_c}{e_c}$$

in above equation, we arrive at

$$\frac{e_0}{A} - e_0 = \frac{E_{bb} - e_c/A}{pR_2C}$$

or $$e_0\left[1 - \frac{1}{A} - \frac{1}{ApCR_2}\right] = -\frac{E_{bb}}{pR_2C} \qquad \ldots(2)$$

or $$e_0\left(1 - \frac{1}{A}\right) = \frac{e_0}{ApCR_2} - \frac{AE_{bb}}{ApCR_2}$$

$$e_0 = \left[\frac{e_0}{ApCR_2} - \frac{AE_{bb}}{ApCR_2}\right]\left(1 - \frac{1}{A}\right)^{-1}$$

$$= \frac{e_0}{ApCR_2} - \frac{AE_{bb}}{ApCR_2}$$

neglecting the product with $1/A$ because $1/A^2pCR_2$ will be small when gain is considerably high. Further,

$$pe_0 = \frac{e_0}{ACR_2} - \left(1 - \frac{AE_{bb}}{e_0}\right)$$

$$-\frac{de_0}{e_0\left(1 - \frac{AE_{bb}}{e_0}\right)} = -\frac{dt}{ACR_2},$$

after putting $$p = \frac{d}{dt}$$

On integration,

$$-\log(e_0 - AE_{bb}) = -\frac{t}{ACR_2} + \text{constant}.$$

We know that at $t = 0$, $e_0 = E_{bb}$, giving

$$\text{constant} = -\log(E_{bb} - AE_{bb})$$

so that

$$-\log(e_0 - AE_{bb}) = -\frac{t}{ACR_2} - \log(E_{bb} - AE_{bb})$$

or $$E_{bb}-AE_{bb}=(e_0-AE_{bb})\ e^{-t/ACR_2}$$

or $$e_0e^{-t/ACR_2}=E_{bb}-AE_{bb}\ (1-e^{-t/ACR_2})$$

or $$e_0\left[\ 1-\frac{t}{ACR_2}+\cdots\ \right]=E_{bb}-AE_{bb}\ \left[1-1+\frac{t}{ACR_2}+\cdots\ \right],$$

$$e_0\left(\ 1-\frac{t}{ACR_2}\right)=E_{bb}\left(\ 1-\frac{t}{CR_2}\right),$$

leaving the higher order terms because A is considerably high. Further,

$$e_0=E_{bb}\left(\ 1-\frac{t}{CR_2}\ \right)\left(\ 1-\frac{t}{ACR_2}\right)^{-1}$$

$$=E_{bb}\left(\ 1-\frac{t}{CR_2}\ \right)\left(\ 1+\frac{t}{ACR_2}\right)$$

$$=E_{bb}\left(\ 1-\frac{t}{CR_2}\ \right)+\frac{E_{bb}\ t}{ACR_2},$$

leaving the term containing the product $\frac{t^2}{AC^2R_2^2}$. Therefore

$$e_0=E_{bb}\left[\ 1-\frac{t}{CR_2}\left(\ 1-\frac{1}{A}\right)\right]$$

Since gain A is sufficiently high we can neglect $\frac{1}{A}$ in comparison to 1, so that

$$e_0=E_{bb}\left(\ 1-\frac{t}{CR_2}\right),$$

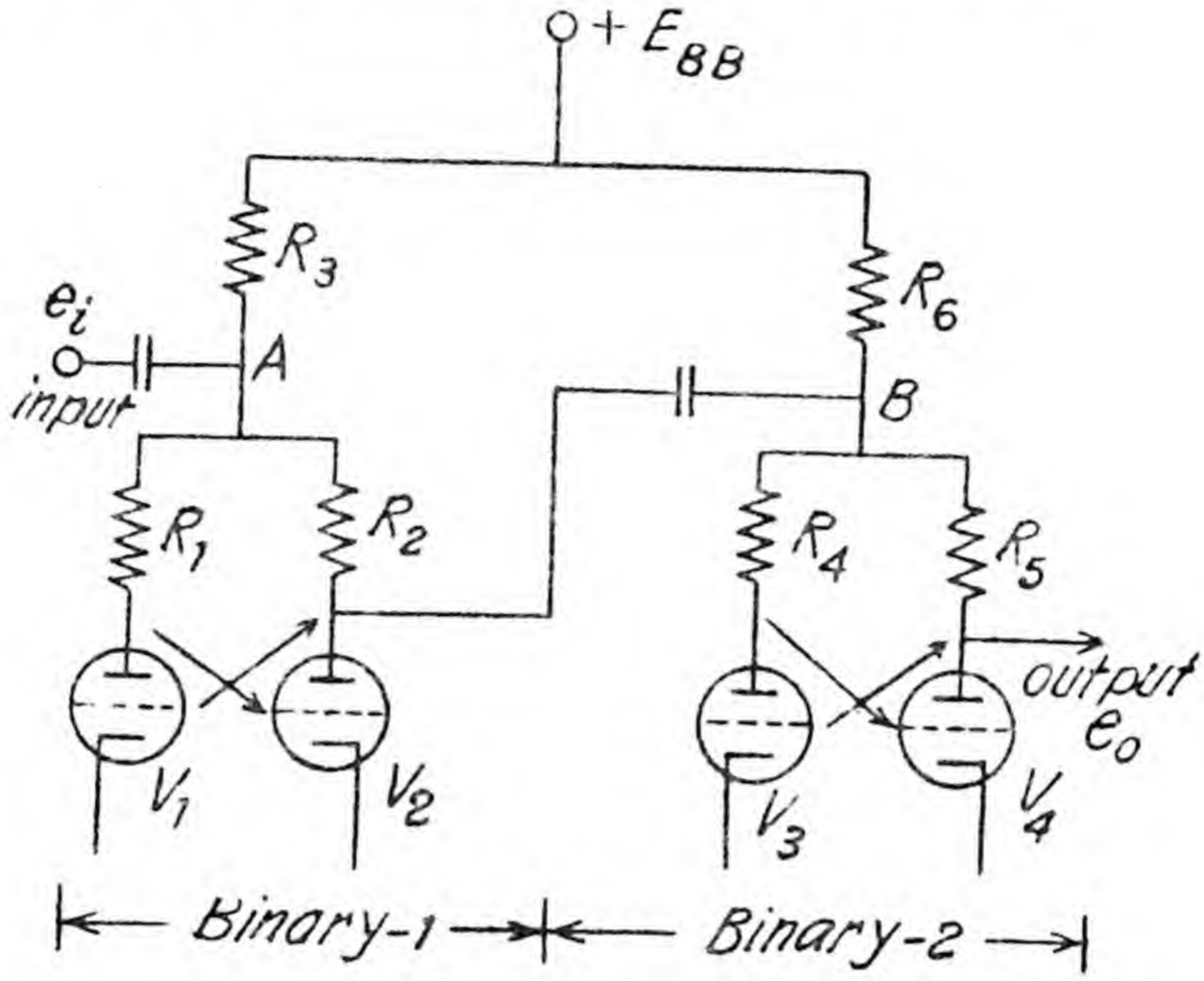
which shows that potential drop is linear from t_1 to t_2.

20·1-1. SCALING CIRCUITS : COUNTING TO THE BASE 2 :

In art. 20·0-3 (bistable multivibrator) we have explained the action of binary as a scale of two counter. It was pointed out there that if two such binaries were arranged in cascade, they would form a scale of four counter. We shall discuss here the latter counter.

Two binaries in cascade : Scale of four counter :

Ref. to fig. 13. Let us designate the two stable states of a binary by the symbol 0 (*zero*) for 'off' state and by the symbol 1 (*one*) for 'on' state. Further, when we say that a binary is in *one* state, it would mean that input stage is on (conducts) *e.g.* tubes that conduct are V_1 and V_3, while output stage is off (non-conducting) *e.g.* tubes V_2 and V_4 are not conducting. Similarly, binary is in *zero* state would mean that input stage is off while output stage is on *e.g.* V_1 off and V_2 conducting.



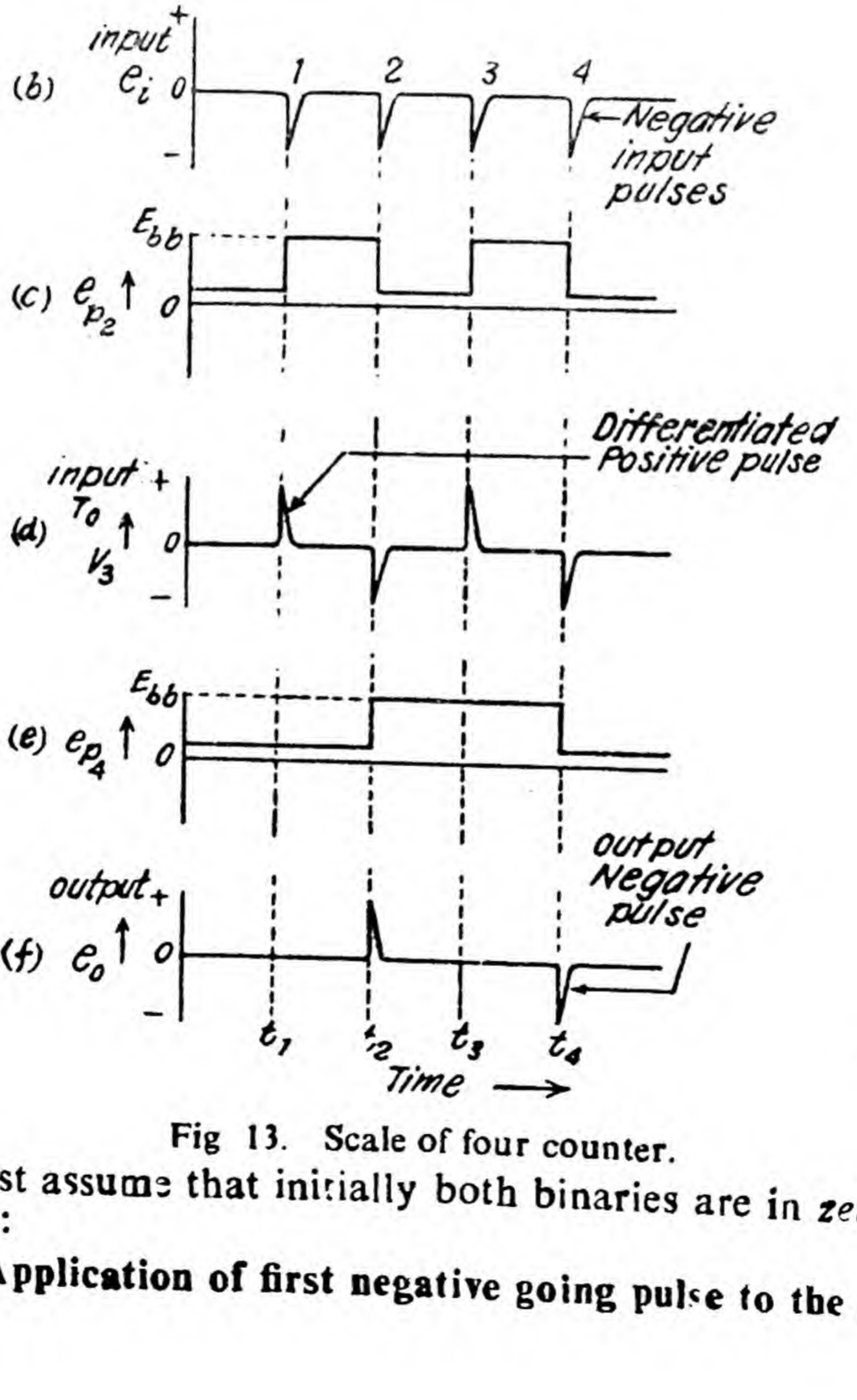


Fig 13. Scale of four counter.

We first assume that initially both binaries are in *zero* state.

Operation :

(i) Application of first negative going pulse to the point A of binary 1.

Binary 1 is in *zero* state, *i.e.* V_1 off and V_2 conducting so that V_1 offers high resistance while V_2 (as it is conducting) a low resistance. Further, branch having resistance R_1 and high resistance offered by V_1 is in parallel to the branch having resistance R_2 and low resistance offered by V_2; therefore portion of any applied voltage across any combination will be

$$\frac{\text{pulse inp ut voltage}}{\text{total resistance of both branches}} \times \text{branch resistance}$$

It means when negative going pulse is applied to the point *A*, most of its portion will appear across R_1, *i.e.*, at the plate of V_1. This negative voltage is therefore applied to the grid of V_2 and switching action starts, so that finaliy V_2 is off and V_1 on. Binary 1 is thus switched to *one* state.

With V_2 off, its plate voltage rises as shown by e_{p2} in fig 13 (*c*). This rise is applied to the point *B* of the plates of birary 2 through condenser C_1 and input resistance of binary 2, R_{in}. Since the time constant $R_{in}\ C_1$ is small, $R_{in}\ C_1$ acts as a differentiator and therefore a positive pulse is produced for the rise in e_{p2} for the period between t_1 and t_2.

As stated previously, most of the portion of this positive pulse will appear at the plate of V_3 since V_4 is conducting offering low resistance. This means grid of V_4 will become more positive but the tube V_4 is already heavily conducting, its plate current is saturated and application of any positive voltage to its grid will not affect its operation. Binary 2 thus remains in previous *zero* state.

(ii) Application of second negative going pulse at the point A.

At instant t_2, when second negative going pulse is applied, the tube V_2 will be again *on* while tube V_1 cutoff. This means binary 1 is again in *zero* state. The plate potential of V_2 will fall as shown in fig. 13 (*c*). This fall in voltage will produce a negative pulse due to the low time constant $R_{in}\ C_1$. Thus negative pulse is applied to the point *B*.

Most of the portion of negative pulse, applied to the point *B*, appears at the plate of V_3 and hence to the grid of V_4. Thus finally V_4 is cutoff while V_3 conducts. The plate potential of V_4 rises and a positive pulse will appear at the output. Binary 2 is now in *one* state.

Thus upto instant t_2, binary 1 has changed its state twice while binary 2 only once. We note that in the output of binary 2, a positive pulse, and not a negative pulse, is produced by the application of two negative pulses at the point *A* of binary 1.

(iii) Application of third negative going pulse at the point A :

At this instant t_3, V_2 will be cut off V_1 is on. *i e*. binary 1 is in *one* state. Plate potential of V_2 will rise, applying a positive pulse to the point *B*. Now since V_3 is conducting, V_4 nonconduct-

ing, most of this pulse will appear at the plate of V_4 consequently, grid potential of V_3 will rise : but as the tube V_3 is already conducting to its plate current saturation, application of this positive voltage to its grid will not affect its performance, *i.e.*, binary 2 does not change its state and is still in *one* state.

(iv) Application of fourth negative going pulse at the point A :

V_1 off, V_2 will conduct, *i.e* , binary 1 is in *zero* state. Plate potential of V_2 falls imparting negative pulse at the point *B* The grid of V_3 will become negative, its plate current will decrease, the plate potential rises. Finally V_3 is cut off and V_4 is conducting. Plate potential of V_4 decreases, producing a negative pulse at the output. Binary 2 thus comes back in *zero* state.

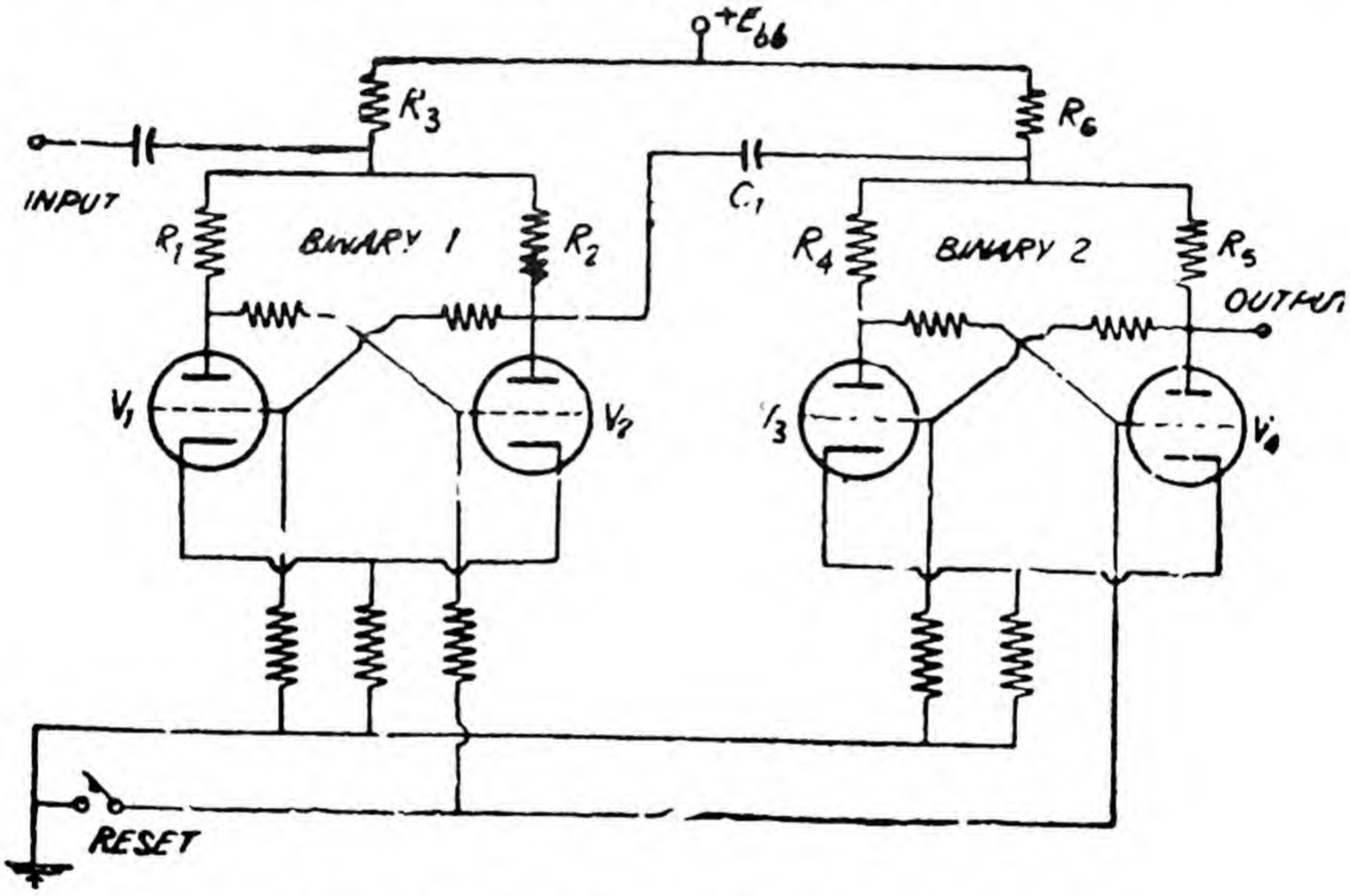


Fig. 14. Circuit of scale of four counter .

Thus *one* negative output pulse is produced by the application of *four* input negative going pulses, one after the other, to the point *A* of binary 1. Thus the circuit counts by *four*--a scale of four counter. Since for 16 inputs it will produce only 4 output pulses, the circuit can be thought of as dividing by four. Complete circuit is shown in fig. 14. The above operation scheme is given in Table—1.

TABLE-1. Operation Scheme of Scale Four Counter.

Number of input pulse	State of binary 1	State of binary 2	Tubes that conduct	Number of negative output pulse
Initial state	0	0	V_2, V_4	0
1st pulse	1	0	V_1, V_4	0
2nd pulse	0	1	V_2, V_3	0
3rd pulse	1	1	V_1, V_3	0
4th pul e	0	0	V_2, V_4	1

To establish the initial condition prevailing in a binary chain a *reset* switch is provided as shown in the figure 14. When this switch is closed, all grid leak resistors are connected to ground Then switch is open, bias is removed from the output stage of each binary and the tubes V_2, V_4 conduct establishing the initial *zero* state.

Thus it is obvious that output of each binary represents a power of 2, the output of first binary 2^1, and output of second binary 2^2 and so on. Therefore if we employ four such binaries, a scale of 2^4 ($= 16$) counter will be obtained. Refer to figure 15 (ignoring the feedback line *FF*) and following the operation of scale of four counter, we find that on the application of 16 pulses to the binary-1. 4 will be recorded by the binary-2, 2 will be recorded by the binary 3 and only 1 will be recorded by the binary 4. Scheme of operation is shown in Table-2 (page 881). The circuit produces one output pulse for sixteen input pulses applied to the binary-1.

20·2-2. DECADE COUNTER : COUNTING TO BASE 10.

A decade counter produces a single output pulse for every ten input pulses.

Suppose we start with every binary in *one* state *i.e.* input tubes conducting (V_1, V_3, V_5, V_7) and output tubes cutoff (V_2, V_4, V_6, V_8). This is possible because suppose after the reset (points *a* and *b* are connected for reset and then disconnected) V_1 conducts more than V_2 and similarly input tube of every binary conducts more so that *one* state is the initial state. The application of pulses in sequences affect the circuit operation in the following way :

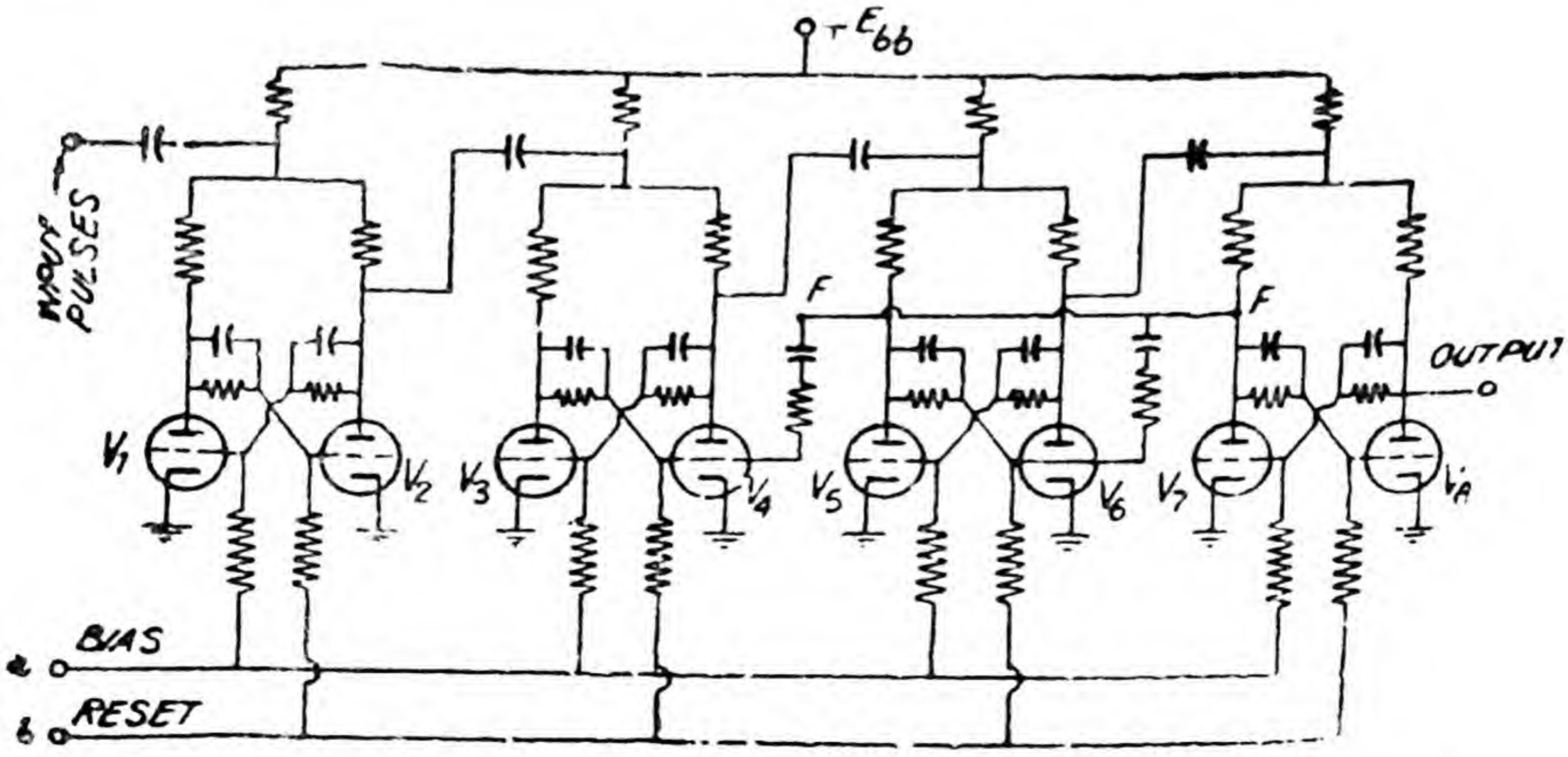


Fig. 15. Circuit of scale of sixteen counter converted into a decade counter by the simultaneous feedback from plate of V_7 to grids of V_6 and V_4. Feedback line is *FF*.

TABLE 2

FOR NORMAL FUNCTION OF SCALE OF SIXTEEN COUNTER

	NO FEED BACK												Effect of feed back from V_7 to V_8 and V_4
Pulse number	V_1	V_2	State of Binary-1	V_3	V_4	State of Binary-2	V_5	V_6	State of Binary-3	V_7	V_8	State of Binary-4	
Initial 0	ON	OFF	1	ON	OFF	1	ON	OFF	1	ON	OFF	1	
1	OFF	ON	0	〃	〃	1	〃	〃	1	〃	〃	1	
2	ON	OFF	1	OFF	ON	0	〃	〃	1	〃	〃	1	
3	OFF	ON	0	〃	〃	0	〃	〃	1	〃	〃	1	
4	ON	OFF	1	ON	OFF	1	OFF	ON	0	〃	〃	1	
5	OFF	ON	0	〃	〃	1	〃	〃	0	〃	〃	1	
6	ON	OFF	1	OFF	ON	0	〃	〃	0	〃	〃	1	
7	OFF	ON	0	〃	〃	0	〃	〃	0	〃	〃	1	
8	ON	OFF	1	ON	OFF	1	ON	OFF	1	OFF	ON	0	→ 8
9	OFF	ON	0	〃	〃	1	〃	〃	1	〃	〃	0	→ 9
10	ON	OFF	1	OFF	ON	0	〃	〃	1	〃	〃	0	→ 10
11	OFF	ON	0	〃	〃	0	〃	〃	1	〃	〃	0	
12	ON	OFF	1	ON	OFF	1	OFF	ON	0	〃	〃	0	
13	OFF	ON	0	〃	〃	1	〃	〃	0	〃	〃	0	
14	ON	OFF	1	OFF	ON	0	〃	〃	0	〃	〃	0	→ 14
15	OFF	ON	0	〃	〃	0	〃	〃	0	〃	〃	0	→ 15
16	ON	OFF	1	ON	OFF	1	ON	OFF	1	ON	OFF	1	→ 16 → Initial state

Due to feed back
8th pulse is equivalent to 14th pulse
9th pulse is equivalent to 15th pulse
10th pulse is equivalent to 16th pulse.

When 8th pulse is applied, binary 4 is in *zero* state so that V_8 conducts and V_7 is cutoff. The rise in plate voltage of V_7 is fed-back to the grids of tubes V_4 and V_6. This makes the grids of V_4 and V_6 positive and thus they conduct bringing the binaries 2 and 3 in *zero* state, *i.e.*, though binary 2 and binary 3 were in *one* state at the 8th pulse without feedback but now due to the feedback, they are bought to *zero* state which is the state with 14th pulse, if there is no feedback as is obvious from the Table-2. Thus 8th pulse (with feedback) is equivalent to 14th pulse (without feedback), 9th pulse will be equivalent to 15th pulse and 10th pulse will bring the final and the initial state of 16th pulse.

Thus with feedback, 10th pulse brings the same operating state as brought by the application of 16th pulse without feedback and therefore an output pulse will be produced when the 10th pulse is applied to the binary-1. The circuit then acts as a scale of ten or decade counter.

EXERCISES

1. What are monostable and bistable circuits ? Explain the function for which they are used. Draw a diagram with approximate value of components of a scalar for a scale of two.
2. Draw the circuit and waveforms of a blocking oscillator. How can it be used as a frequency divider.
3. What are disadvantages of a neon and thyratron sawtooth generator ?
4. Draw the circuit diagram of a vacuum tube sweep generator and explain its function. What modification will you bring to linearize the sweep ?
5. What are disadvantages associated with the bootstrap sweep generator? Draw the circuit and waveforms of Miller linear sweep generator.
6. Explain the function and draw the circuit diagram of a scale of sixteen counter employing binary. How can it be modified to work as a decade counter ?

21 WAVE SHAPING CIRCUITS

The process of deriving a desired waveform from a given waveform is termed as wave shaping.

In this chapter we shall describe two types of wave shaping circuits :

(A) Linear Wave Shaping : It has been observed that when a sinusoidal wave is input to *RLC* network (passive elements), then output waveform is not changed in shape. Such *RLC* networks, as they do not distort sine wave, are called linear networks. But if a *non sinusoidal* wave is transmitted through such networks, then shape of output waveform is different from the shape of the input wave. Such a shaping of input waveform by linear networks is called linear wave shaping.

(B) Non-linear Wave Shaping ; When shaping of the input wave to a circuit such as clipping, clamping is done by taking advantage of the non-linearity of semi-conductor or thermionic devices, then it is termed as a non-linear wave shaping.

(A) LINEAR WAVE SHAPING

As stated earlier in linear wave shaping, the wave is passed through linear network and the operations involved are linear operations such as integration, differentiation and summation etc. We shall deal here the sinusoidal input and shaping of non-sinusoidal, *viz.* step, pulse and square wave inputs through *RC* network.

21·1. THE HIGH PASS RC CIRCUIT :

Fig. 1 shows a high pass *RC* circuit. Since the reactance of capacitor, *C*, decreases with increasing frequency, the capacitor acts almost as a short circuit at high frequencies. Thus all the high frequency input appears at the output Thus the circuit acts as high pass filter. We shall now discuss the response of the circuit to various types of input.

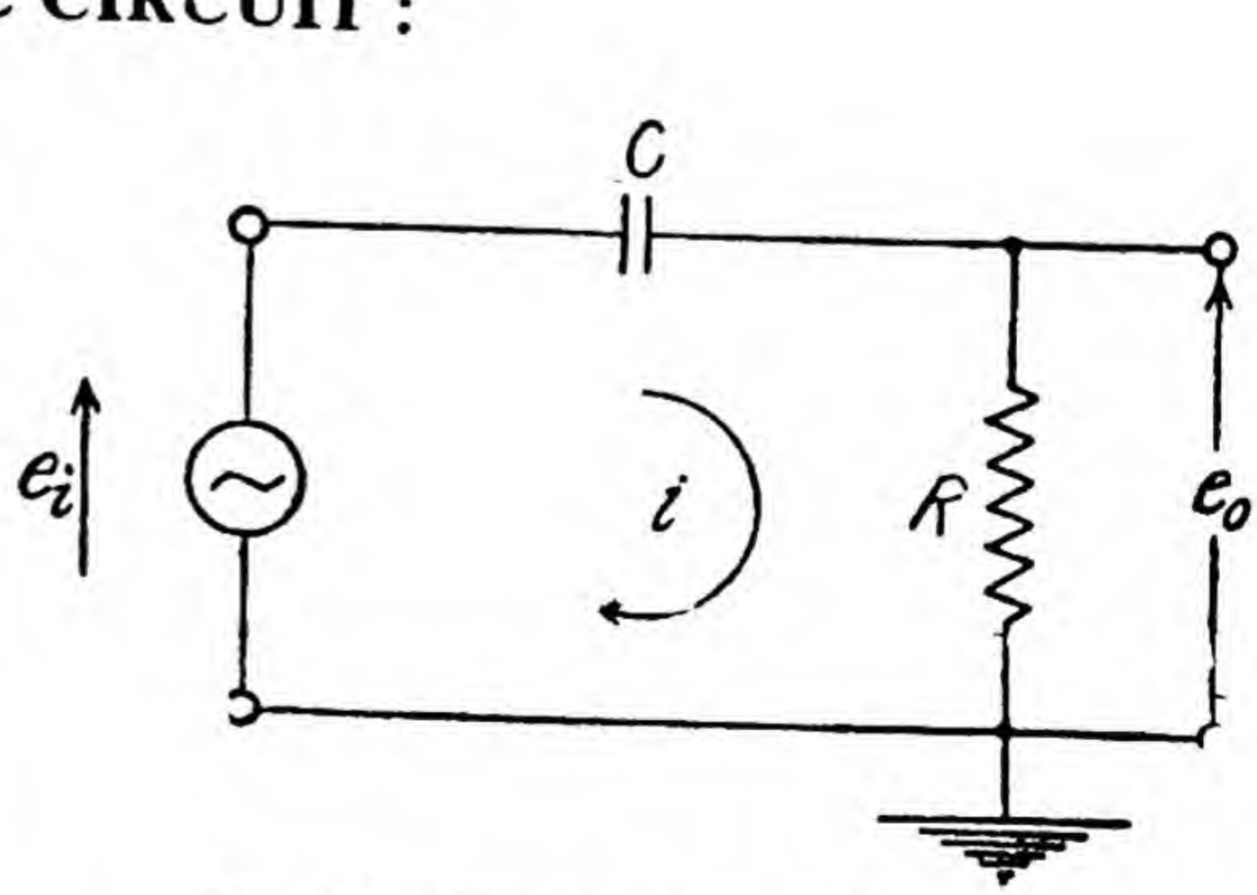


Fig. 1. High pass *RC* circuit.

(1) Sinusoidal Input : Let the input be

$$e_i = E_{im} \sin \omega t,$$

then the current, i, flowing through the circuit will be

$$i = \frac{e_i}{R + \frac{1}{j\omega C}}$$

and therefore output across R will be

$$e_0 = iR = \frac{e_i R}{R + \frac{1}{j\omega C}} = \frac{E_{im} R \sin \omega t}{R + \frac{1}{j\omega C}} \quad \ldots(1)$$

or

$$e_0 = E_{0m} \sin \omega t, \quad \ldots(2)$$

where

$$E_{0m} = | E_{0m} | e^{j\theta} = \frac{E_{im} R}{R + \frac{1}{j\omega C}}$$

or

$$| E_{0m} | \cos \theta + j | E_{0m} | \sin \theta = \frac{E_{im}}{1 - \frac{j}{\omega CR}}$$

or

$$= \frac{E_{im}\left(1 + \frac{j}{\omega CR}\right)}{1 + \frac{1}{\omega^2 C^2 R^2}}$$

Equating real and imaginary parts, we get

$$| E_{0m} | \cos \theta = \frac{E_{im}}{1 + \frac{1}{\omega^2 C^2 R^2}}$$

$$| E_{0m} | \sin \theta = \frac{E_{im}/\omega CR}{1 + \frac{1}{\omega^2 C^2 R^2}},$$

so that

$$| E_{0m} | = \frac{E_{im}}{\left[1 + \frac{1}{\omega^2 C^2 R^2}\right]^{1/2}}$$

and

$$\tan \theta = \frac{1}{\omega CR}$$

Writting again equation (2) for output voltage, we find

$$e_0 = E_{0m} \sin \omega t$$
$$= | E_{0m} | e^{j\theta} \sin \omega t$$
$$= \frac{E_{im}}{\left(1 + \frac{1}{\omega^2 C^2 R^2}\right)^{1/2}} \sin (\omega t + \theta).$$

Thus output voltage leads the input voltage by an angle θ.

Further ratio of peak values of output and input voltage is given by

$$\frac{E_{0m}}{E_{im}}=\frac{R}{R+\frac{1}{j\omega C}}=\frac{1}{1-\frac{j}{\omega CR}}.$$

Let us put

$$\omega_1=\frac{1}{CR},$$

then

$$\frac{E_{0m}}{E_{im}}=\frac{1}{1-\frac{j\omega_1}{\omega}}$$

$$=\frac{1}{\left\{1+\frac{\omega_1^2}{\omega^2}\right\}^{1/2}}\angle\theta$$

where $$\theta=\tan^{-1}\left(\frac{\omega_1}{\omega}\right)$$

We can also write

$$\frac{E_{0m}}{E_{im}}=\frac{1}{\left[1+\left(\frac{f_1}{f}\right)^2\right)^{1/2}}\angle\ \tan^{-1}\left(f_1/f\right),$$

where $f_1=\frac{1}{2\pi CR}$ is called lower cut off frequency at which this ratio falls to 70·7% of its maximum value. At this frequency, the magnitude of capacitive reactance is equal to the resistance. The maximum possible value of the gain (unity) is approached asymptotically at high frequencies (fig. 2).

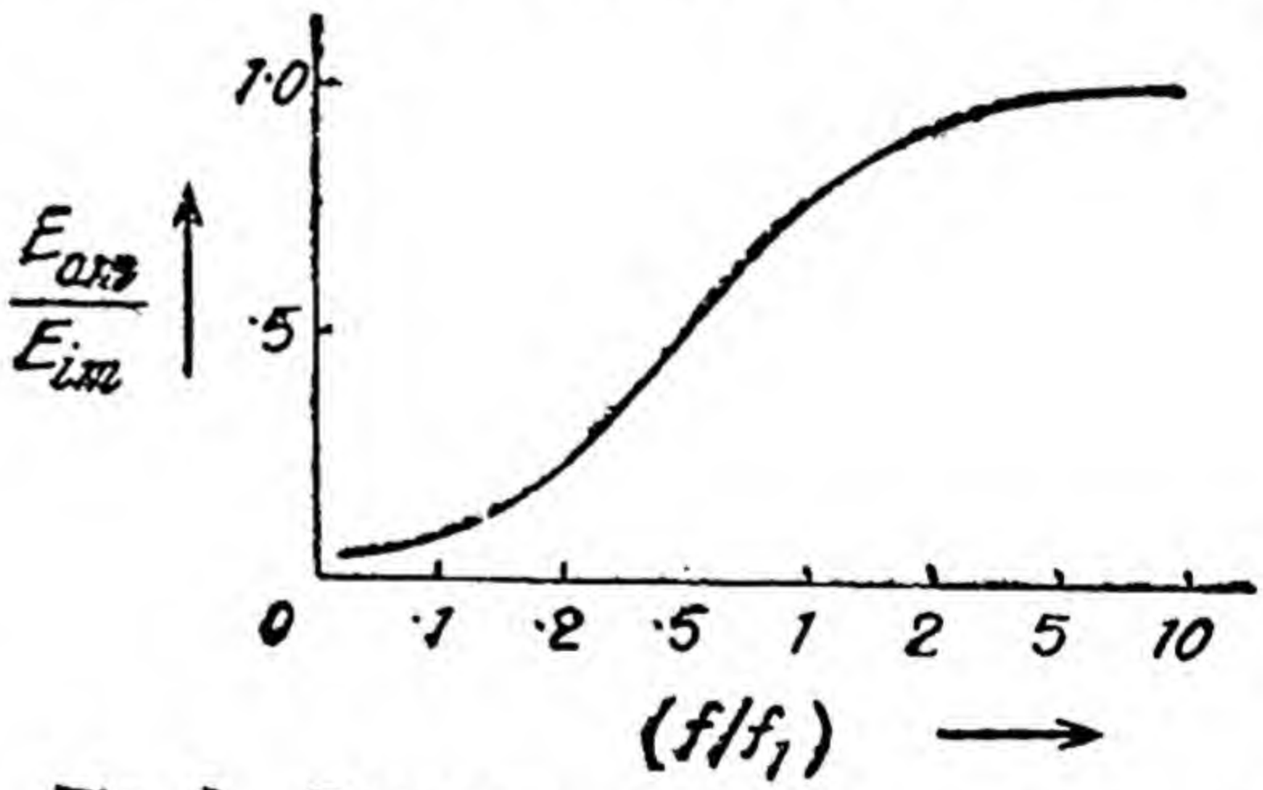


Fig. 2. Response of high pass *RC* circuit for step-voltage.

(2) Step-voltage Input : A step voltage (fig. 3) is one which maintains the value zero for all times $t<0$ and maintains the value E_0 for all times after $t=0$. Rise of voltage from zero to E_0 occurs very rapidly at $t=0$ so that voltage curve is vertical at $t=0$. Thus for step voltage, we can write

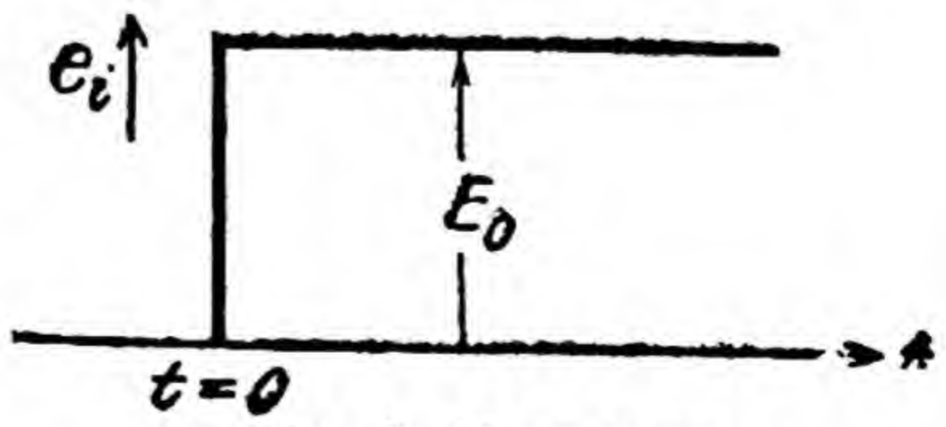


Fig. 3. Step-voltage.

$e_i=0$ at $t=0_-$ (immediately before $t=0$),
$e_i=E_0$ at $t=0_+$ (immediately after $t=0$).

Such a voltage is applied at the input of a high pass *RC* circuit (fig. 1)

We know that the response of a high pass *RC* circuit to step input voltage is exponential with a time constant *RC*, we can write output voltage in the form

$$e_0(t) = B_1 + B_2 e^{-t/RC}.$$

In order to find constant B_1 and B_2, we have the boundary conditions

$$\left.\begin{aligned} e_0(t) &= 0 \text{ when } t = \infty \\ e_0(t) &= E_0 \text{ when } t = 0 \end{aligned}\right\}$$

which when applied to above equation yield

$$B_1 = 0$$

$$B_2 = E_0$$

Therefore output voltage will be of the form

$$e_0 = 0 + E_0\, e^{-t/RC}$$
$$= E_0\, e^{-t/RC}$$

For $t = RC$,

$$e_0 = e^{-1}\, E_0 = 0{\cdot}63\, E_0,$$

i.e., output falls to 0·63 of maximum input voltage. Fig 4 shows the input and output voltages. The drop from the signal value is called the *sag*. The sag value after t second is

$$\text{Sag.} = E_0 - e_0(t)$$
$$= E_0 - E_0\, e^{-t/RC}$$

and fraction sag

$$= \frac{\text{Sag.}}{E_0}$$
$$= (1 - e^{-t/RC})$$
$$= \left[1 - \left(1 - \frac{t}{RC} + \frac{t^2}{2R^2C^2} - \frac{t^3}{6R^3C^3} + \dots\right)\right]$$
$$= \frac{t}{RC},$$

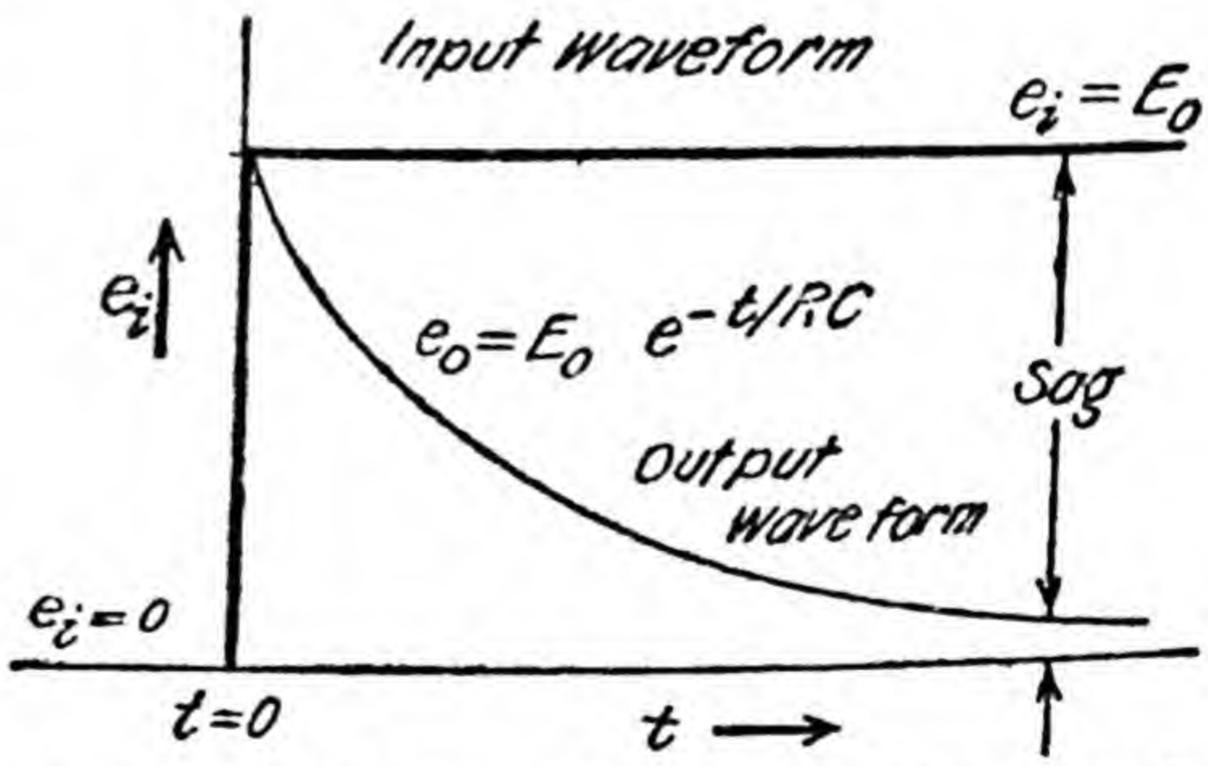


Fig 4 Response of high pass *RC* circuit for step-voltage.

when t is far less than *RC*. In fig. 5, output voltage is plotted for *RC* high enough and *RC* very low. We observe that in case *RC* is made high, e_0 resembles e_i except for a small sag.

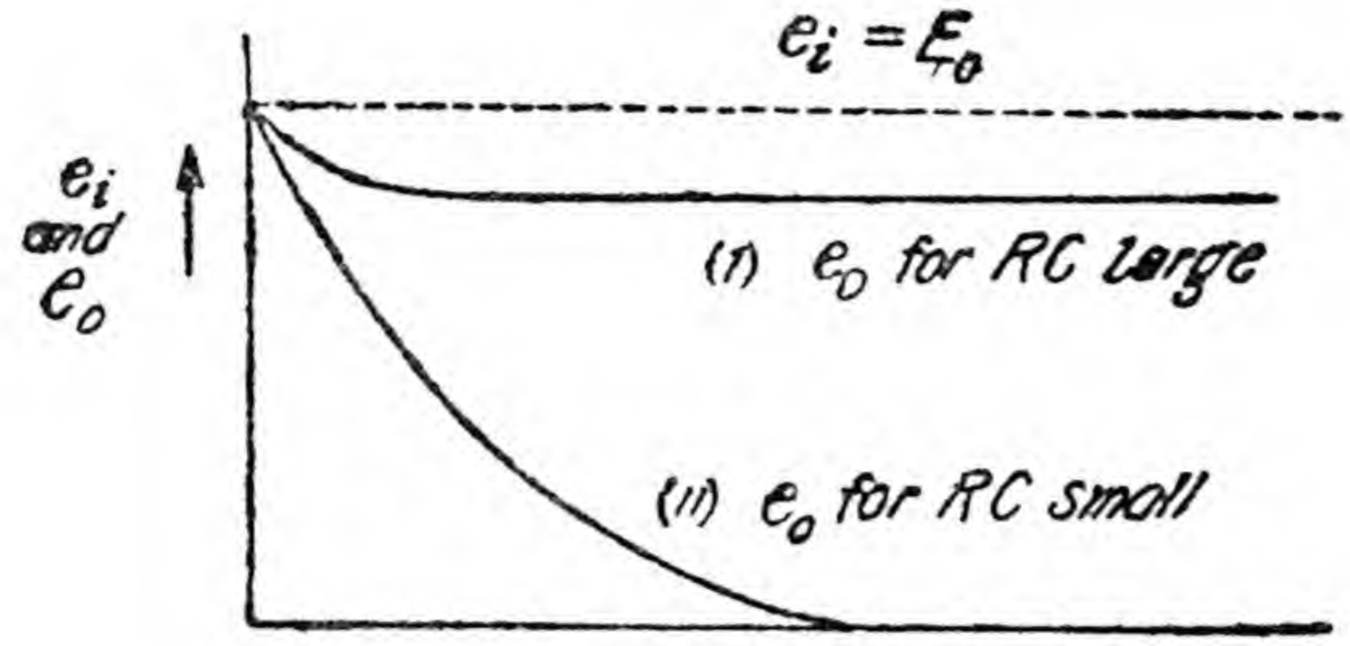


Fig. 5. Plot of output voltage for
(*i*) *RC* high (*ii*) *RC* low.

(3) Pulse Input :

Fig. 6 (*a*) shows an ideal pulse with amplitude E_0 and duration t_p. It appears from figs. 6 (*a*), (*b*) and (*c*) that the pulse may be considered to be the sum of two step voltages; first step voltage E_0 whose discontinuity occurs at $t=0$ and other step voltages $-E_0$ whose discontinuity occurs at $t=t_p$..

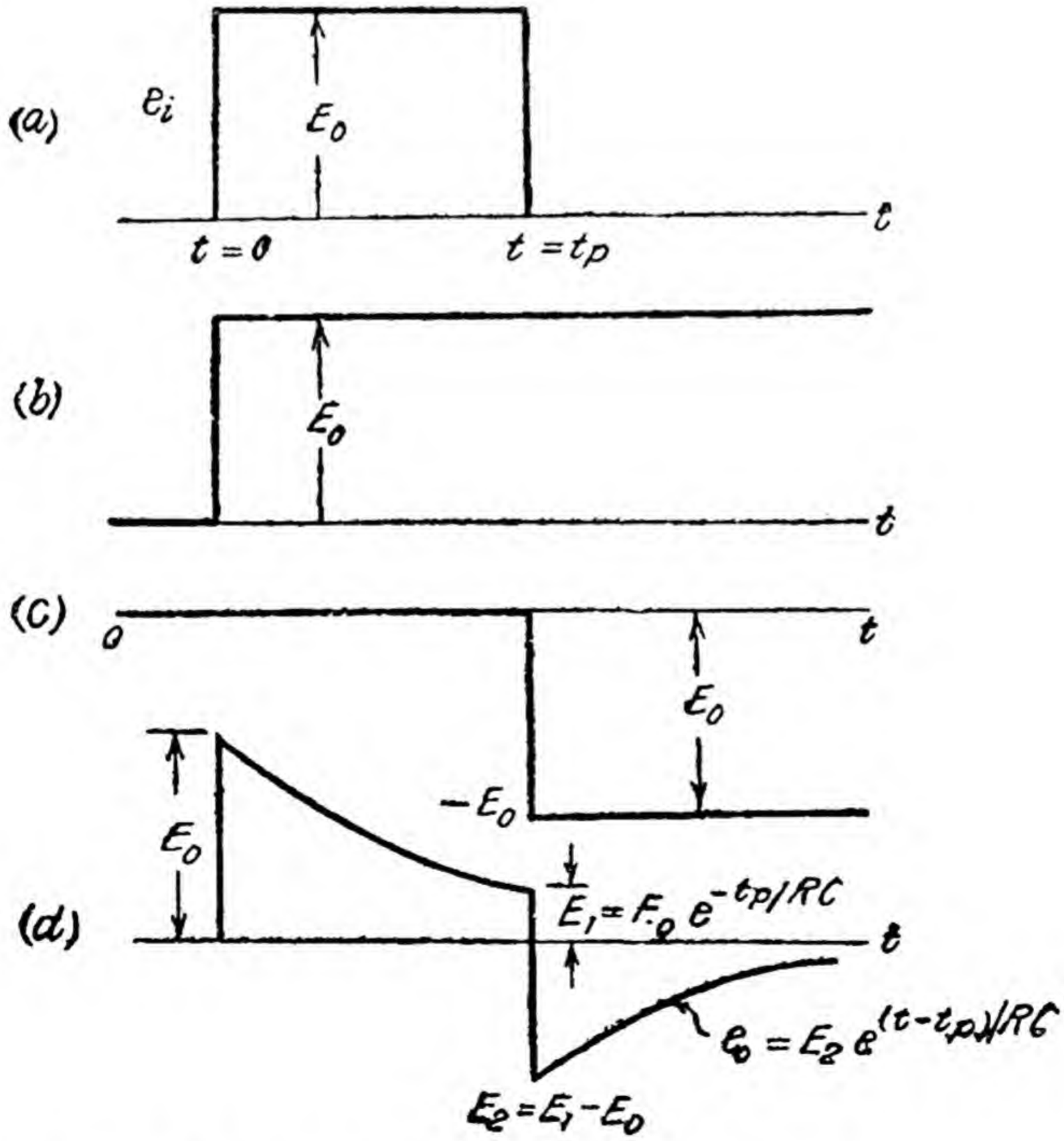


Fig. 6. (*a*) An ideal square pulse of width, t_p,
(*b*), (*c*) Step voltages which make up the pulse,
d) pulse after transmission through high pass *RC* circuit.

When the pulse of fig. 6 (*a*) is applied at the input of fig. (1), the output voltage for times less than t_p *i e.* in the interval $t=0$ to t_p will be the same as that of a step voltage input. Therefore for $0 < t < t_p$, output is given by

$$e_0 = E_0\, e^{-t/RC} \qquad 0 < t < t_p$$

At the end of the pulse, $t=t_p^-$, the height of the output pulse be E_1 so that we can write

$$E_1 = E_0\, e^{-t_p/RC}$$

At the end of the pulse, $t=t_p^+$, the input falls abruptly by an amount E_0, and, since the voltage across the capacitor can not change instantaneously, the output voltage must also drop by the same amount E_0. Therefore at $t=t_p^+$ the output voltage is given by

$$\begin{aligned} E_2 &= E_1 - E_0 \\ &= E_0\, e^{-t_p/RC} - E_0 \\ &= E_0\,(e^{-t_p/RC} - 1) \end{aligned}$$

Since E_1 is less than E_0, the voltage E_2 becomes negative and after $t=t_p$, the output voltage decays exponentially (with time constant CR) towards zero as shown in fig. 6 (*d*).

For $t=t_p$, output voltage is given by

$$\begin{aligned} e_0 &= E_2\, e^{-(t-t_p)/RC} \\ &= E_0\,(e^{-t_p/RC} - 1)\, e^{-(t-t_p)/RC} \end{aligned}$$

We observe from fig. 6 (*d*) that the transmitted pulse is distorted – there is a tilt to the top of the pulse and an under-shoot at the end of the pulse. If these distortions are to be minimised, then the time constant RC must be very large compared with the width t_p. If the time constant is very large ($RC/t_p \gg 1$), there is a slight tilt to the output pulse and the undershoot is very small (fig. 7 *b*). If the time constant is very small ($RC/t_p \ll 1$) the output consists of a positive *spike* or *pip* of amplitude E_0 at the beginning of the pulse and a negative spike of the same size at the end of the pulse (fig 7*c*). This process of converting pulses into pips by means of a circuit of short time constant is called *peaking* or *differentiating*.

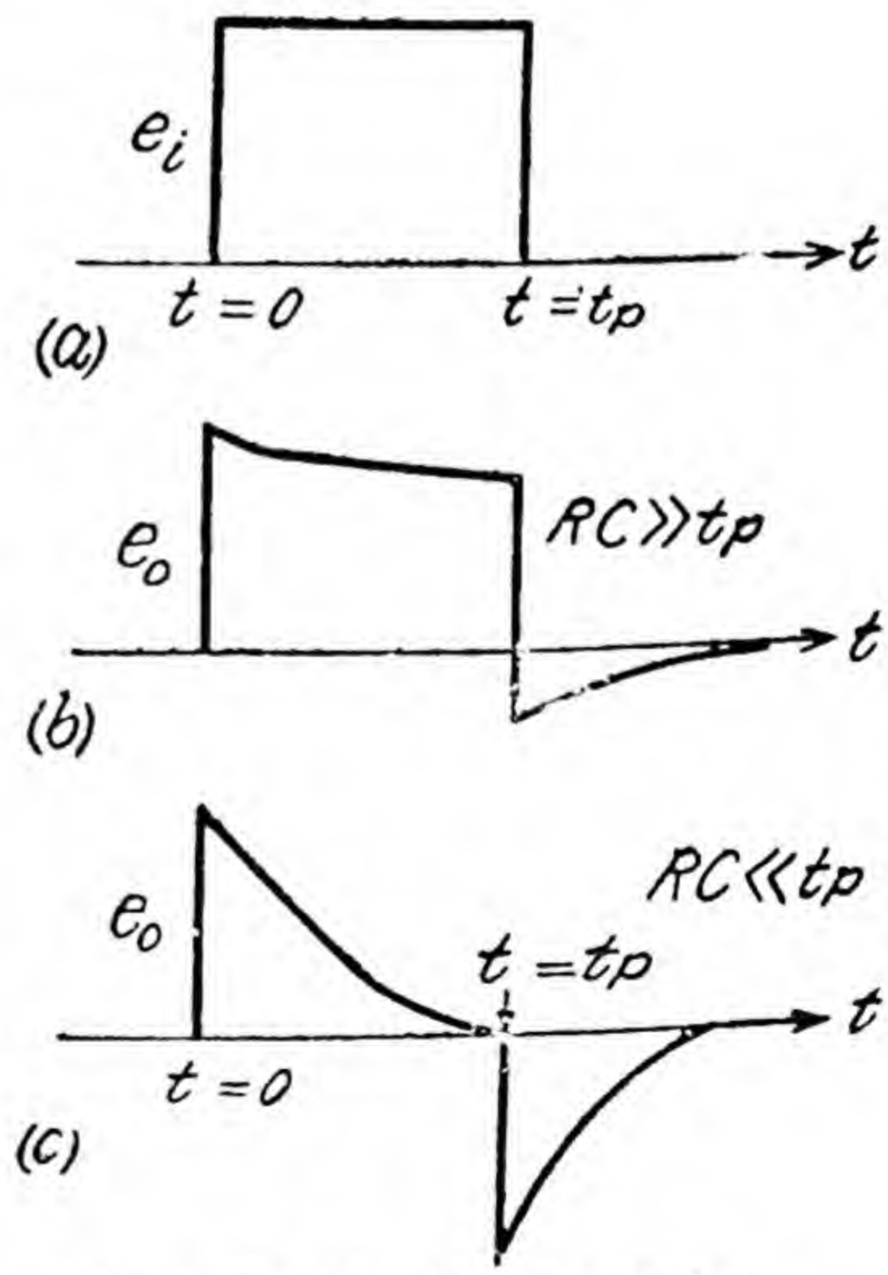


Fig 7 (a) Input pulse
(b) Output pulse with $RC \gg t_p$.
(b) Output pulse with $RC \ll t_p$.

(4) Square Wave Input :

A waveform which maintains itself at one constant level for a time T_1 and

at another constant level for a time T_2 and which is repetitive with a period (T_1+T_2), T_1 being equal to T_2, is called a square wave (fig. 8 *a*).

Let a square wave having a period comparable to the RC value of the circuit is applied at the input of fig. 1. We shall now study the response of the circuit (fig. 8*b*) at different levels of input wave :

(*i*) First, soon as the positive spike voltage appears at the input, it directly appears across the load R as output, because initially condenser C behaves as short circuit.

(*ii*) Now the capacitor starts charging through R, and thus capacitor voltage increases. As a result, positive voltage across R goes on decreasing exponentially with time constant RC of the circuit towards zero. But this exponential decay is slow as the time constant RC is comparable to half period, T_1, of the wave and as such, at $t=t_1=T_1$, the output voltage can not approach a value zero.

(*iii*) Now at $t=t_1$ a negative spike voltage equal to E_0 appears at the input, due to which output voltage atonce falls by the same amount. Thus output voltage becomes negative. Now as the capacitor C, discharges exponentially with time constant RC, output voltage across R increases exponentially towards zero. But since time constant RC is comparable to T_2, the output voltage does not rise to a value zero at $t=t_2=T_1+T_2$.

(*iv*) At $t=t_2$, again positive spike appears across input so that output rises by an amount E_0 and it becomes positive. State of affair is now like points (i) and (ii) described above. The tip values

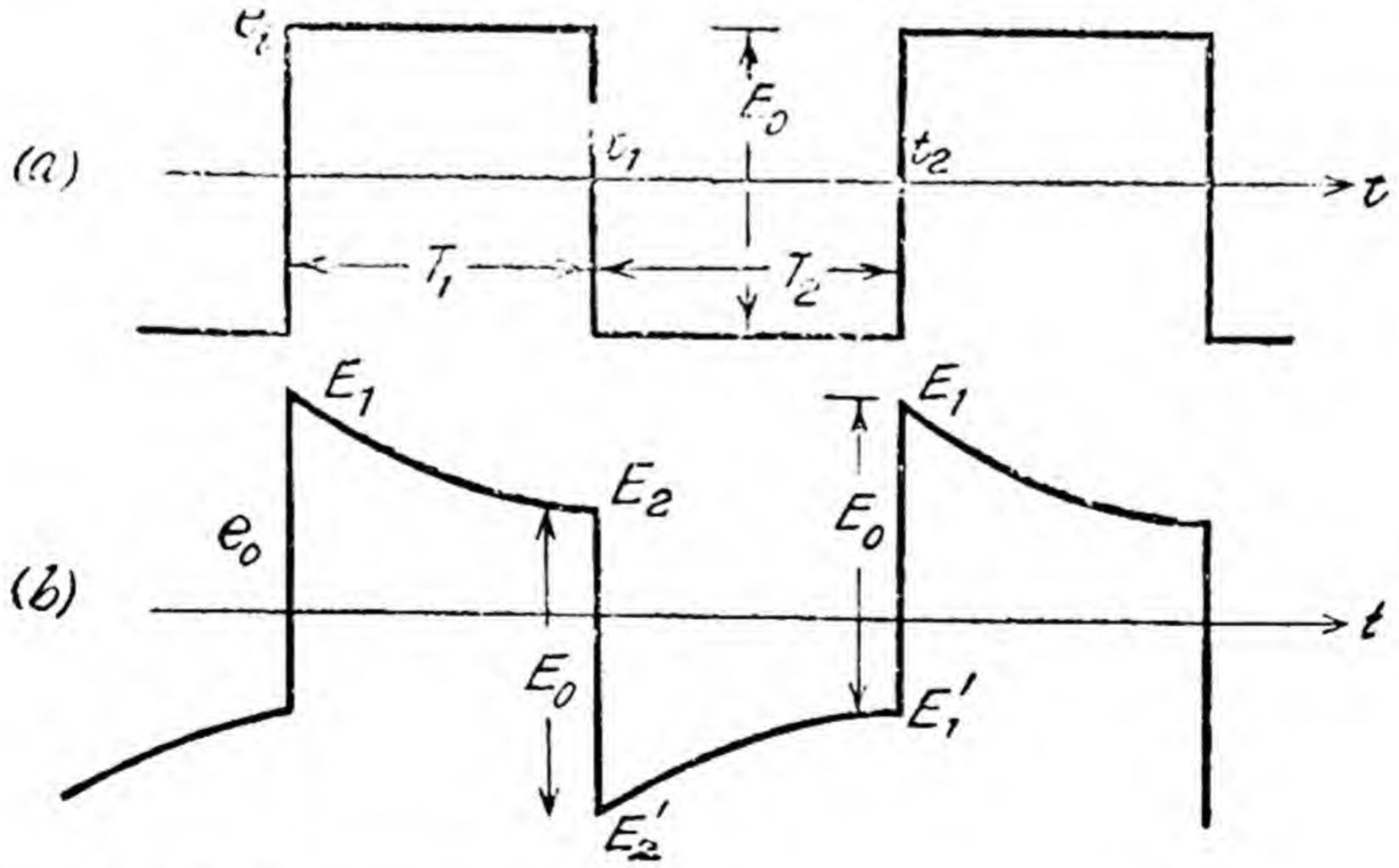


Fig. 8. (*a*) Square wave input.
(*b*) Output wave (response of the circuit).

E_1, E_2, E_1', and E_2' of output voltage as shown in fig. 8 (*b*), can be calculated as follows :

(*i*) The output voltage from E_1 to E_2 is given by

$$e_0=E_1\, e^{-t/RC}$$

when $t = T_1$

$e_0 = E_2$

so that $E_2 = E_1 \, e^{-T_1/RC}$ ...(1)

(*ii*) The equation of the curve from E_2' to E_1' is

$$e_0 = E_2' \, e^{-(t-T_1)/RC}$$

when $t = t_2 = T_1 + T_2$

$e_0 = E_1'$

so that $E_1' = E_2' \, e^{-T_2/RC}$ ...(2)

(*iii*) At instant $t = t_1$, E_2 drops by an amount E_0 so that the output voltage is E_2'. Therefore

$$E_2 - E_0 = E_2'. \quad ...(3)$$

(*iv*) At instant $t = t_2$, the output voltage E_1' abruptly increases by an amount E_0 to attain a value equal to E_1. Therefore

$$E_1' + E_0 = E_1 \quad ...(4)$$

Eqs. (1), (2), (3) and (4) on solving, provide four tip values E_1, E_2, E_1' and E_2' of the output voltage.

21·2. HIGH PASS RC CIRCUIT AS DIFFERENTIATOR :

In fig. 9, a high pass *RC* circuit is shown. As arleady mentioned, the process of converting pulses into pips by means of a circuit of short time constant is called differentiating. Thus here time constant *RC* is far less than the half time (T_1 or T_2) period of the wave.

If e_i is input voltage, e_0 output voltage, and i current flowing in the combination, then with zero initial voltage across *C*, we can write

$$e_i = \frac{q}{C} + iR,$$

where q is the charge on *C*,

or $$e_i = \frac{1}{C}\int idt + iR$$

If time constant *RC* is small compared to the interval of applied voltage, then condenser will charge and discharge rapidly which means that out of two terms in equation for e_i, the first term will predominate and hence

$$e_i \approx \frac{1}{C}\int idt$$

or $$i = C\,\frac{de_i}{dt}$$

Therefore output across *R* will be

$$e_0 = iR$$

$$= RC\frac{de_i}{dt}$$

Because *RC* is small and constant, we have

$$e_0 \propto \frac{de_i}{dt}.$$

Obviously, output voltage is proportional to the derivative of the input voltage. Thus the wave is said to be differentiated since it develops another wave, the shape of which is determined by the derivative of the input voltage.

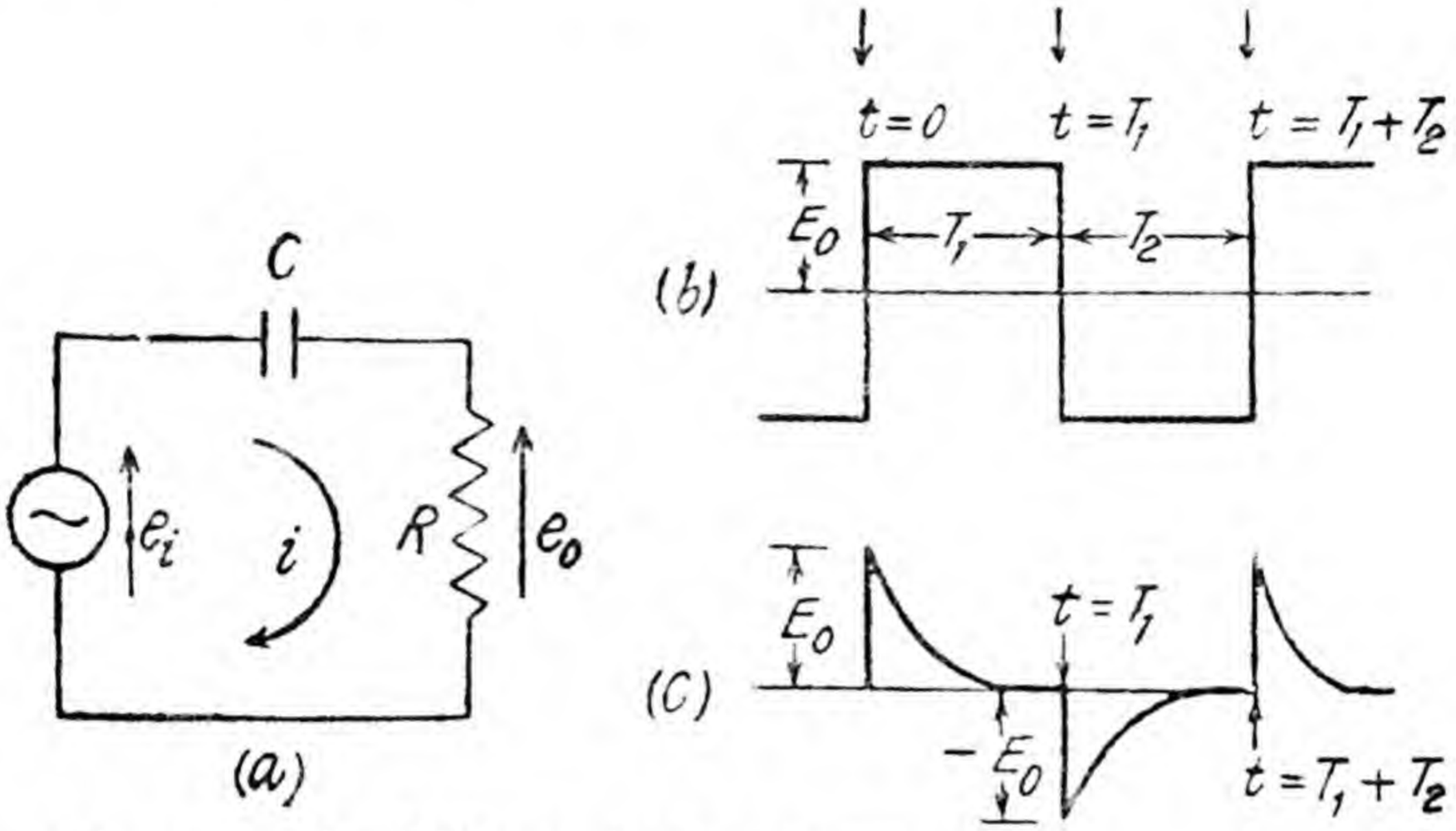


Fig. 2. (*a*) Differentiating circuit,
(*b*) Pulse input.
(c) Differentiated output.

We can study the response of the circuit as follows :

(*i*) At $t=0$, when step voltage is applied to the circuit, the whole input voltage appears across load R as the capacitor initially acts as short circuit.

(*ii*) At $t > 0$, capacitor charges exponentially and voltage across R, therefore, decreases exponentially with time constant CR. Since $CR \ll T_1$, charging of C and therefore decrease of output voltage across R takes place very rapidly so that in a short time output voltage is reduced to zero and is maintained at zero till the instant $t=T_1$.

(*iii*) At $t=T_1$, negative step voltage ($-E_0$) appears at input so that atonce output goes negative by an amount equal to the amplitude of the input step voltage ($=-E_0$). This voltage then increases to zero as the capacitor discharges rapidly before time $t=T_1+T_2$ is reached.

(*iv*) At $t=T_1+T_2$, a positive step voltage is input to the circuit and state of affairs is like that described in points (i) and (ii) above. We can obtain a criterion for good differentiation in terms of the time period of a sinusoidal input voltage. Refer to art. 21·1, case of sinusoidal input. The output voltage is

$$e_0=|\, E_{0m} \,| \sin(\omega t+\theta),$$

where output voltage leads the input voltage by θ, given by

$$\tan\theta=\frac{1}{\omega CR}.$$

We note that for better differentiation CR should be as small as possible. It can be very near to zero or in other words tan θ should approach infinity for which $\theta=90°$. But the angle can not be exactly 90°, because for this either $R=0$ or $C=0$ which is practically impossible. Thus $\theta=89{\cdot}4°$ serves as a better selected value for good differentiation. Thus

$$\tan 89{\cdot}4=\frac{1}{\omega CR}$$

$$100=\frac{1}{\omega CR}$$

or $$\omega CR={\cdot}01$$

or $$2\pi fCR={\cdot}01$$

or $$\frac{2\pi CR}{T}={\cdot}01,$$

where f and T are the frequency and the time period of the inp t sine wave. Thus

$$RC=\frac{{\cdot}01T}{2\pi}$$

$$\leqslant 0{\cdot}0016T$$

which is a criterion for good differentiation.

21·3. LOW PASS RC CIRCUIT :

Fig. 10 shows a low pass RC circuit. Since the reactance of the capacitor, C, decreases with increasing frequency, this circuit passes low frequencies readily but attenuates high frequencies. Thus the circuit acts as a low pass filter. We shall now discuss the response of the circuit to various types of the input.

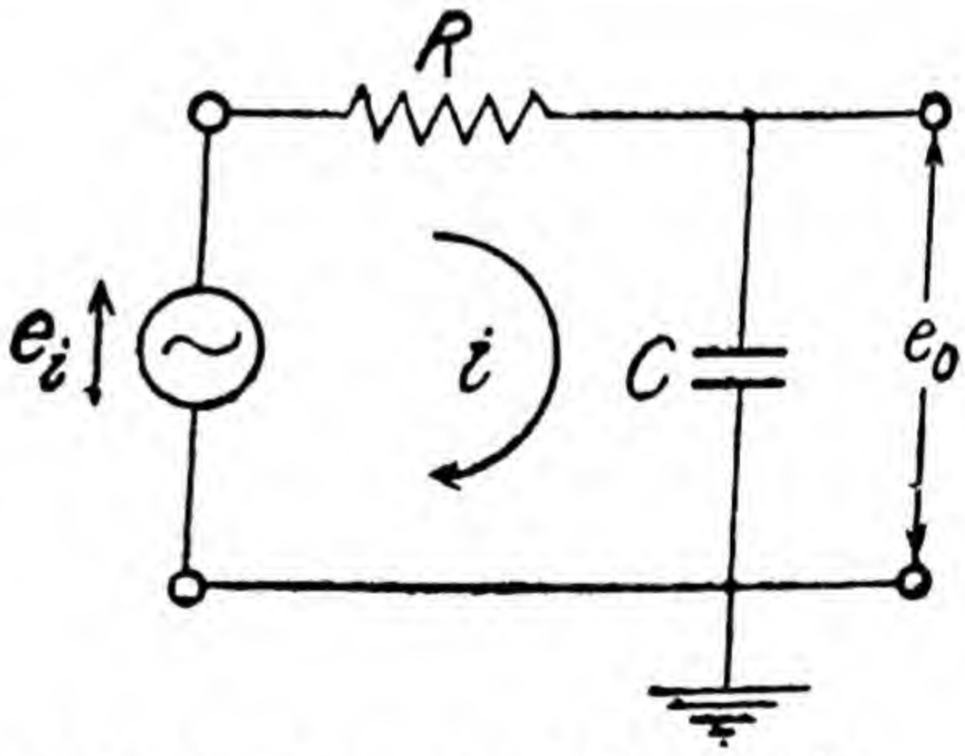


Fig 10. Low pass RC circuit.

(1) Sinusoidal Input :

Let the input be

$$e_i=E_{lm}\sin\omega t,$$

then the current, i, flowing through the circuit will be

$$i=\frac{e_i}{R+\frac{1}{j\omega C}}$$

and therefore output across C will be

$$e_0=i.\frac{1}{j\omega C}=\frac{e_i}{R+\frac{1}{j\omega C}}.\frac{1}{j\omega C}$$

$$= \frac{E_{im} \sin \omega t}{1 + j\ \omega CR} \qquad \ldots(1)$$

or $$e_0 = E_{0m} \sin \omega t, \qquad \ldots(2)$$

where $$E_{0m} = | E_{0m} | e^{j\theta} = \frac{E_{im}}{1 + j\omega C R}$$

or $$| E_{0m} | \cos \theta + j | E_{0m} | \sin \theta = \frac{E_{im}\ (1 - j\omega CR)}{1 + \omega^2 C^2 R^2}$$

or $$| E_{0m} | \cos \theta = \frac{E_{im}}{1 + \omega^2 C^2 R^2}$$

$$| E_{0m} | \sin \theta = \frac{-\omega CR\ E_{im}}{1 + \omega^2 C^2 R^2}$$

$$| E_{0m} | \cos (-\theta) = \frac{E_{im}}{1 + \omega^2 C^2 R^2}$$

$$| E_{0m} | \sin (-\theta) = \frac{\omega CR\ E_{im}}{1 + \omega^2 C^2 R^2}$$

so that $$| E_{0m} | = \frac{E_{im}}{[1 + \omega^2 C^2 R^2]^{1/2}}$$

and $$\tan (-\theta) = \omega CR$$

or $$\theta = -\tan^{-1} (\omega CR).$$

Thus θ should be taken as $-\theta$ in equation for E_{0m}, so that

$$E_{0m} = | E_{0m} | e^{-j\theta}$$

and output voltage will be

$$e_0 = E_{0m} \sin \omega t$$
$$= | E_{0m} | e^{-j\theta} \sin \omega t$$
$$= | E_{0m} | \sin (\omega t - \theta).$$

Therefore output voltage lags behinds the input voltage by an angle θ.

Further ratio of peak value of output and input voltage is given by

$$\frac{E_{0m}}{E_{im}} = \frac{1}{1 + j\omega CR} = \frac{1}{1 + \frac{j\omega}{\omega_2}}$$

by taking

$$\omega_2 = \frac{1}{CR}.$$

Thus $$\frac{E_{0m}}{E_{im}} = \frac{1}{\left(1 + \frac{\omega^2}{\omega_2^2}\right)^{1/2}} \angle (-\theta)$$

where $$\theta = \tan^{-1} \left(\frac{\omega}{\omega_2}\right).$$

We can also write

$$\frac{E_{0m}}{E_{im}} = \frac{1}{[1+(f/f_2)^2]^{1/2}} \angle \tan^{-1}(f/f_2),$$

where $f_2 = \frac{1}{2\pi CR}$ is called the upper cut off frequency at which this ratio falls to 70·7% of its maximum value. Variation of (E_{0m}/E_{im}) as a function of (f/f_2) is shown in fig. 11. We observe that gain at high frequencies is dropped to lowest value whereas it is unity at low frequencies.

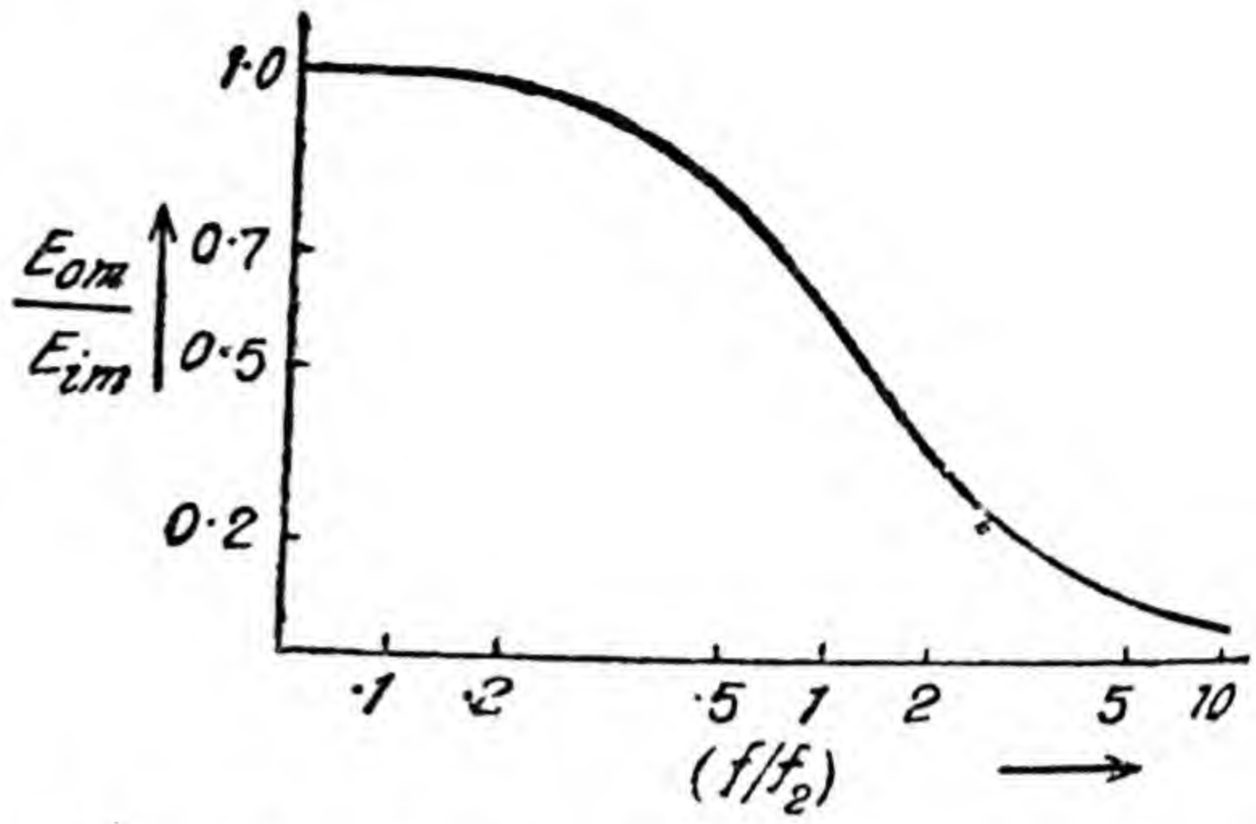


Fig. 11. Graph showing gain as a function of frequency.

(2) Step voltage Input :

We have already explained in art. 21·1 the meaning of step-voltage. We know that the response of low pass *RC* circuit to step input voltage is exponential with a time constant *RC*. That is, as charge across the capacitor cannot change instantaneously, the output starts from zero and rises exponentially with time constant *RC* towards steady state value as shown in figure 12.

The output voltage is the same as the voltage across the condenser during charging and is given by

$$e_0 = E_0\,[1 - e^{-t/RC}].$$

Time, t_1, in which output falls to 10% of its final value, E_0, will be

$$e_0 = 0{\cdot}1\ E_0 = E_0\,[1 - e^{-t_1/RC}]$$

or $$e^{-t_1/RC} = 0{\cdot}9$$

or $$-\frac{t_1}{RC} = \log 0{\cdot}9$$

or $$\frac{t_1}{RC} = 2{\cdot}303 \log_{10}(1/0{\cdot}9)$$

$$= 2{\cdot}303\,(\log 10 - \log 9)$$
$$= 2{\cdot}303 \times {\cdot}045$$
$$= 0{\cdot}1$$

or $$t_1 = 0{\cdot}1\ RC.$$

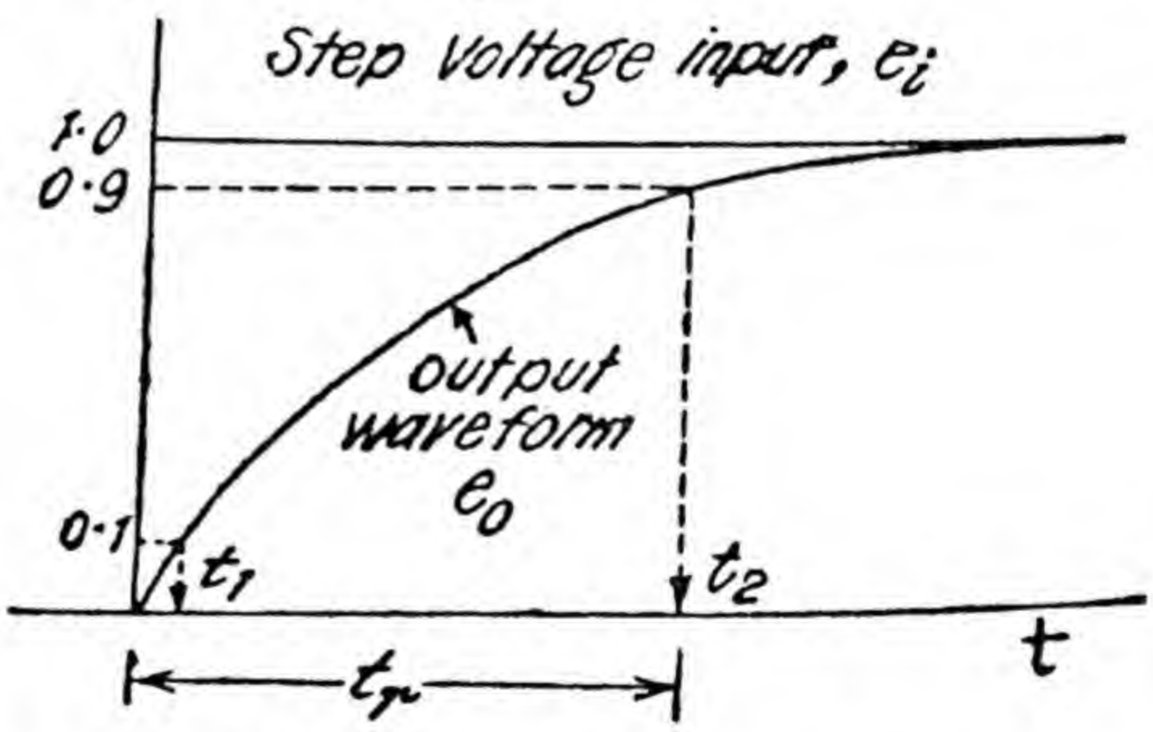


Fig. 12. Response of low pass *RC* circuit for step voltage.

And the time, t_2, in which the output falls to 90% of its final value E_0, is given by

$$e_0 = 0{\cdot}1\ E_0 = E_0\,(1 - e^{-t_2/RC})$$

or $$e^{-t_2/RC}=1-0\cdot9$$

or $$\frac{t_2}{RC}=2\cdot303\ (\log 10-\log 1)$$

$$=2\cdot3$$

or $$t_2=2\cdot3\ RC.$$

The difference between t_1 and t_2 is called *rise time*, t_r, of the circuit. It gives an indication of how fast the circuit can respond to an abrupt voltage. Thus it is given by

$$t_r=t_2-t_1=2\cdot3\ RC-0\cdot1\ RC$$

$$=2\cdot2\ RC$$

$$=\frac{2\cdot2}{\omega_2}$$

$$=\frac{0\cdot35}{f_2}.$$

Thus rise time is directly proportional to the time constant and inversely proportional to the upper cut off frequency, f_2.

(3) Pulse Input :

A square pulse of width t_p and height E_0 is applied at the input of low pass RC circuit of fig. 10.

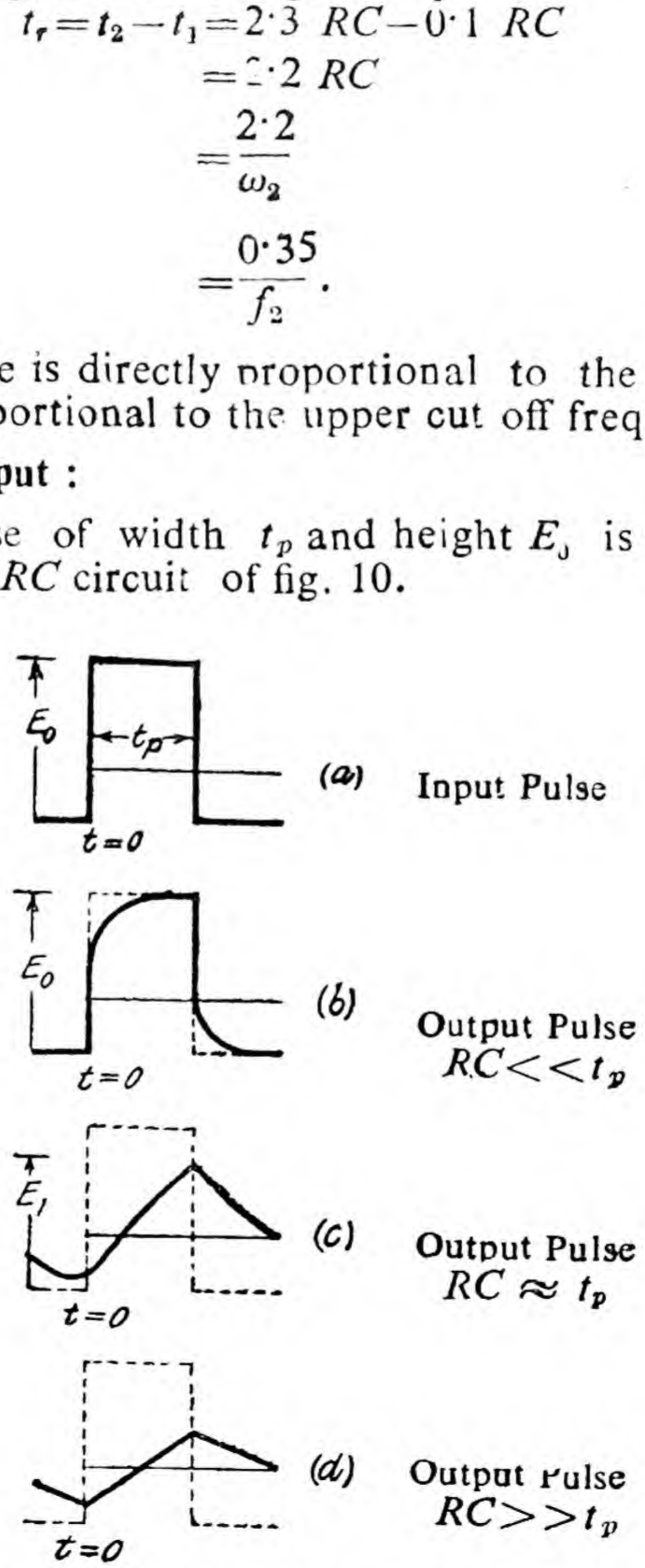


Fig. 13. Response of low pass RC circuit for pulse input.

The response of the circuit for such a pulse for time less than the pulse width, t_p, is the same as that for a step input and is given by

$$e_0 = E_0 (1 - e^{-t/RC}), \qquad 0 < t < t_p.$$

At the end of the pulse *i.e.*, at $t = t_p$, the output rises exponentially to a value E_1 where

$$E_1 = E_0 (1 - e^{-t_p/RC}).$$

After it, the output decreases exponentially to zero with time constant RC. We observe that, in case of low pass RC circuit, output waveform is distorted and extends beyond the pulse width t_p. This is because the charges accumulated on the condenser during charging (0 to t_p) can not leak off instantaneously. For the distortion to be minimum, *risetime*, t_r, *should be small compared with pulse width* t_p. If upper cut off frequency, f_2, is chosen equal to $1/t_p$ then

$$t_r = 0{\cdot}35\ t_p.$$

Thus a pulse shape will be preserved if the upper cut off frequency is approximately equal to the reciprocal of the pulse width. For a pulse of duration, $t_p = 0{\cdot}5\ \mu$ sec., we shall then require upper cut off frequency, f_2, equal to

$$= \frac{1}{0{\cdot}5 \times 10^{-6}} = 2 \text{ mc's.}$$

(4) Square Wave Input :

Let a square wave having a period greater than rise time, be applied at the input of low pass RC circuit of fig. 10. We get steady state response. We can describe, now, the response of the circuit as follows (Refer to fig. 14) :

(i) First, soon as the positive step voltage is applied, almost all voltage drops across R and no voltage appears across C. Therefore output at $t = 0$ is zero.

(ii) Now the cap acitor starts charging through R and the output rises exponentially with time constant RC towards E_0 and reaches a value E_1' at $t = t_1 = T_1$.

(iii) Now at $t = t_1$, a negative step voltage appears at the input so that condenser C starts discharging exponentially till next positive step voltage appears at the input at $t = t_2 = T_1 + T_2$.

(iv) After $t = t_2$, the state of affairs is now like points (i) and (ii) described above.

The tip values E_1' and E_2' of the output voltage can be calculated as follows :

Let E_1 and E_2 denote the tip values of input voltage then :

(i) the exponential curve of output voltage from E_2' to E_1' (in period $t = 0$ to $t = t_1$) is given by

$$e_0 = Ae^{-t/RC} + B,$$

where A and B are the constants, to be determined by the boundary conditions :

$$t=\infty \qquad e_0=E_1$$

and

$$t=0 \qquad e_0=E_2'$$

which when applied to the above equation for e_0, yield

$$E_1=B$$

and

$$E_2'=A+B$$

so that

$$E_2'-B=A$$

or

$$E_2'-E_1=A$$

Puting the values of A and B, we get

$$e_0=(E_2'-E_1)\, e^{-t/RC}+E_1 \qquad \ldots(1)$$

when $t=T_1$, $e_0=E_1'$, therefore

$$e_0=E_1'=(E_2'-E_1)\, e^{-T_1/RC}+E_1 \qquad \ldots(2)$$

Similarly the exponential curve of output voltage from E_1' to E_2' (in period $t=t_1$ to $t=t_2$) is given by

$$e_0=A'\, e^{-t/RC}+B'$$

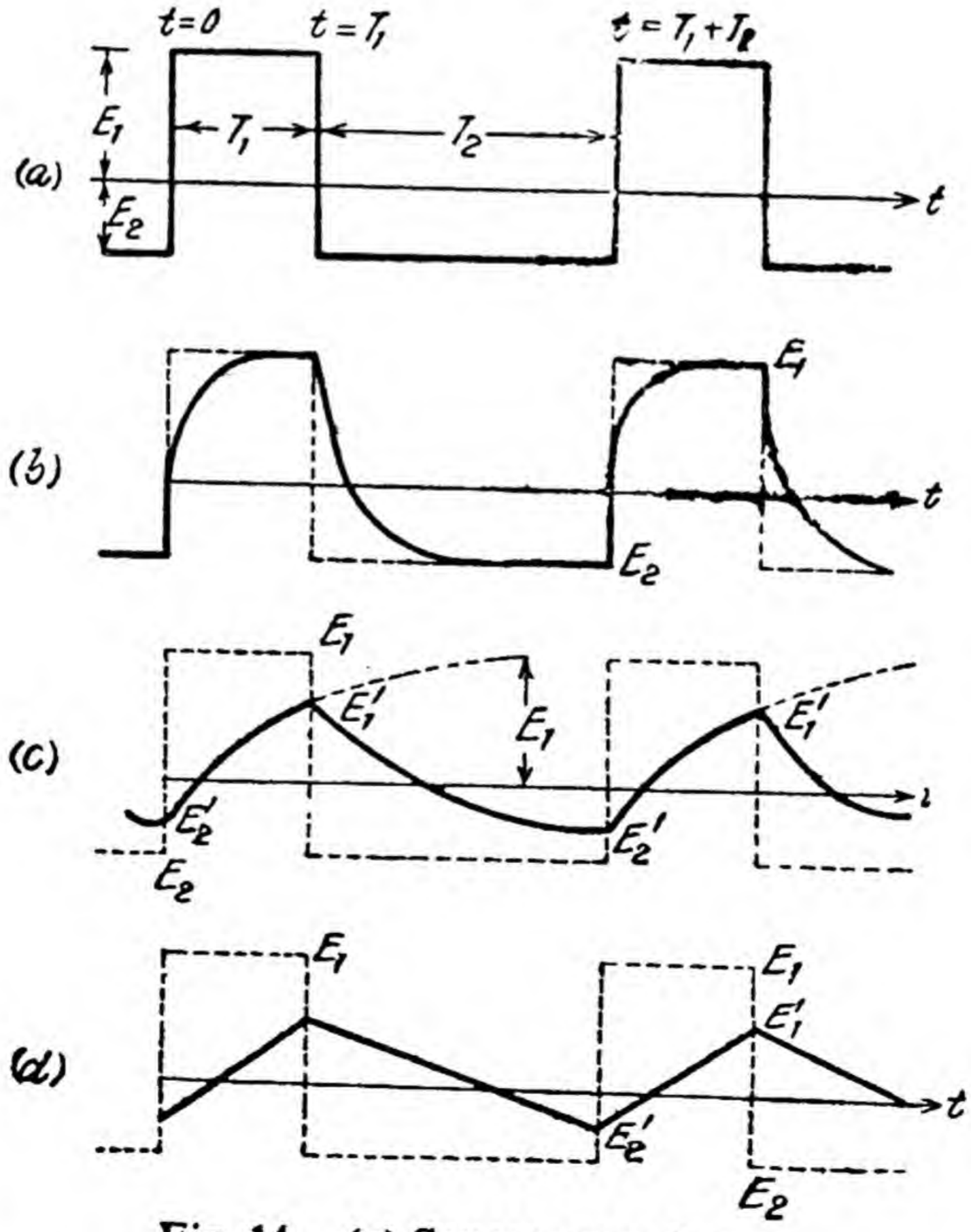


Fig. 14. (*a*) Square wave input.
(*b*), (*c*) and (*d*) Output waveforms of low pass *RC* circuit. Time constant, *RC*, is smallest for (*b*) and largest for (*d*).

Applying

$$t=\infty \qquad e_0=E_2$$
$$t=T_1 \qquad e_0=E_1'$$

we get

$$A'=(E_1-E_2)\, e^{T_1/RC}$$
$$B'=E_2$$

so that

$$E_0=(E_1'-E_2)\, e^{T_1/RC}.e^{-t/RC}+E_2$$
$$=(E_1'-E_2)\, e^{-(t-T_1)/RC}+E_2 \qquad ...(3)$$

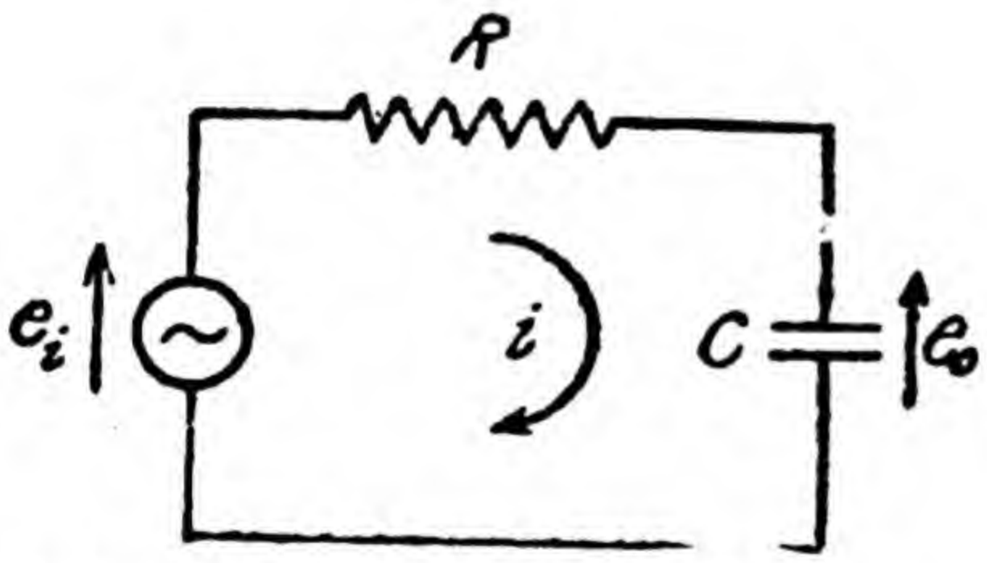
$t=(T_1+T_2)$, $e_0=E_2'$; therefore

$$E_2'=(E_1'-E_2)\, e^{-T_2/RC}+E_2 \qquad ...(4)$$

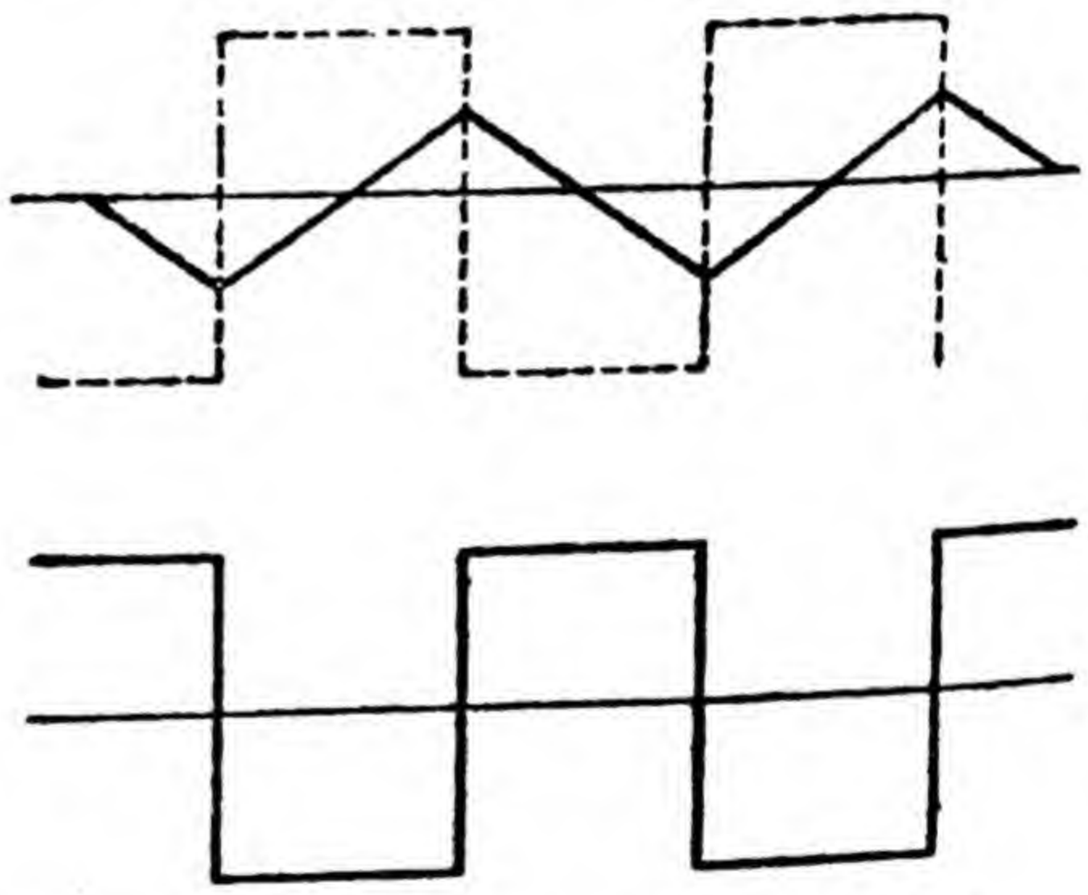
Eqs. (2) and (4) can be applied to determine the tip values of output voltage.

21·4. LOW-PASS RC CIRCUIT AS AN INTEGRATOR :

A low pass *RC* circuit is shown in fig. 15 (*a*). It consists of a high resistance in series with large capacitance, output being taken across *C*.



(*a*) A low pass *RC* circuit acting as integrating circuit with $RC > 15\ T$.



(*b*) Input square wave to an integrating circuit and integrated output.

Fig. 15. Integrating Circuit and Waveforms.

The expression for input wave can be written as

$$e_i=\frac{1}{C}\int idt+iR$$

or

$$i=\frac{e_i}{R}-\frac{1}{RC}\int idt.$$

If *time constant RC' is very large compared* to the interval of input signal then the second term of above equation will be very small and hence

$$i=\frac{e_i}{R}$$

since output of this circuit is taken across C, output voltage will be

$$e_0=\frac{1}{C}\int i\,dt$$

$$=\frac{1}{C}\int\frac{e_i}{R}\,dt=\frac{1}{RC}\int e_i\,dt$$

Obviously, the output voltage is proportional to the integration of input voltage. Thus the wave is said to be integrated when it develops another wave, the amplitude of which is proportional to the integral of input wave, (fig. 15 *b*). But two important points are to be noted :

(i) If $e_i=\alpha t$ then

$$e_0=\frac{\alpha t^2}{2RC}.$$

As time increases, drop across condenser will **not** remain negligible compared with that across resistor, R and the output will not remain the integral of the input.

(ii) The integral of a constant is a linear function and this agrees with the curve of fig. 14 (*d*) which corresponds to $RC \geqslant T$. As the value of RC/T decreases, the departure from true integration increases as indicated in figs. 14 (*b*) and 14 (*c*).

We can obtain a criterian for good integration interms of the time period of a sinusoidal input voltage. We note for better integration CR should be as large as infinity. For which either R should be infinity or C should be infinity. This is, however, not possible. But from case of sinusoidal input we observe that

$$\tan(-\theta)=\omega CR$$

so that if CR is infinity, $\tan\theta$ should be infinity. In other words θ should be of 90°. For all practical purposes, we can choose value of θ to be 89·4° so that

$$\tan 89\cdot4=\omega CR$$

$$100=2\pi fCR$$

$$100=\frac{2\pi CR}{T}$$

$$CR=\frac{100}{2\pi}T$$

$$\approx 15T,$$

where T is the time period of input sine wave.

Thus $$RC \geqslant 15\,T$$

is a criterian for good integration.

(B) NON-LINEAR WAVE SHAPING

In non-linear wave shaping, we shall discuss clipping and clamping circuits.

21·5-1. CLIPPING OR LIMITING CIRCUITS :

The shaping a wave in which it is flattened off or limited to a particular reference voltage level is called clipping. According to operation, clippers can be divided into three classes :

(a) Peak clipper or peak limiter : In such type of clipping circuit, the top portion of wave amplitude (positive or negative both) is limited so as not to exceed a value set up by clipper. Such clipping can be performed by a diode or triode.

In figure 16, a diode clipper is shown. It is obvious that when $e_i < E_1$, the diode T_1 will not conduct at positive peaks of the input wave. Further at negative peaks, if $e_i < E_2$, the diode T_2 will not conduct. Therefore when e_i lies between E_1 and $-E_2$, neither of the diodes will conduct and the input voltage will be transmitted to

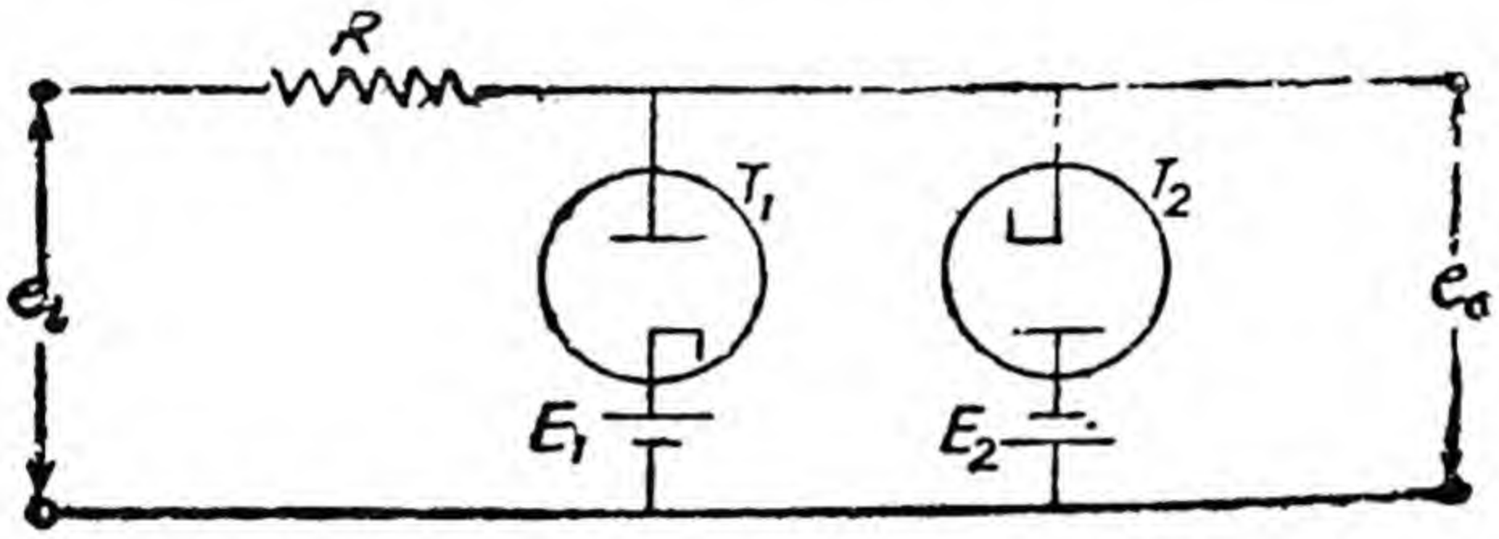


Fig. 16. Diode peak limitter.

output terminals. But when input voltage e_i is more positive than E_1 the diode T_1 will conduct and if series resistor R is sufficiently high† this conduction will limit the positive peak of output voltage approximately to the level of E_1 (fig. 17). Similarly, when e_i is more negative than E_2, the diode T_2 will conduct and with the same assumption for R, this conduction will limit the negative peaks of the output voltage to the level of E_2 (fig. 17). Thus both peaks are clipped.

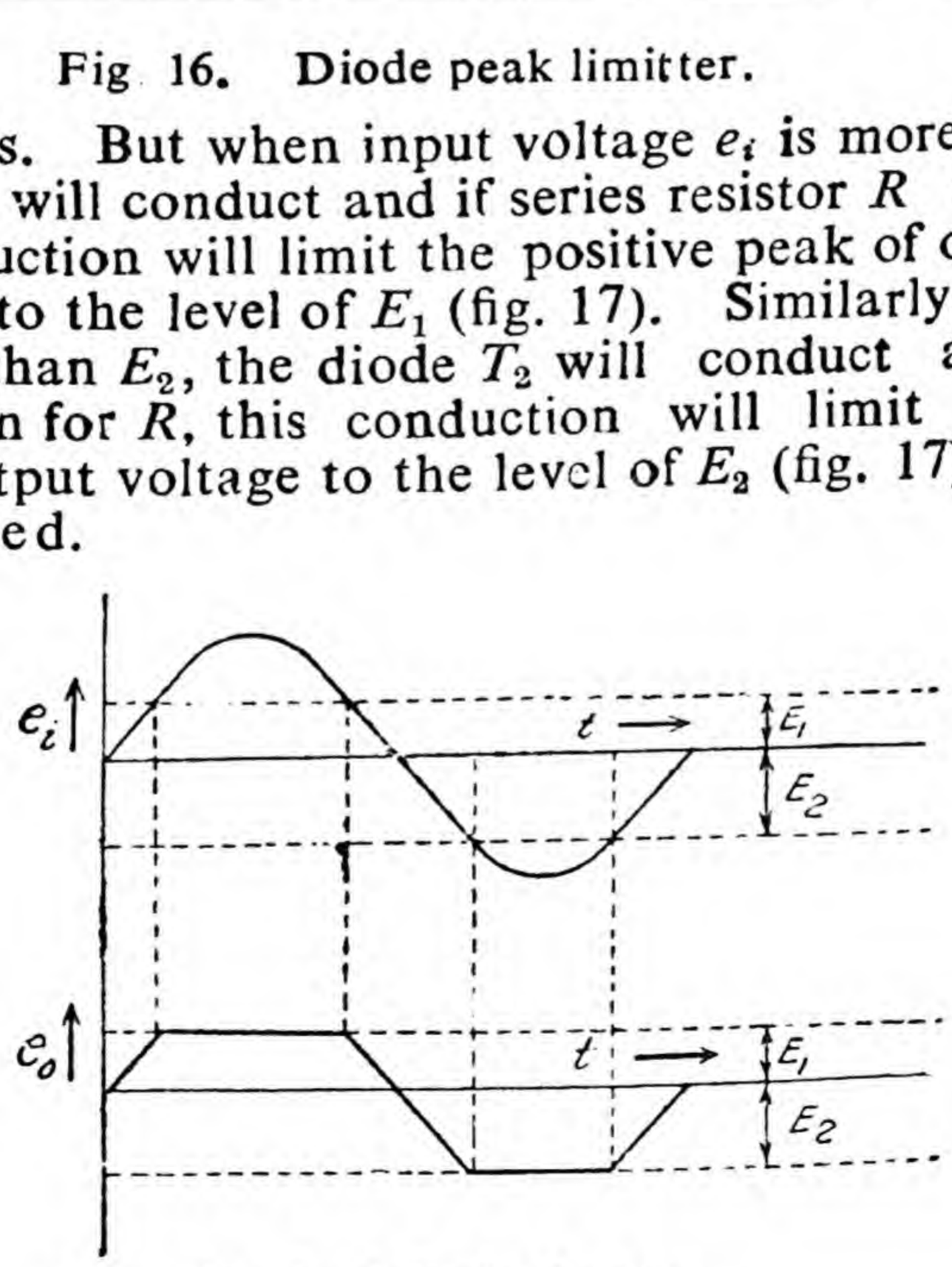


Fig. 17. Peak limited wave.

†If R is high compared to diode resistance then drop will be almost across R and hence linear.

In figure 18, a *triode clipper* is shown. In this circuit, grid and cathode function like diode T_1 of fig 16 to limit the positive peaks of the output voltage to a level E_1. This is termed as grid limiting.

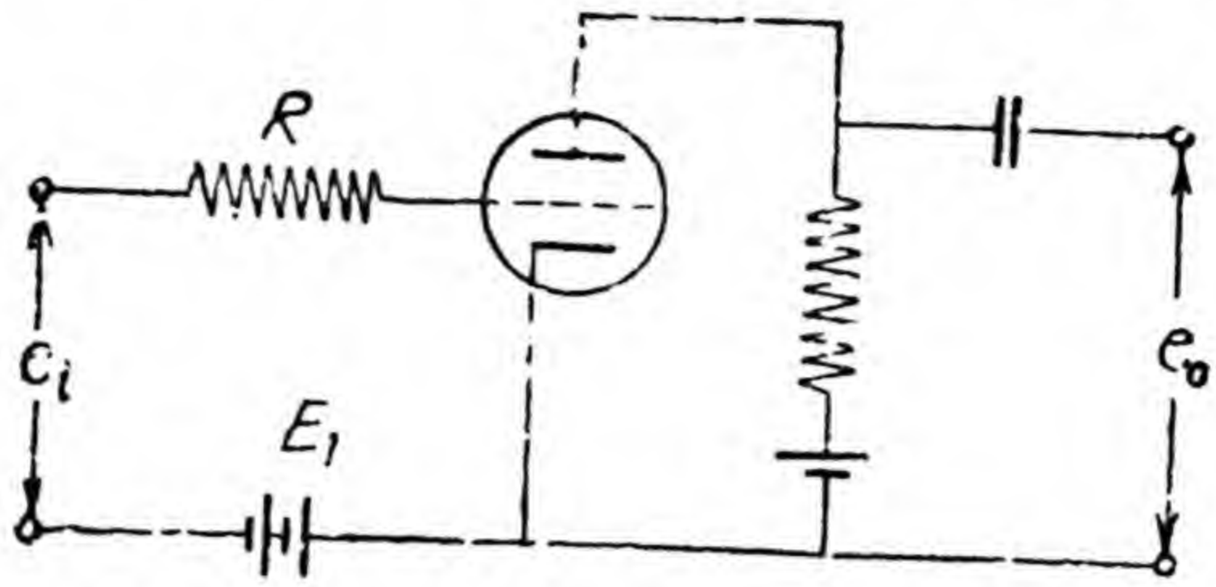


Fig. 18. Triode peak limitter.

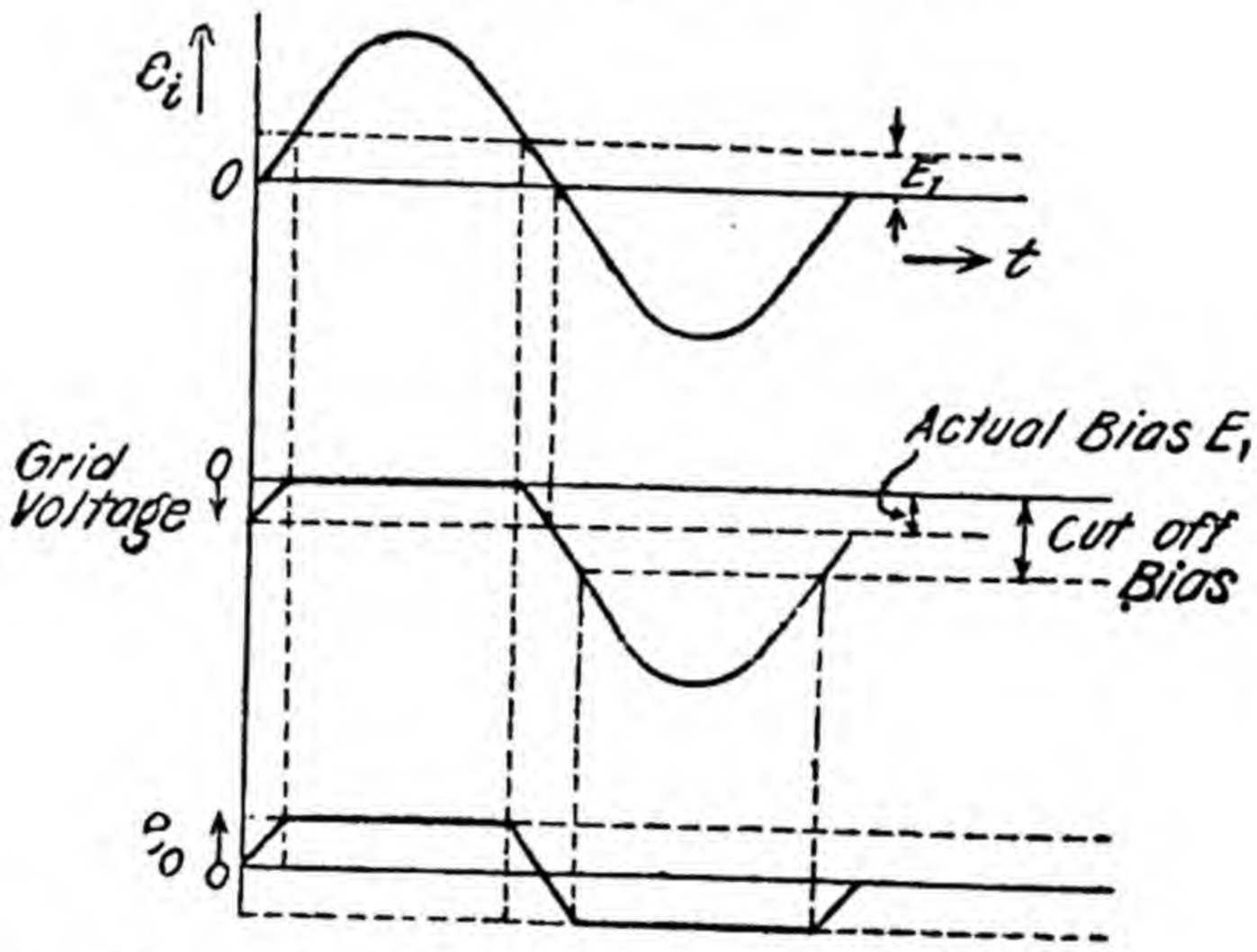


Fig. 19. Figure illustrating the clipping of a sine wave by triode circuits.

Now at negative peaks of input voltage e_i, the amplitude of e_i is large compared with the cut off grid voltage for the circuit and thus tube stops operation. This results in flattening of waveform at cut-off bias level (fig. 19). This is termed as cut-off limiting. Limited waveforms are shown in fig. 19.

(b) Base Clipper : Circuit of a diode base clipper is shown in figure 20. The waveform in such a clipper, is modified such as to remove the amplitudes that are less than an arbitrarily chosen level, say E_1.

It is obvious from figure 20, that output voltage will be obtained only when the amplitude of input voltage exceeds E_1. Thus only peak portion of the wave results in the output (fig. 21).

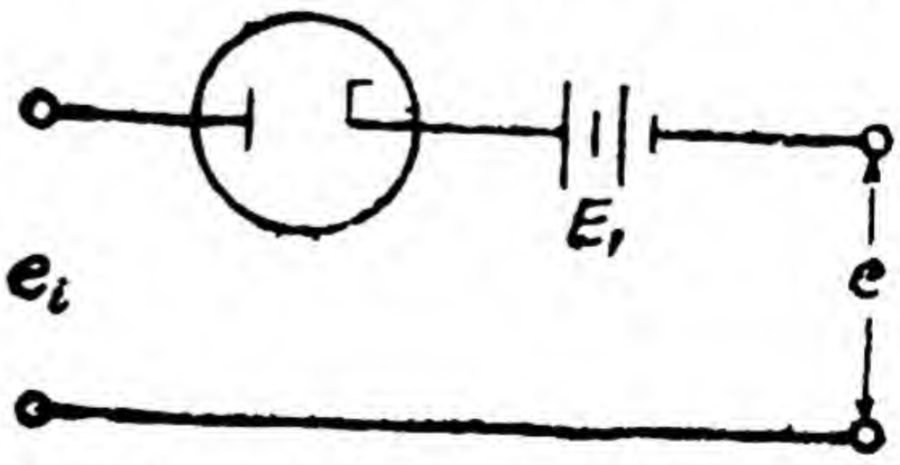


Fig. 20. Diode base clipper.

Such base clipping may also be affected by using a triode instead of diode (fig. 22). Suppose the level below which we want

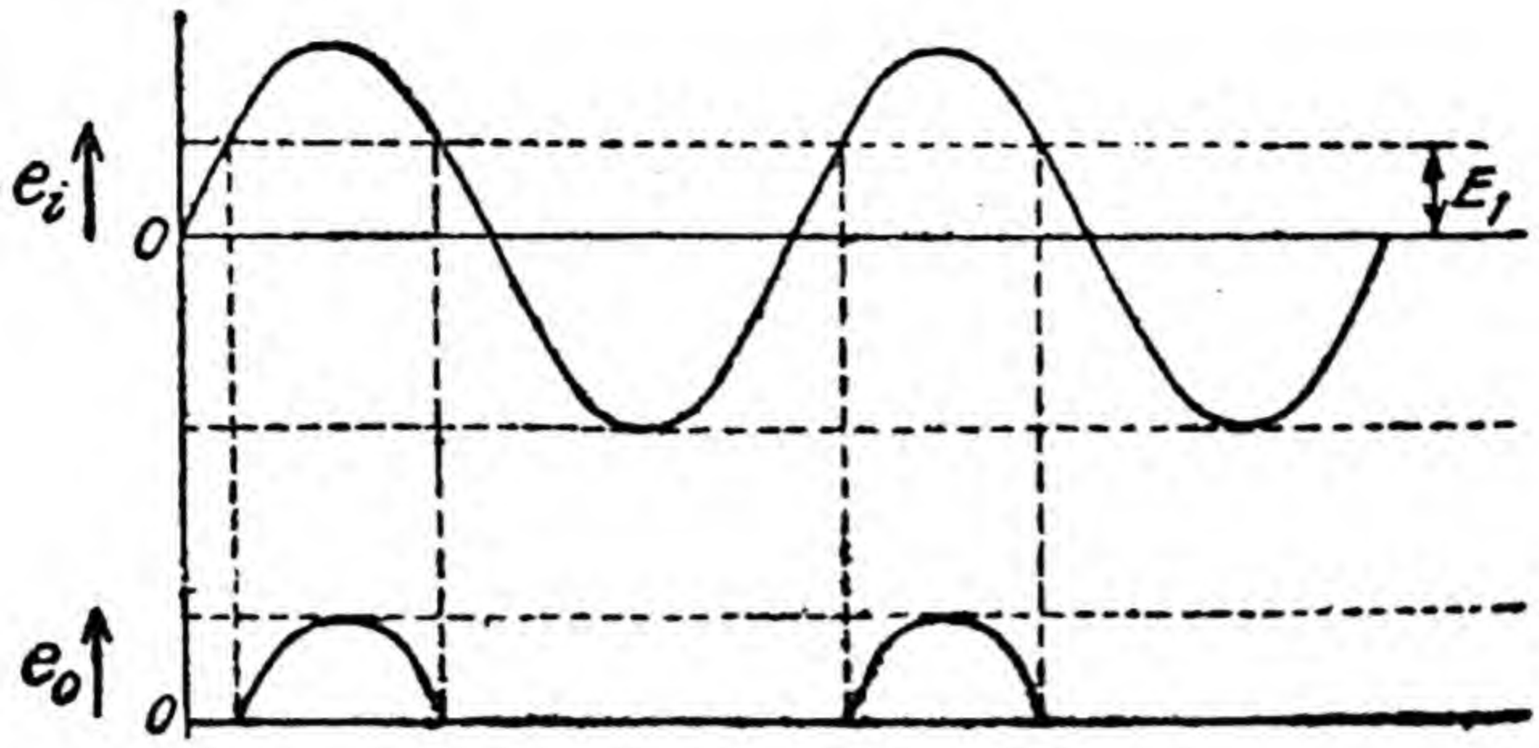


Fig 21. Base clipped waveform.

the waveshape to be removed is E_1. For this, the tube is biased to a value that is more negative than cut-off by an amount of this desired level. The shaping is shown in fig. 23.

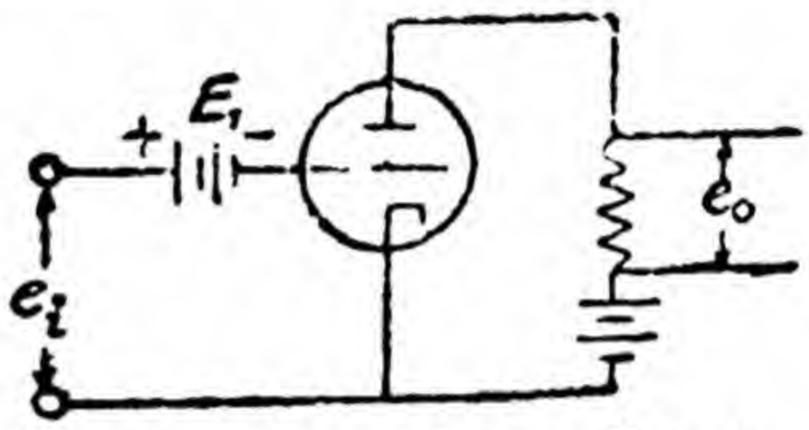


Fig. 22. Triode base clipper.

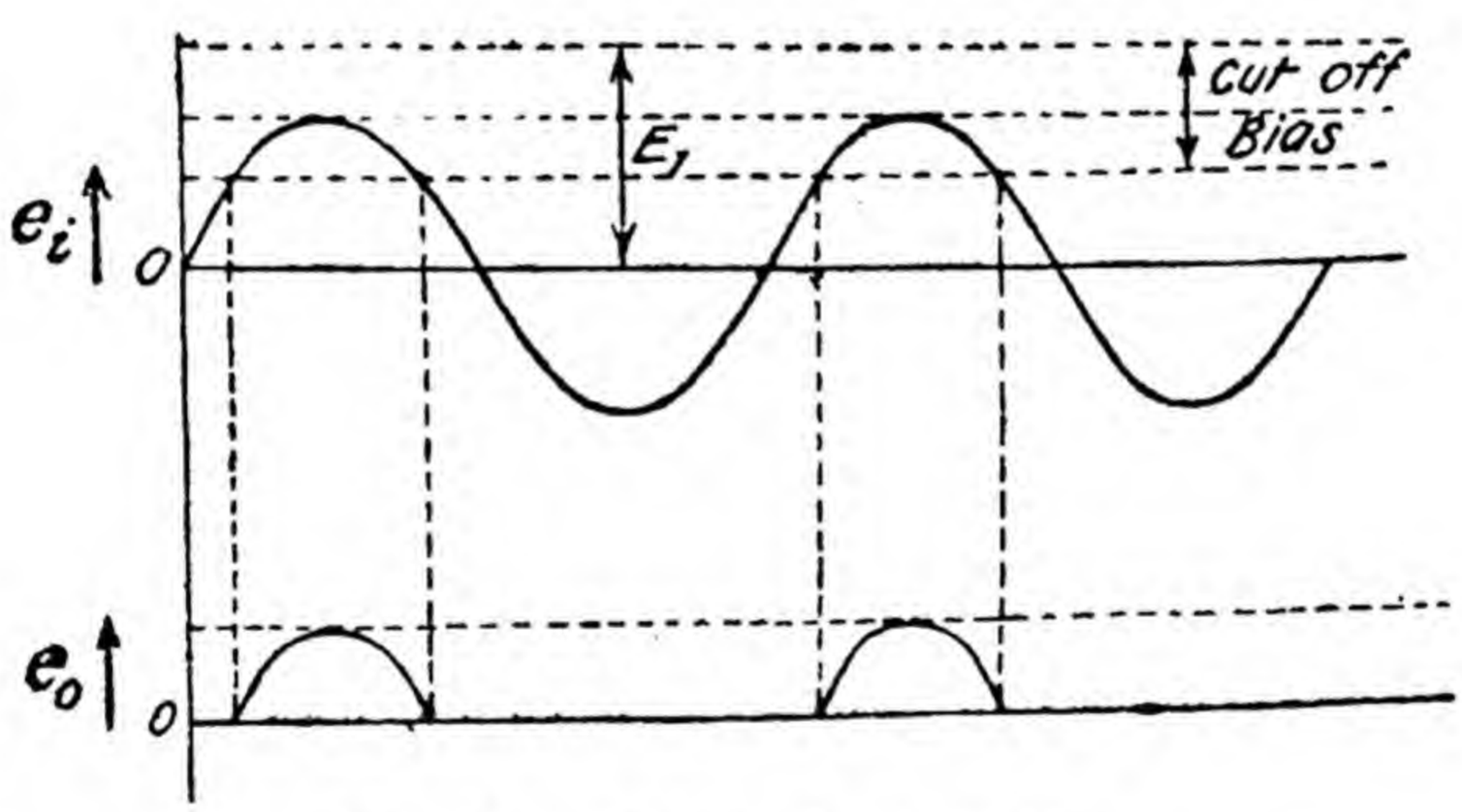


Fig. 23. Clipped waveform.

(c) Slicer : An arrangement in which both peaks and base are removed, *i.e.*, a slice of wave appears at the output, is termed as slicer. Such a circuit is shown in fig. 24, in which two diodes T_1 and T_2 are used and they are biased to the two desired levels E_1 and

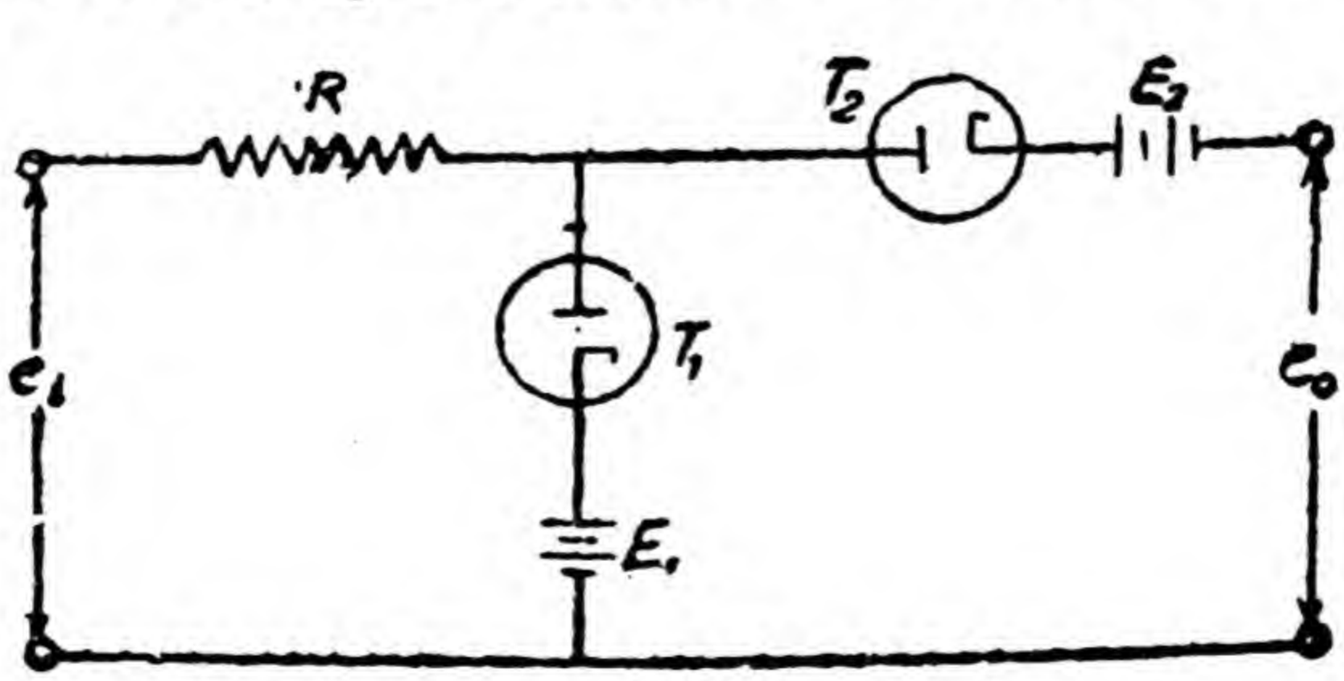


Fig. 24. A slicer circuit.

E_2. How slicing action takes place is shown in figure 25. Diode T_1 (high resistance R and bias E_1 unit) serves as a peak limiter while diode T_2 in association with E_2 acts as a base clipper. Therefore when input voltage exceeds E_1, output of amplitude $(E_1 - E_2)$ is produced while all input voltages with amplitude less than E_1 are reduced to zero amplitude in the output.

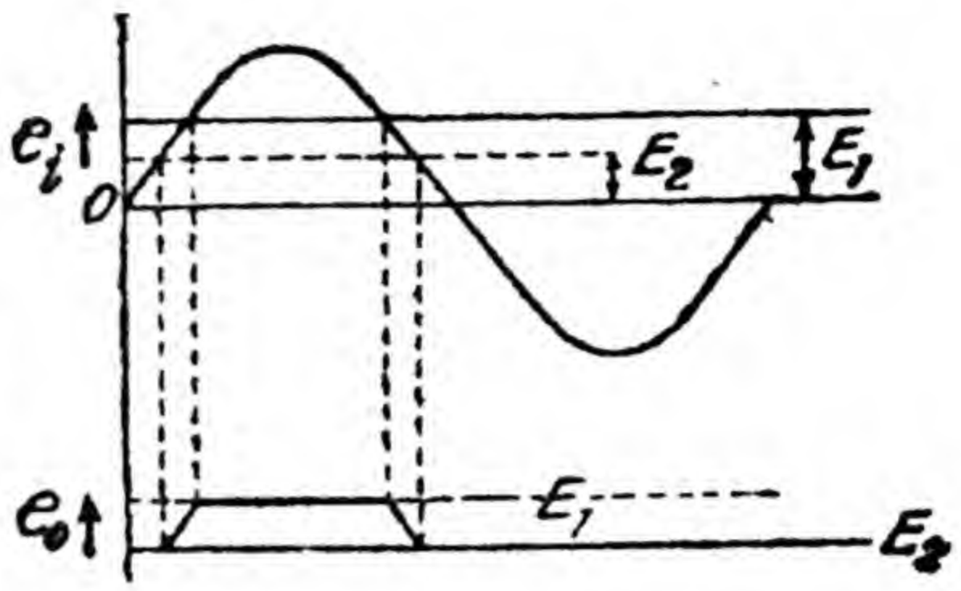


Fig. 25. Sliced waveform.

21·5-2. CLAMPING CIRCUITS :

Circuits which hold either the positive or negative extremeties of any waveform at some predetermined (fixed) reference level are called clamping circuits.

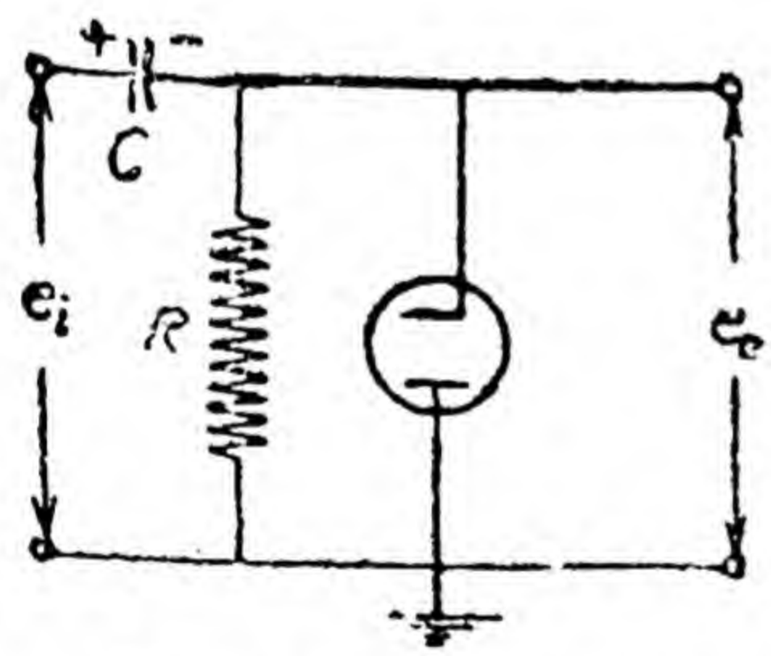


Fig. 26. Diode positive clamping.

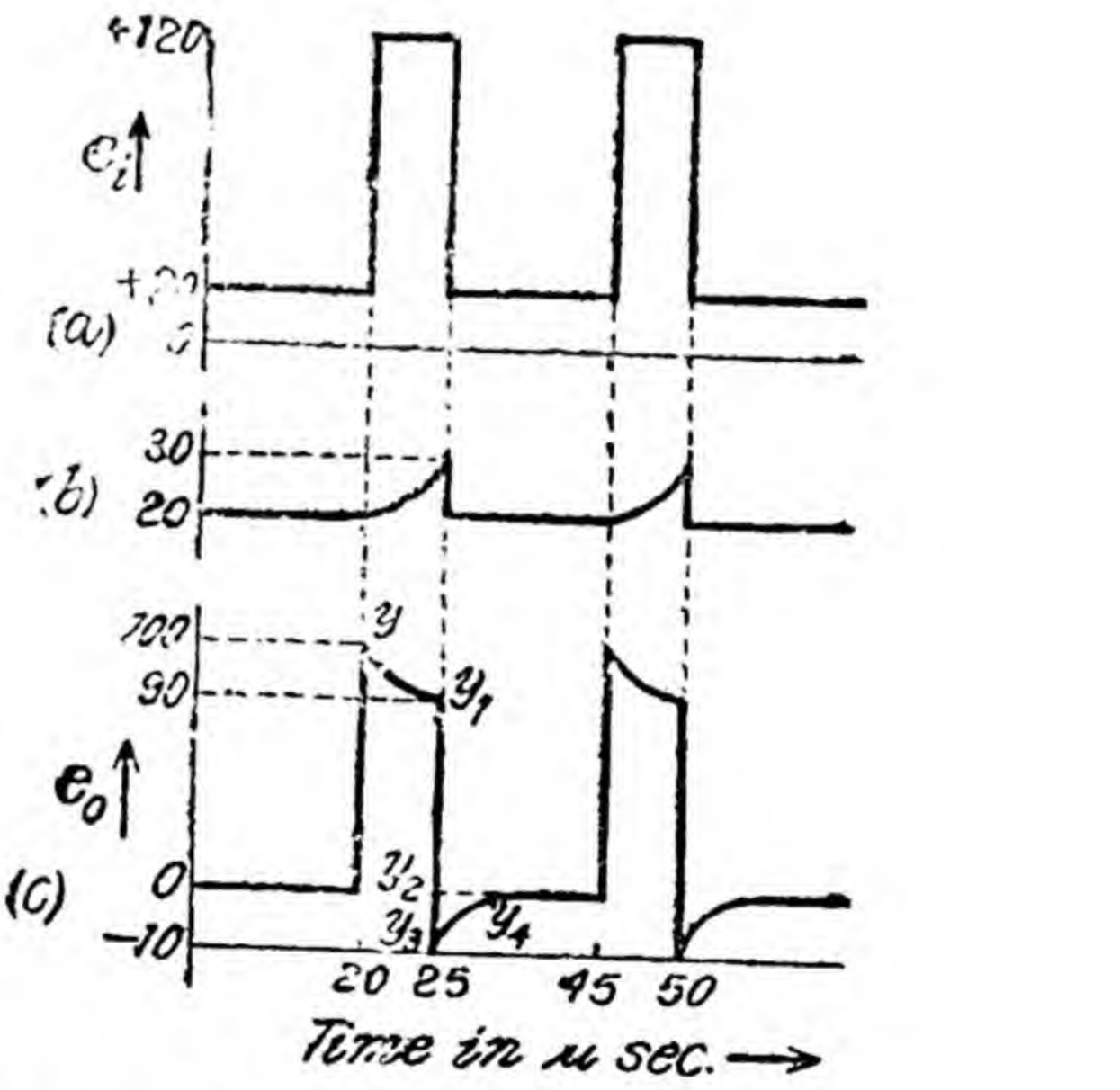


Fig 27. Positive Clamping
(a) waveform at input e_i.
(b) charge and discharge curve of capacitor.
(c) output e_0.

Sometimes it is required to amplify a complex waveform having a high d.c. component value. Such a wave can not be applied as such to the grid of the amplifier since d.c. value of waveform would affect the grid bias and hence the operation of the amplifier. Therefore, it is necessary to remove this d.c. level which can easily be accomplished by passing the wave through a capacitor which will block this d.c. component. After amplification, we again want the same d.c. level to be present in the wave. To restore such a d.c. level, we use the above clamping circuit. Therefore, these circuits are also called as *d.c. restorer.*

In the present discussion, we shall describe positive and negative diode clampers. Both the circuits consists of a diode with *RC* network. The essential conditions for such circuits are :

(*i*) source and diode should be of low resistance,

(*ii*) time constant *RC* should be considerably large compared to input wave period.

(i) Positive Clamping Circuit : Positive clamping circuit is shown in fig. 26. In fig. 27 (*a*), a square wave of 100 volt amplitude with a d.c. component (average value) of 70 volts [20+(120−20)/2] is shown.

When this wave is applied to the clamping circuit, then during the interval of 20μ sec (from 0 to 20μ sec.) the capacitor charges to full 20 volts through *R* with polarity shown in fig. 26. The diode does not conduct since its plate and cathode are at the same potential. The output is zero.

Now at the instant $t=20\mu$ sec, applied voltage increases to 120 volts. Since the charge on condenser (which is already charged to 20*V*) cannot change instantaneously, a voltage of 100*V*=(120−20) will appear across *R* as output.

As shown in fig. 27 (*b*) because of large time constant *RC* condenser charges itself only to a small extent, say 10 volts more, so that total voltage across condenser becomes 30 volts. This reduces the output voltage from 100 to 90 volts as shown in fig. 27 (*c*) by the position yy_1.

At $t=25\mu$ sec. input voltage drops down to 20 volts level (fig. 27*a*) with a fall of 100 volts. But net voltage across *R* at this instant is 90*V* as a fall of 100 volts (curve y_1y_3) will leave a voltage (90−100), *i.e.* −10 volts across *R*. This explains the negative undershoot y_2y_3 seen in fig. 27 (*c*). This means cathode of diode is at −10 volts with respect to the plate, allowing the diode to conduct. Since the diode resistance is small, the output terminals are short circuited. The output is then zero (y_3y_4 portion in fig. 27 *c*). Since d.c. input of 20 volts is still at the input, capacitor remains charged to 20 volts till the next rise in pulse amplitude comes at $t=45\mu$ sec. *i.e.* output will remain at zero for the interval for (45 - 25), *i.e* , 20μ sec. and after $t=45\mu$ sec. operation repeats itself.

(ii) Negative Clamping Circuits : As shown in fig. 28, a square wave of 100 volts amplitude with a d.c. component of 70 volts is applied to the circuit. The only difference between this circuit and that shown in fig. 26 for positive clamping is that the diode is reversed in polarity.

The condenser is charged to full +20 volts through the resistor *R*. The output is zero since cathode and plate of diode are at the same potential. Then at $t=20\mu$ sec input voltage increases to 120 volt. Since the charge on condenser, already charged to +20 volts can not change instantaneously, full 100 volts will appear across *R*

and consequently, makes the plate of diode 100 volts positive with respect to its cathode. The diode then conducts and the condenser

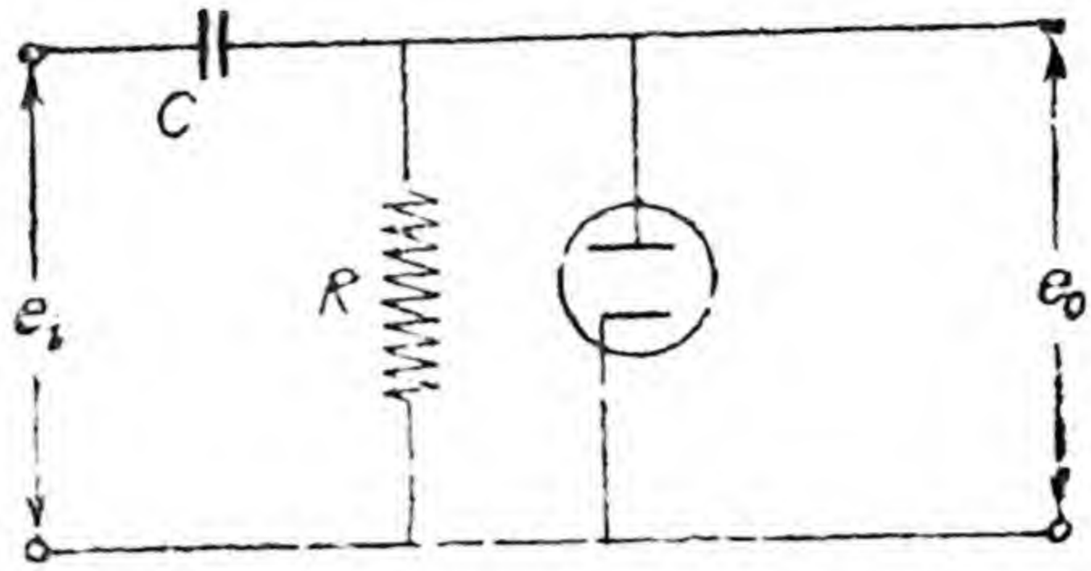


Fig. 28. Negative Clamper.

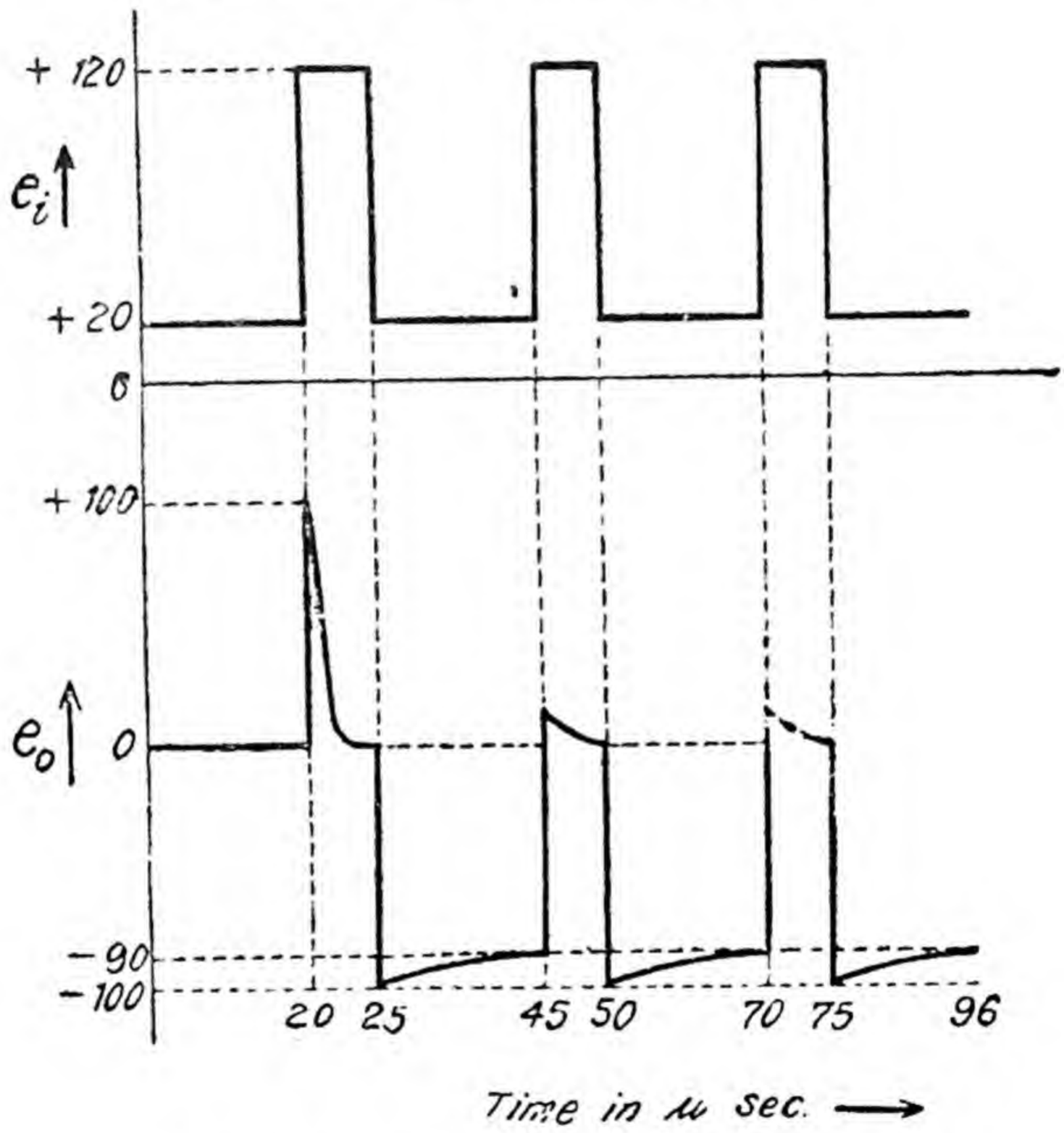


Fig. 29. (*a*) input waveform, e_i,
(*b*) output waveform, e_0.

is charged to +120V, quickly through its low forward resistance. Since the diode resistance while in conduction is small, output terminals are short circuited and output goes down to zero. This operation is shown by a sharp pulse at $t=20\mu$ sec. in fig. 29 (*b*).

Output remains zero for a period of 5μ sec. from 20μ sec. to 25μ sec. At $t=25\mu$ sec. the input goes down to 20 volts level, falling by 100 volts in amplitude. Diode stops conduction. Since condenser is already charged to 20 volts, its charge will not change instantaneously and a full drop of 100 volts will appear across R, the output. This is shown in fig. 29 (*b*) at $t=25\mu$ sec. where output goes from 0 to −100 volts.

For $t=25\mu$ sec. to $t=45\mu$ sec. capacitor discharges to very small extent as the time constant RC is high. We take that

potential across C falls through R by '0 volts so that output comes down from -100 to -90 volts

At $t=45\mu$ sec. input pulse again rises to 120 volts, an increase of 100 volts. This full increase appears across R which already has -90 volts across it. So now voltage across R will become (100−90) *i.e.* 10 volts, making the plate at + 10 volts. The diode conducts and output is reduced to zero as shown by an overshoot at $t=45\mu$ sec. in fig. 29 (*b*).

Output remains at zero for the period $t=5$ to $t=50\mu$ sec. For next periods, cycle described above is repeated.

If the cathode and resistor R are returned to some voltage other than ground, as shown in fig. 30, clamping level will be shifted accordingly. For example if cathode (also resistor R) is biased to 5 volts negative with respect to ground, the output wave-shape would shift down by 5 volts.

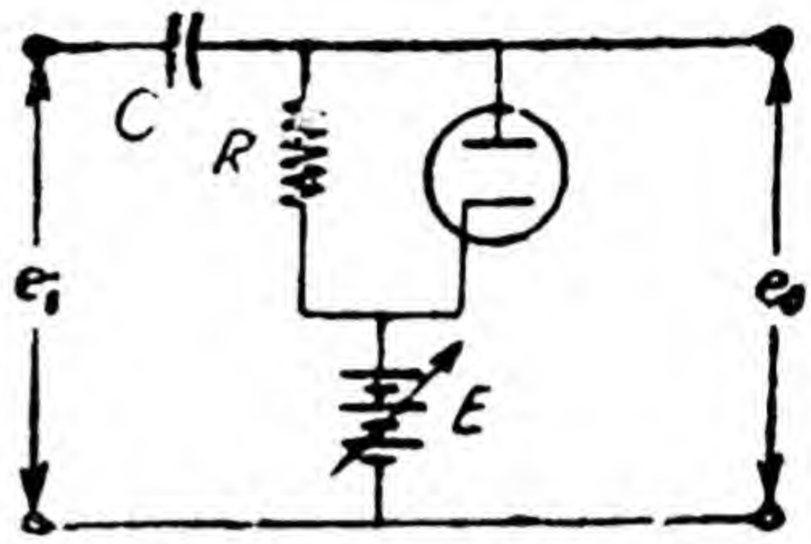


Fig. 30. Circuit for negative clamping at a level other than zero.

21·6. PN JUNCTION CLIPPERS :

A diode volt-amp. characteristic is shown in fig. 31 (*a*). The slope is discontinuous at the voltage V_B. This point is called a *break point* (for Si, $V_B=0{\cdot}6V$; for Ge, $V_B=0{\cdot}2V$). Now refer to the clipping circuit of fig. 31 (*b*). The plot of its output voltage, v_0, as a function of input voltage, v_i, is shown in fig 31 (*c*). Here break point occurs at (V_B+V_R).

(*i*) *For positive peak of input voltage*, v_t : For $v_t < (V_B+V_R)$ diode is reverse biased (OFF) and therefore for this portion input will appear as such at output. But for $v_i > (V_B+V_R)$, diode will conduct. If a silicon diode is employed then drop across diode will be approximately $0{\cdot}6V$ during this internal and the rest of the input will be developed across the resistor, R. Infact, increment in output is given by

$$\Delta v_0=\Delta v_i\left(\frac{R_f}{R_f+R}\right),$$

where R_f is diode forward resistance. If $R_f \ll R$, then positive

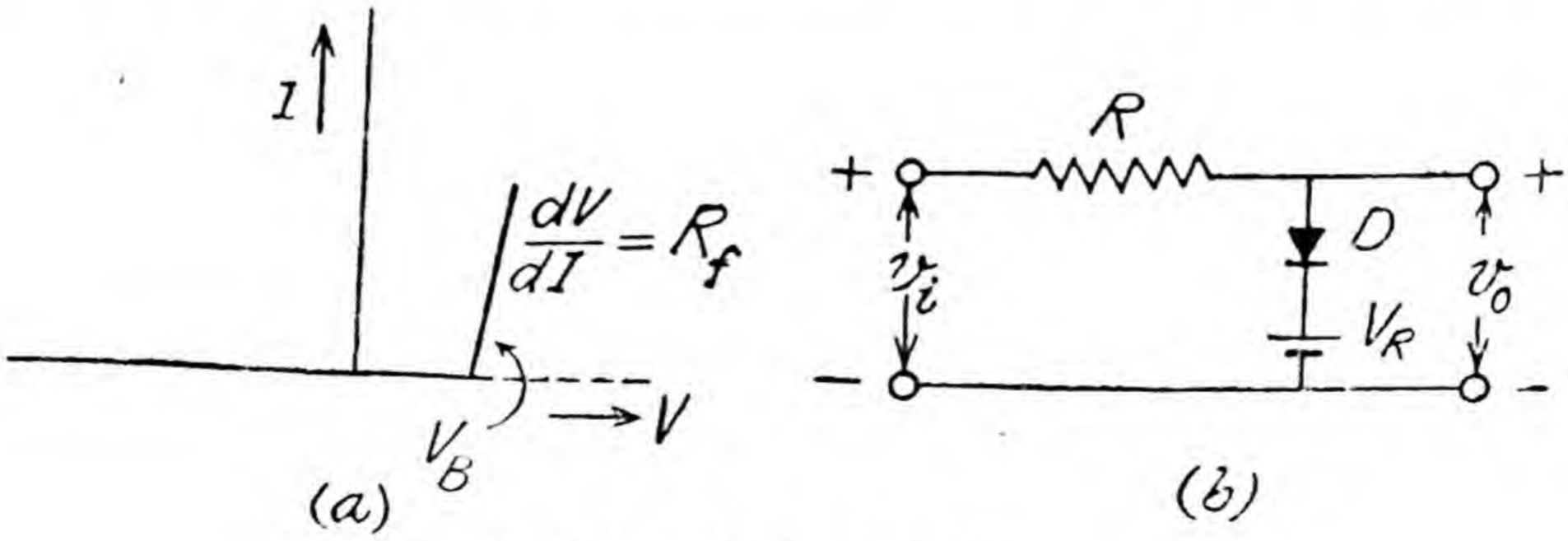


Fig. 31 (a), (b). a diodə clipper.

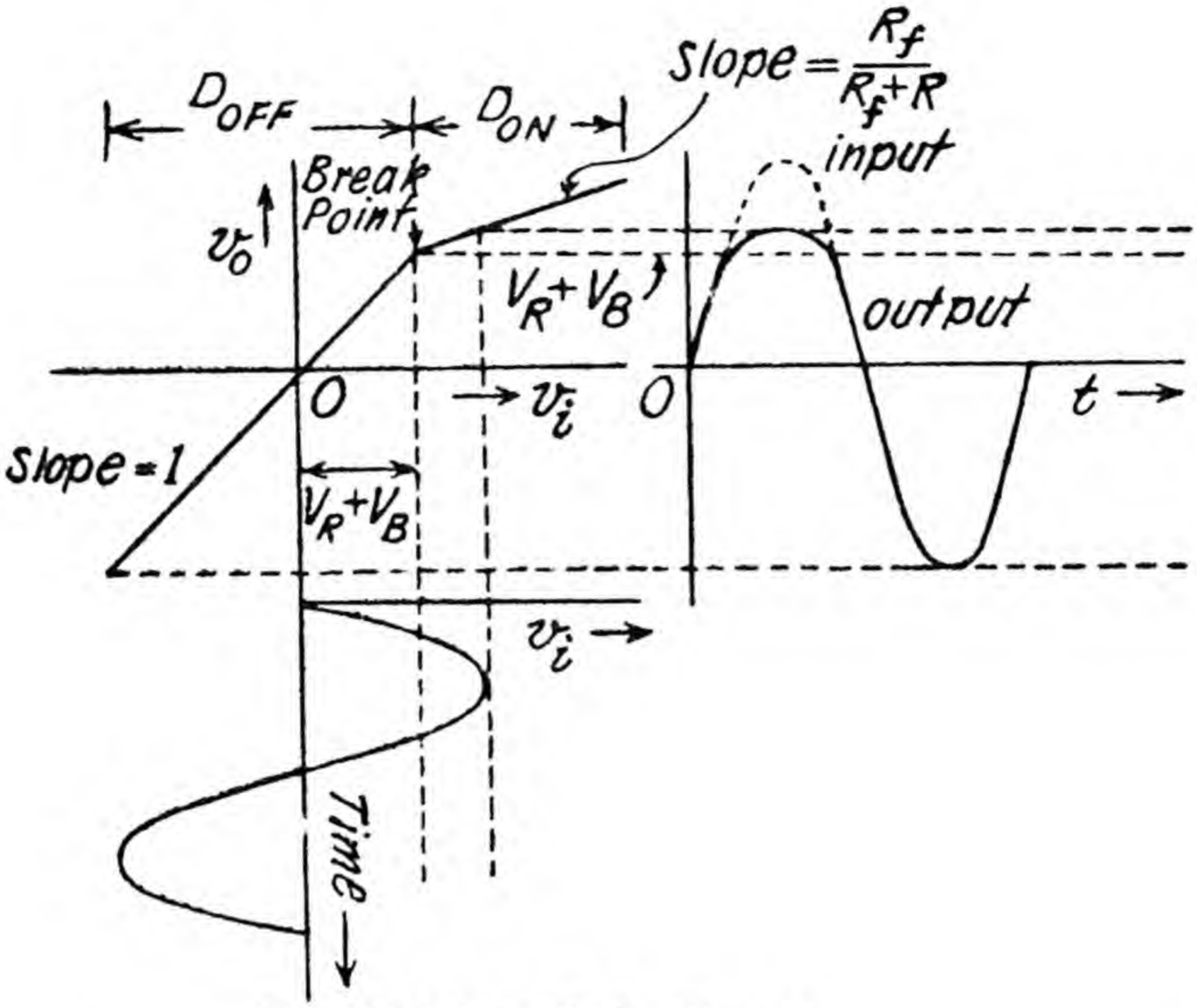


Fig. 31 (c) A diode clipper.

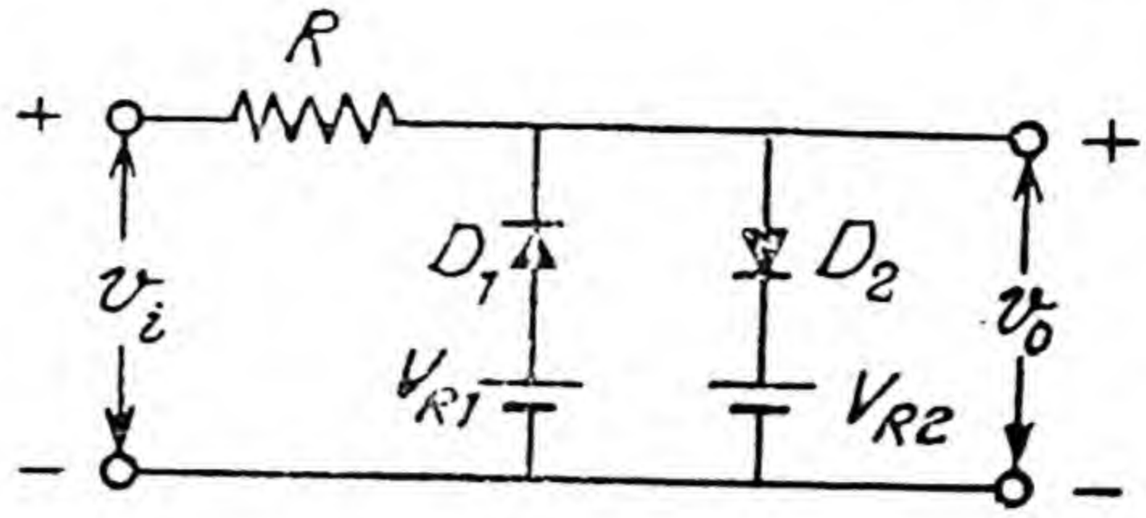
peak of the output will be sharply limited at the voltage (V_R+V_B) and it will appear as though the positive peak has been clipped off. When $V_B \ll V_R$, positive peak is clipped at the voltage V_R and is thus *limiting reference voltage*. Thus output positive peak can be clipped at the desired level, V_R.

(*ii*) *For negative peak of input voltage, v_i :* When the input goes negative, diode will be reverse biased and if small reverse leakage current is ignored, input will be transmitted directly to the output.

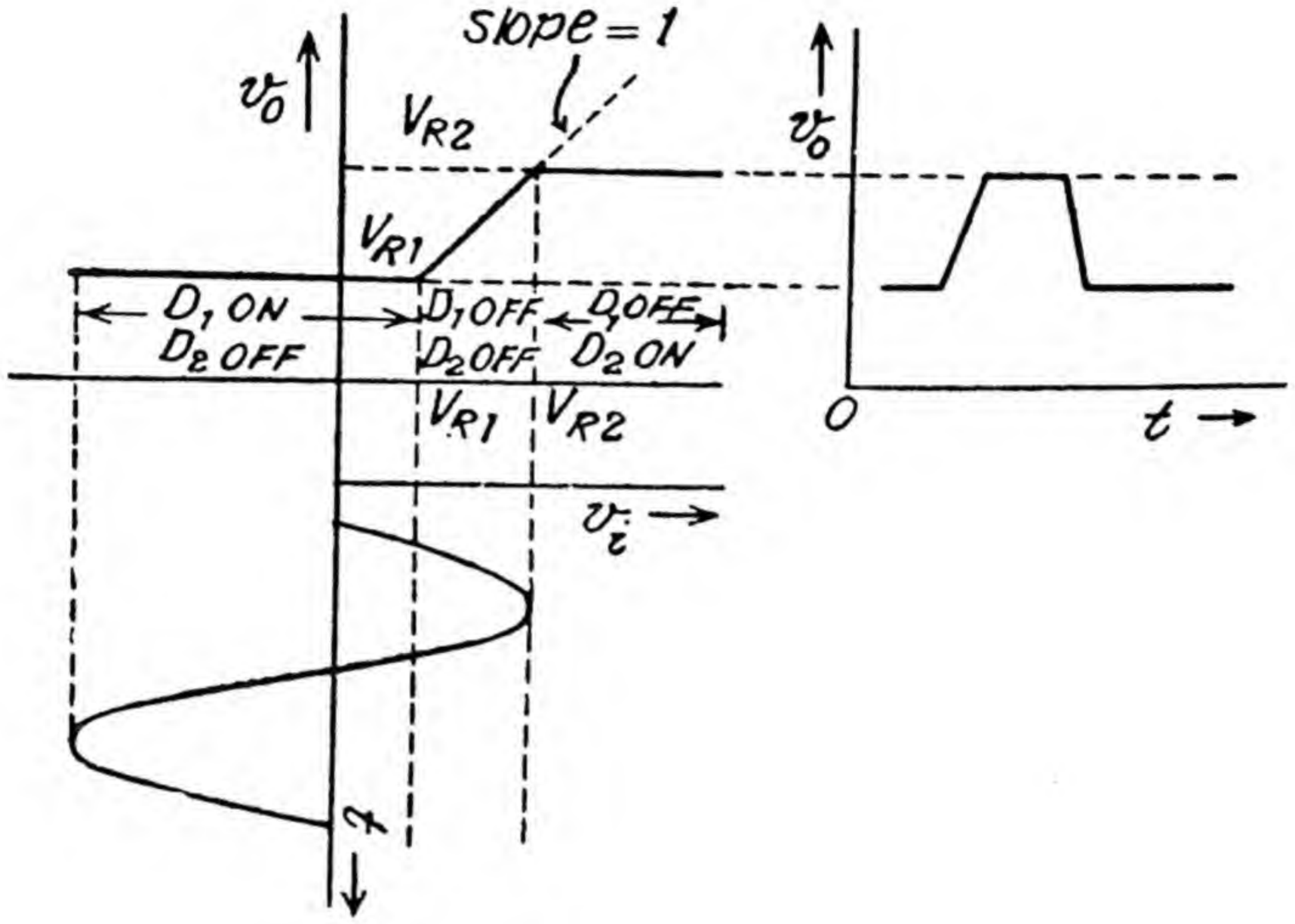
Clipping at two different levels :

A parallel arrangement of two diodes shown in fig. 32 (*a*) is used to clip the input sinusoidal wave at two independent levels. The transfer curve has two break points, one at $v_0=v_i=V_{R1}$ and a second at $v_0=v_i=V_{R2}$, assuming $V_{R2} > V_{R1}$. Operation can be briefed as follows ;

Input, v_i	Output, v_0	Diode states	Description
$v_i \leqslant V_{R1}$	$v_0=V_{R1}$	D_1 ON, D_2 OFF	As diode D_1 conducts, it will limit this portion of input waveform to a level V_{R1}.
$V_{R1} < v_i < V_{R2}$	$v_0=v_i$	D_1 OFF, D_2 OFF	As both diodes are OFF, this portion of input will be transferred directly to the output.
$v_i \geqslant V_{R2}$	$v_0=V_{R2}$	D_1 OFF, D_2 ON	As diode D_2 conducts, it will limit this portion of input waveforms to a level V_{R2}.



(a)



(b)

Fig. 32. (a), (b) Diode Clipping at two independent levels.

The circuit of fig. 32 (*b*) is referred to as *slicer*.

21·7. PN JUNCTION CLAMPING CIRCUIT (d.c. inserter) :

Commonly used dc inserter circuit is shown in fig. 33 (*a*).

Operation :

(*i*) Let initially capacitor be discharged so that $v_{ac}=0$. At $t=0$, input jumps to -5V and remains at this value for a period $T/2$. This input will forward bias the diode and capacitor will rapidly charge upto $5-V$ (provided we neglect forward voltage drop of the diode). Since R_L is shunted by the ON diode, it will

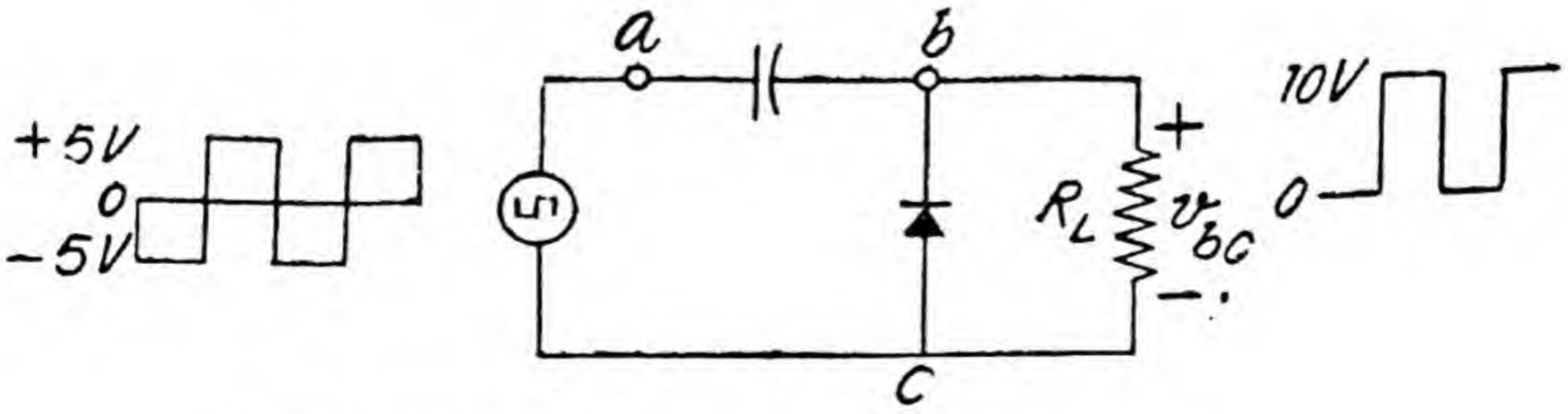
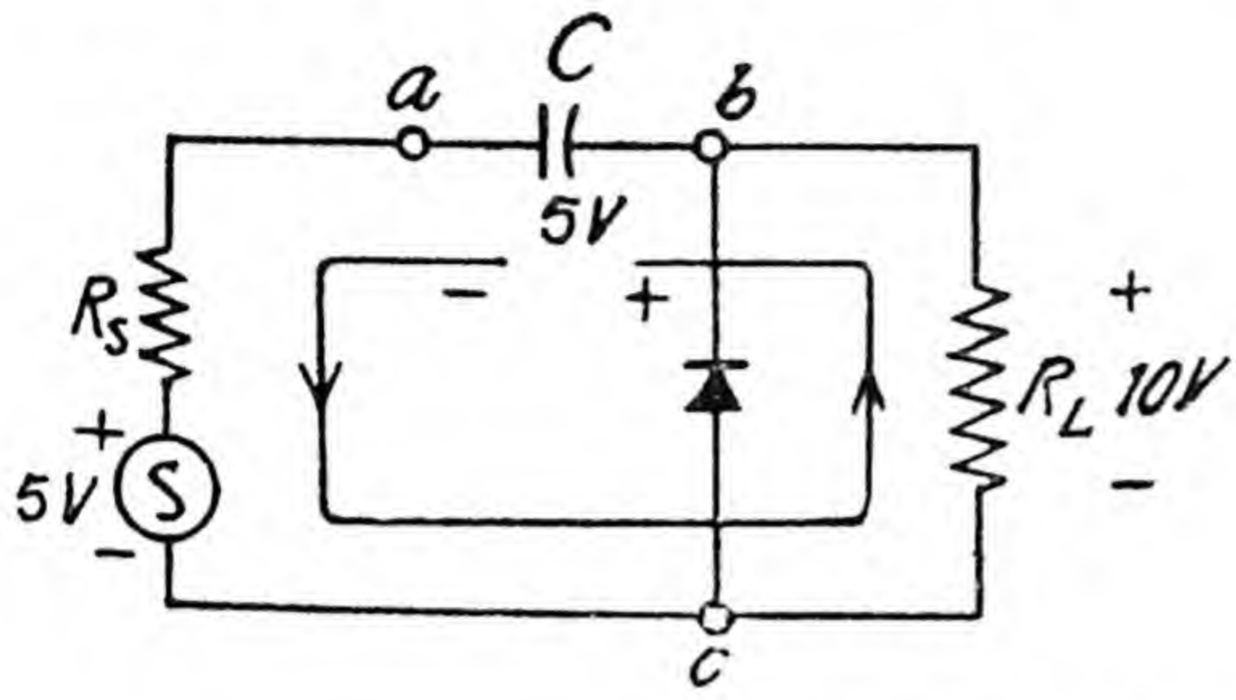


Fig. 33 (a) Diode Clamping circuit.

not be effective in this period and thus output will be close to zero for the first half period ($T/2$) of the input.

(*ii*) At $t=T/2$, input jumps to $+5V$ while *the capacitor will hold its previous charge at* $-5V$. *If time constant is made much longer than* $T/2$, *the capacitor will not have time to change from its previous charge i.e.* capacitor will serve as a source of a constant 5V. As the input +5V and capacitor 5V are in series, they will



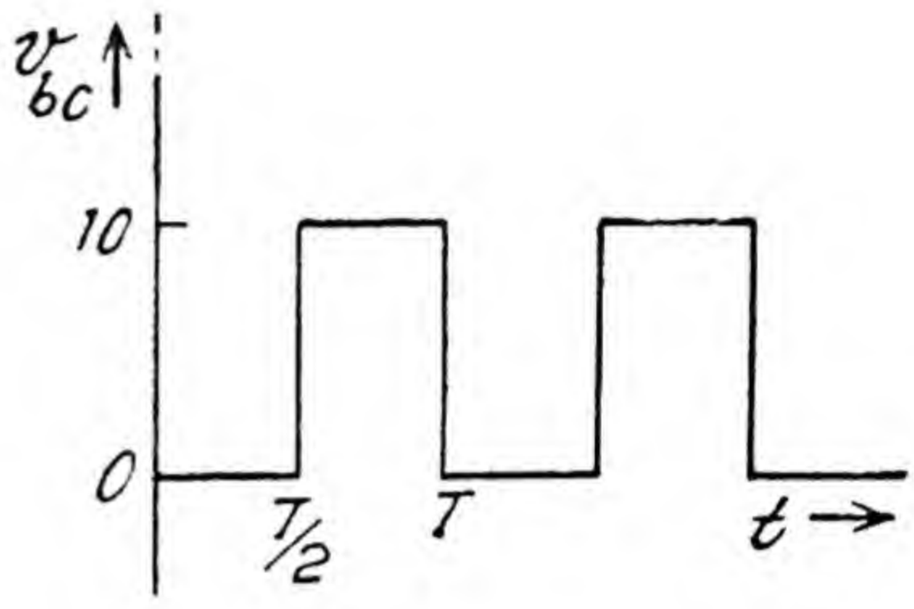


Fig. 33 (b) Clamping circuit waveforms.

add to give an output of 10V (provided R_L is much larger than source resistance fig. 33 *b*). It may happen that during this period, capacitor may lose a slight charge.

(*iii*) When the input reverts back to −5V again, capacitor has already a charge almost upto−5V. Therefore there will be no current flow and the output will be zero. The slight charge lost by condenser during positive peak time will also be quickly made up again. Thus capacitor constantly behaves as a constant source of −5V.

21·8. CIRCUITS TO PRODUCE SHARP PULSES :

Timing markers : In various receivers and transmitters use of very accurate sharp timing markers is made to measure the difference in the time of arrival between two pulses or to produce accurate timed pulse waves by triggering the generating circuits. We shall discuss here the following two techniques :

(*i*) *First Technique* : The block diagram is shown in fig. 34. A high frequency sine wave produced by crystal oscillator is first sliced and then differentiated to produce pulses of short duration (positive which removes pips). These pulses are then applied to

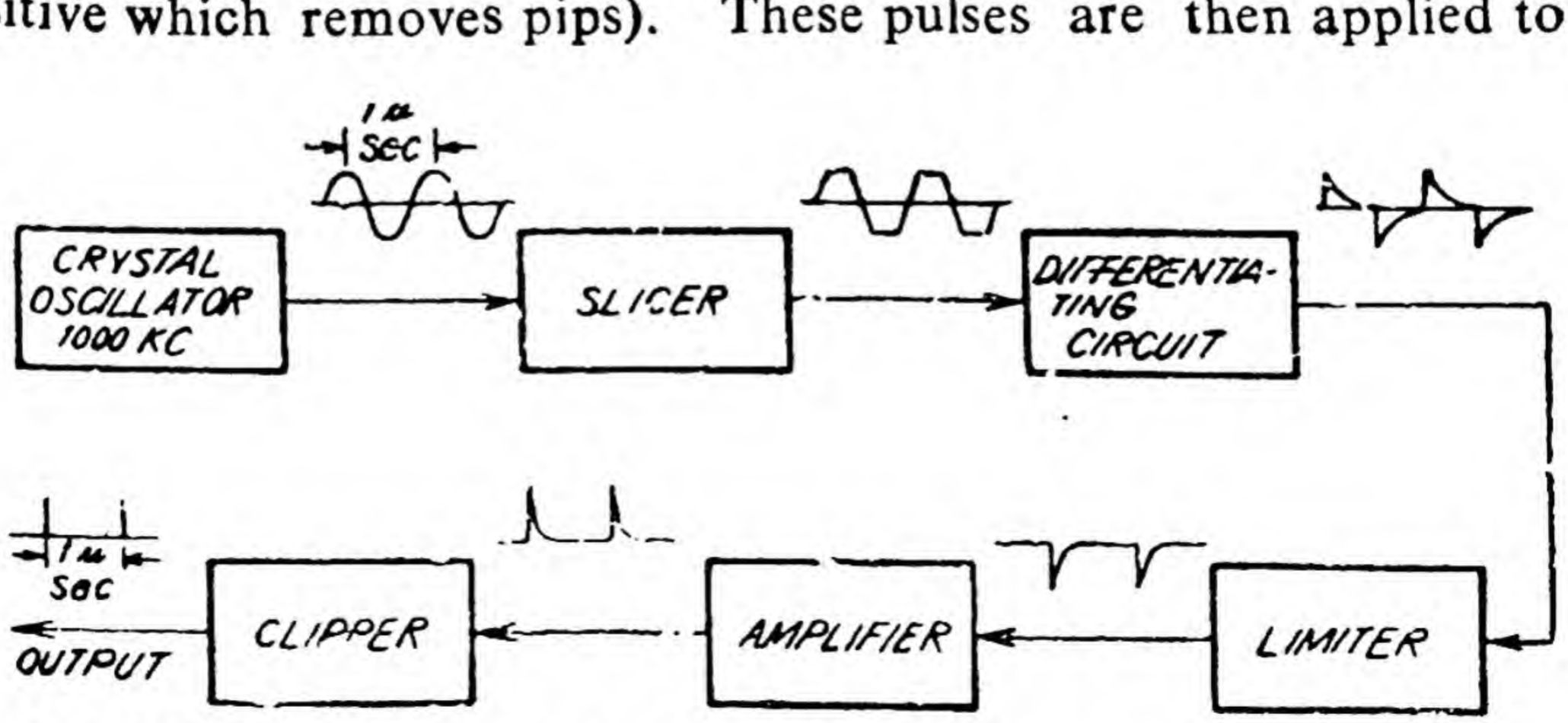


Fig. 34. Timing markers generation with the help of wave shaping circuit.

limiter which removes the positive pips. Suppose the frequency of crystal oscillator is 1000 kc/s. then these negative pips will be of 1μ sec. interval.

To get sharper pips of positive polarity, the output of limiter is fed to an amplifier which will amplify and invert the wave. To remove the broad base these pips are clipped at the suitable level by means of a clipper. We thus obtain '*hair line*' marker pips at 1-μ sec intervals.

(*ii*) *Second Technique* ; *Ringing Circuit* : This circuit is used typically to generate a sequence of pulses spaced regularly in time. Circuit is shown in fig. 35. The grid is slightly positive when no signal is applied. The tube conducts heavily and the steady current builds up a strong magnetic field around the inductor *L*. Output voltage is zero because d.c. resistance of the coil is zero.

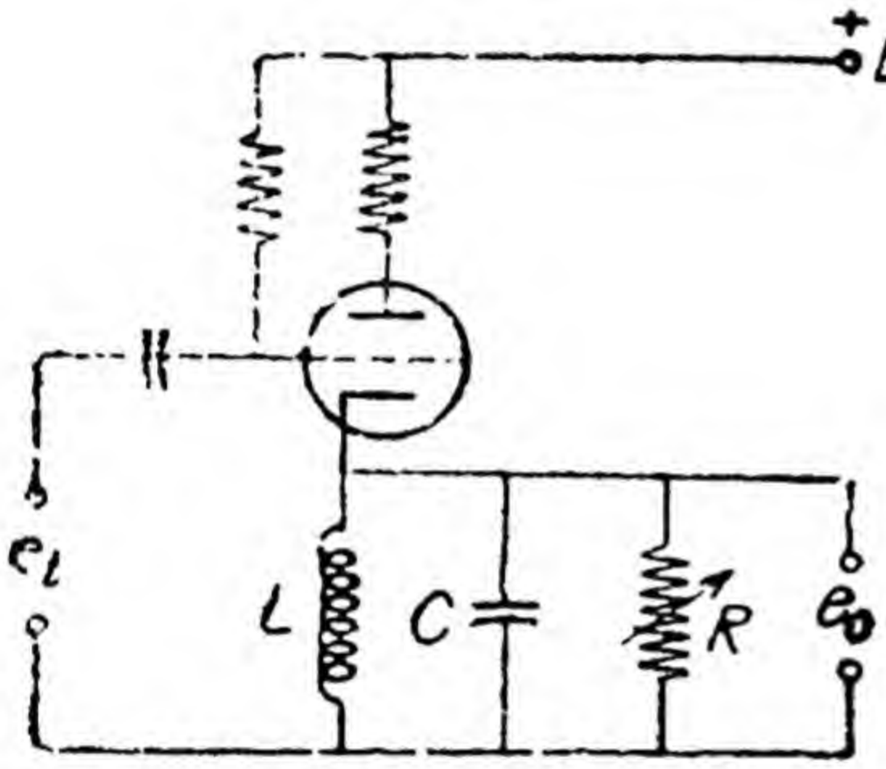


Fig. 35. Ringing circuit.

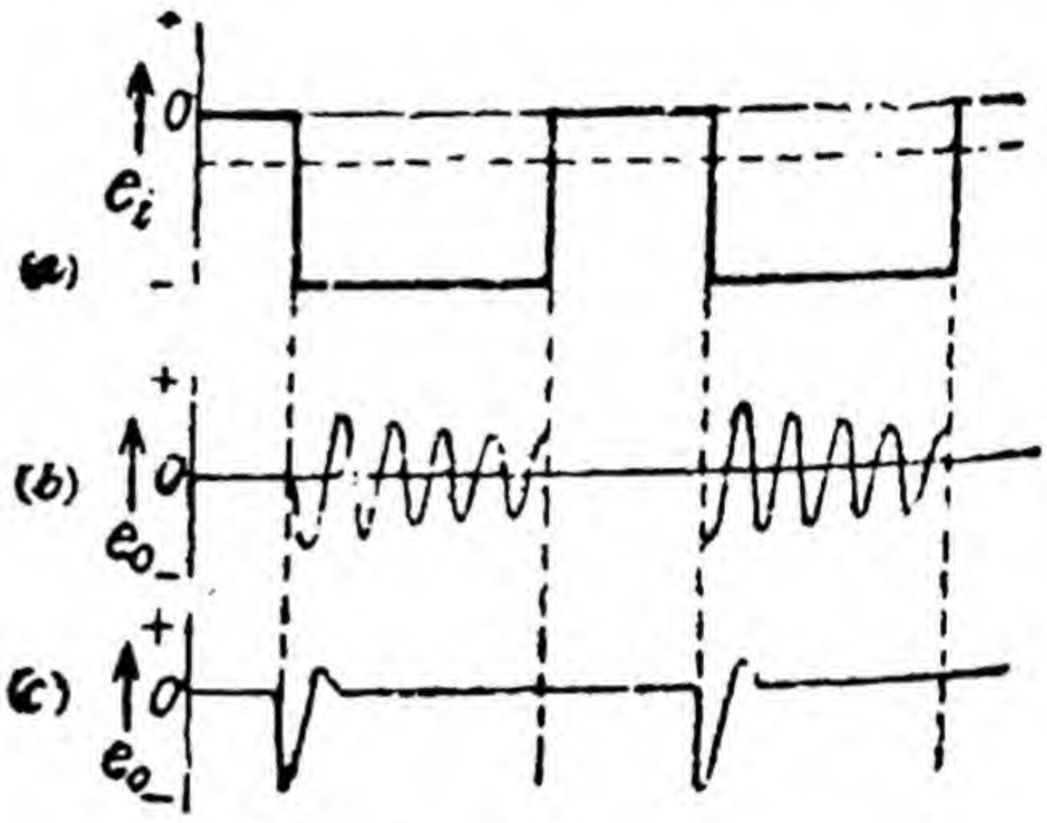


Fig. 36. Waveform.

When the negative going square wave, fig. 36 (*a*), sufficient to drive the tube beyond cutoff, is applied, current stops to flow and energy in the tank circuit oscillates back and forth between the inducator *L* and capacitor *C*, *i.e.*, *LC* circuit is left free to oscillate at its natural frequency. During each oscillation some energy is lost in the resistance *R* so that output is a damped wave shown in fig. 36 (*b*). Oscillations are damped out soon by decreasing the value of *R* , its critical value is given by

$$R=\tfrac{1}{2}\sqrt{(L/C)}.$$

With such a value of *R*, most of the energy is lost in first cycle of operation and the resulting output is of the form shown in fig. 36 (*c*). Thus a pulse is formed which can be further sharpened using clipping and differentiating circuit as shown in fig. 34. It should be noted that

(*i*) width of the pulse, fig. 36 (*c*), produced depends upon *LC*.

(*ii*) pulse duration is shorter and its rise time is faster when the resonant frequency is higher.

Another circuit of ringing oscillator is shown in fig. 37. In the absence of an input signal, the transistor is biased to produce a steady current through the high Q inductor and steady magnetic field is built up around the inductor. When the negative going square wave, sufficient to cause cut off, is applied to the base circuit of the transistor, the steady magnetic field collapses and the LC circuit executes free oscillations at its natural frequency. The collasping field tends to maintain current which charges C and oscillations continue until the end of the input pulse, at which the transistor again conducts. Due to conduction of the transistor oscillations are damped out and magnetic field starts building up around inductor. Thus the process continues.

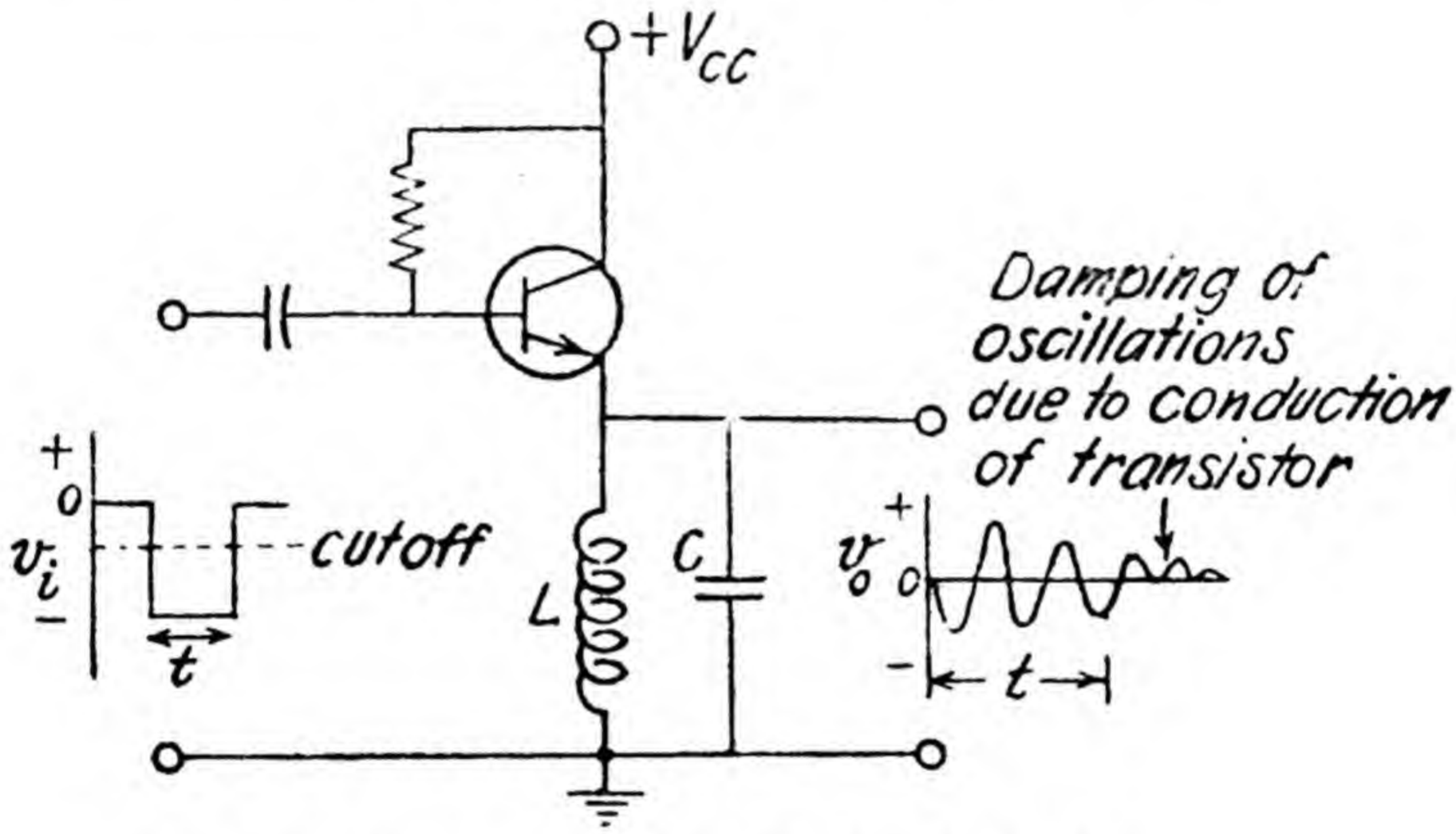


Fig. 37. Transistor ringing oscillator.

EXERCISES

1. Discuss the response of high pass RC circuit to various types of input. When can it work as a good differentiator ?
2. Discuss the response of low pass RC circuit to various types of input. When can it work as a integrator ?
3. Discuss diode and triode clipping circuits when it is desired to clip (*a*) the peak (*b*) the base of the input wave.
5. Discuss positive and negative clamping circuits.

22 VOLTAGE AND CURRENT SWEEP GENERATORS

In chapter on 'Pulse generation and scaling circuits', we have described these generators with circuits using vacuum tubes. As they find wide applications in instrumentation and communication systems, we are describing them in detail with transistor circuits.

22·1. TERMINOLOGY USED TO DESCRIBE SWEEP GENERATORS :

Prior to discussing sweep circuits, we shall explain in brief the terminology used for such a description.

Ideal voltage ramp : In a linear ramp (sweep) generator, the voltage increases or decreases *linearly* with time, and has a duration. At the end of duration, it returns *abruptly* to initial value. In fig 1 (a) ideal voltage sweep waveform (with positive slope) is shown. Voltage ramp (sweep) starts everytime after a certain repetition period T_S, called ramp duration. Such an ideal ramp (sweep) waveform can be expressed as

$$V_S\Big|_t = V_S\Big|_{t=0} + \dot{V}.t \quad \text{with} \quad 0 < t < T_S$$

where $V_S\Big|_{t=0}$ is the initial value of the ramp voltage.

and $$\dot{V} = \frac{dV_S\Big|_{t=0}}{dt}$$

$\dot{V}$ is constant for ideal voltage ramp waveform.

In fig. 1 (b), ramp waveform having a negative slope is shown.

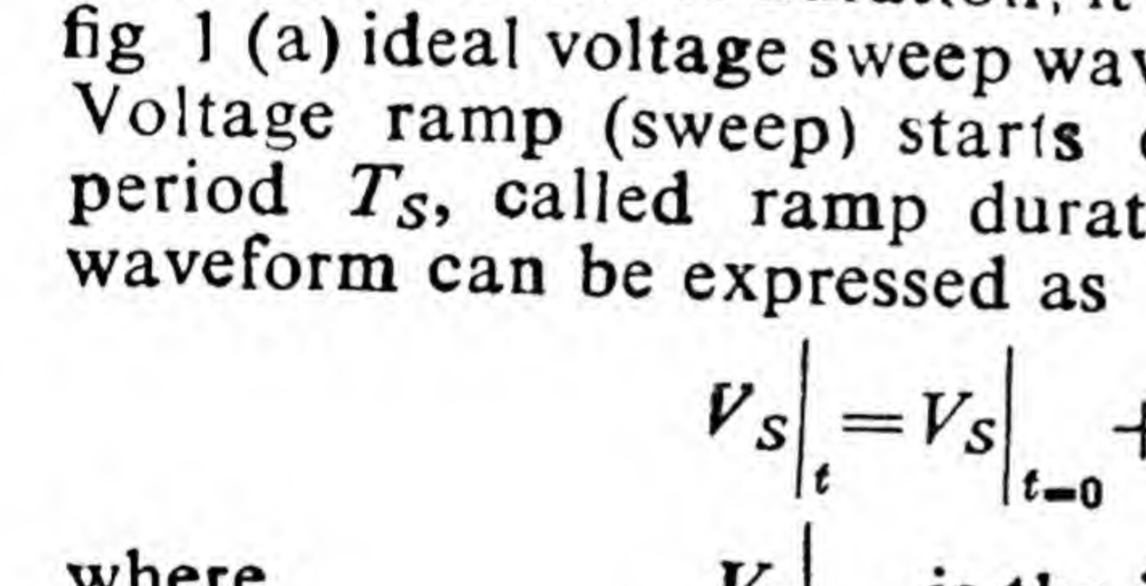


Fig. 1 (a)

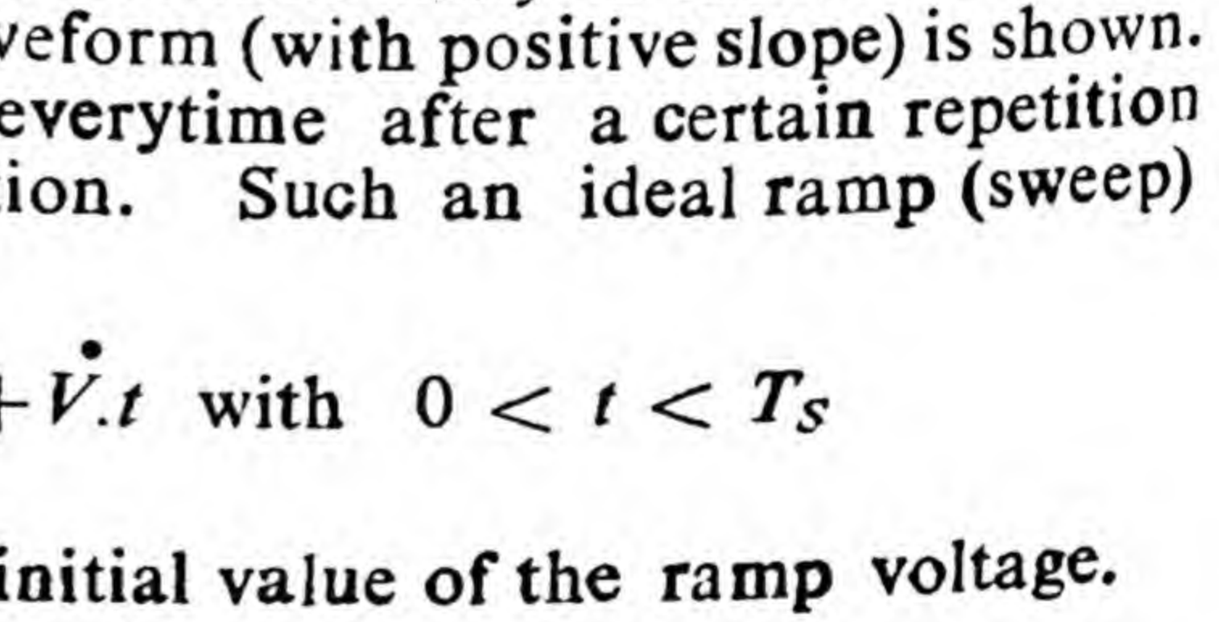


Fig. 1 (b).

In fig. 2, practical sweep voltage waveform is shown. In it we note that voltage *does not return abruptly* to its initial value after the end of sweep duration, T_S. Thus it takes some more

time after the end of sweep duration, T_S, to come to its initial value. This is called *retrace interval* or *flyback period* or *restoration time period*, and is denoted by T_r. It is desirable that retrace time be as short as possible. If fig. 3, another sweep waveform is shown in which the next ramp takes some time duration even after coming to initial value *i.e.* after the retrace time. This is called *reset time* or *rest period* and is denoted by T_{reset}. Thus for a practical sweep waveform, the repetition period is

$$T_{rep.} = T_S + T_r + T_{reset}.$$

For ideal waveform T_r and T_{reset} are zero. We may mention here that the time base generators which produce a sweep waveform without any trigger signal are called *free running time base generators*, whereas time base generators in which start of sweep or its form is governed by external signal are called *triggered time base generators*.

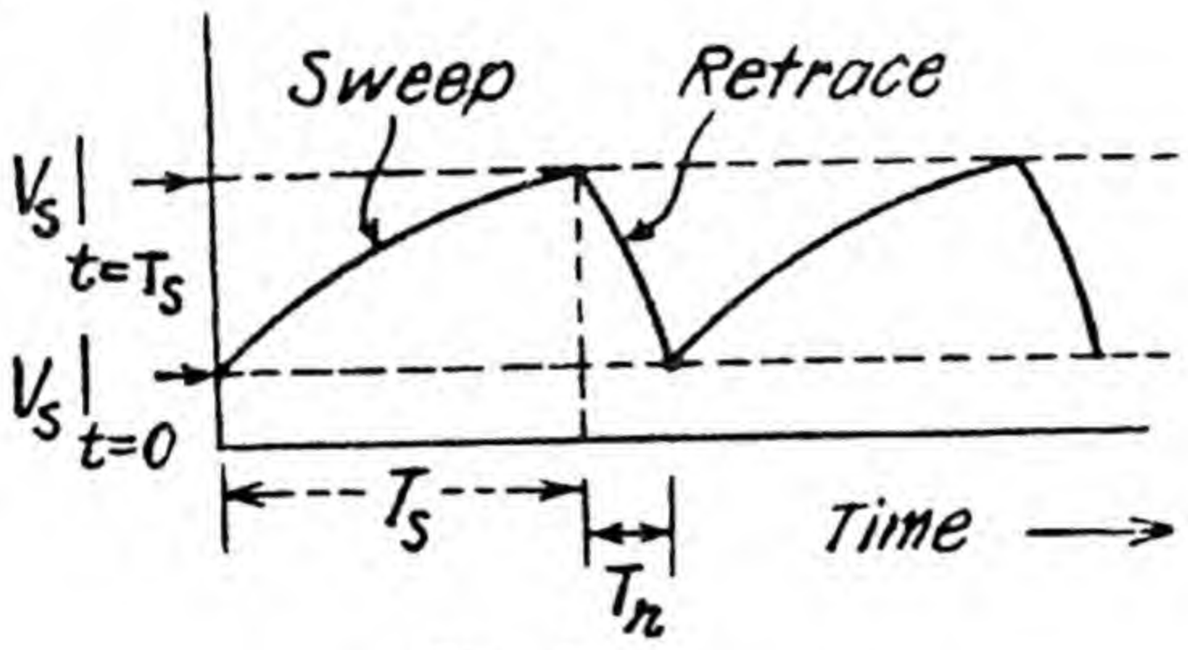


Fig. 2.

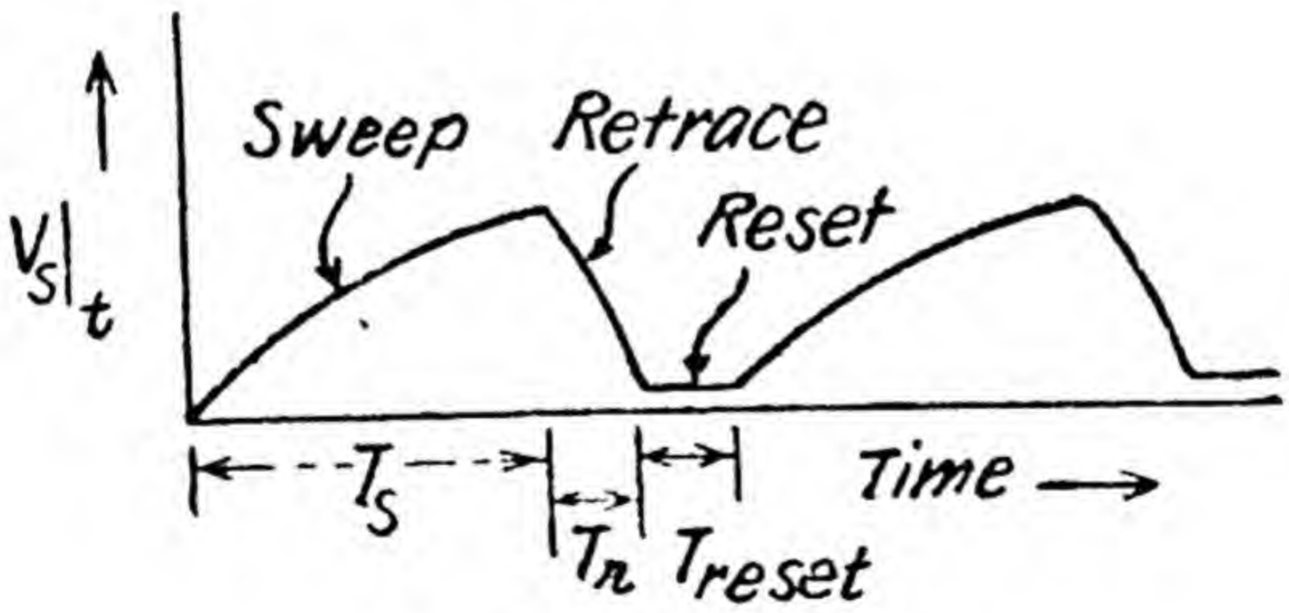


Fig. 3.

Sweep amplitude : It is defined as the modulus of the difference of ramp voltage at $t = Ts$ and the initial value of the voltage

$$V_{sweep} = \left| V_S|_{t=T_S} - V_S|_{t=0} \right|$$

Sweep or Ramp Speed : It is the rate at which the sweep voltage changes with time. It is constant for an ideal ramp waveform.

Slope or Sweep speed error : It is defined as

$$e_{speed} = \frac{\text{Initial sweep speed} - \text{Final sweep speed}}{\text{Initial sweep speed}}$$

It gives an idea of the deviation of sweep waveform from an ideal linear one.

Displacement error : This error also accounts for the departure of actual ramp waveform from an ideal perfectly linear ramp waveform. This deviation from linearity is accounted in two ways :

(a) It is expressed interms of the difference, δ, between the straight line tangent at the initial point of sweep and the actual sweep as shown in fig. 4 (*a*). The displacement error is then defined as

$$e_{disp.} = \frac{\delta_{max.}}{V_{sweep.}}$$

This deviation δ will be maximum for the tangents drawn at initial and final sweep points :

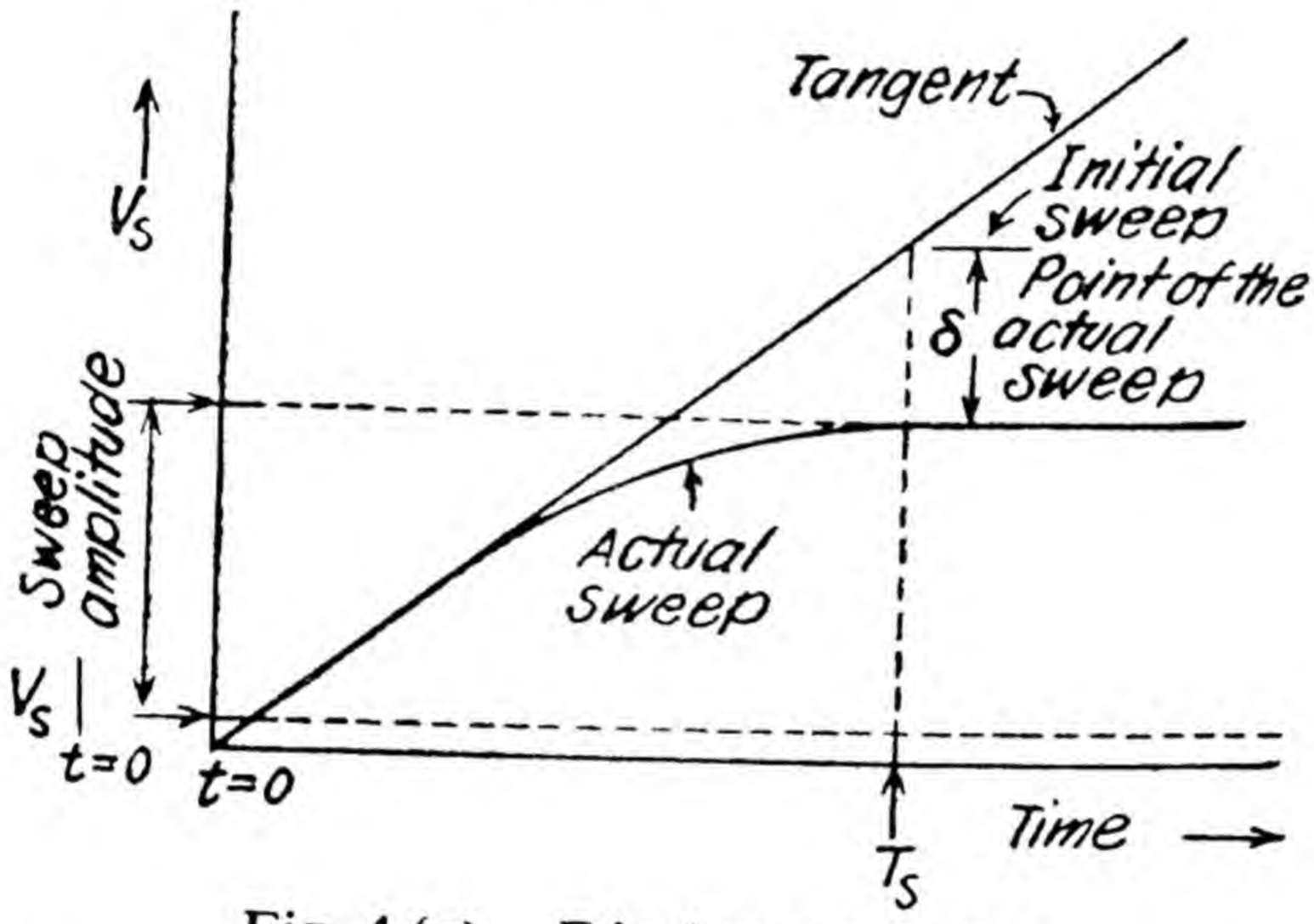


Fig. 4 (*a*). Displacement error.

(b) It is also expressed as the maximum of the difference between actual sweep and linear ramp passing through the initial and final sweep points fig. 4 (*b*). The displacement error is then defined as

$$e_{disp.} = \frac{[V_S - V_{S,\,lin.}]_{max.}}{V_{sweep.}}$$

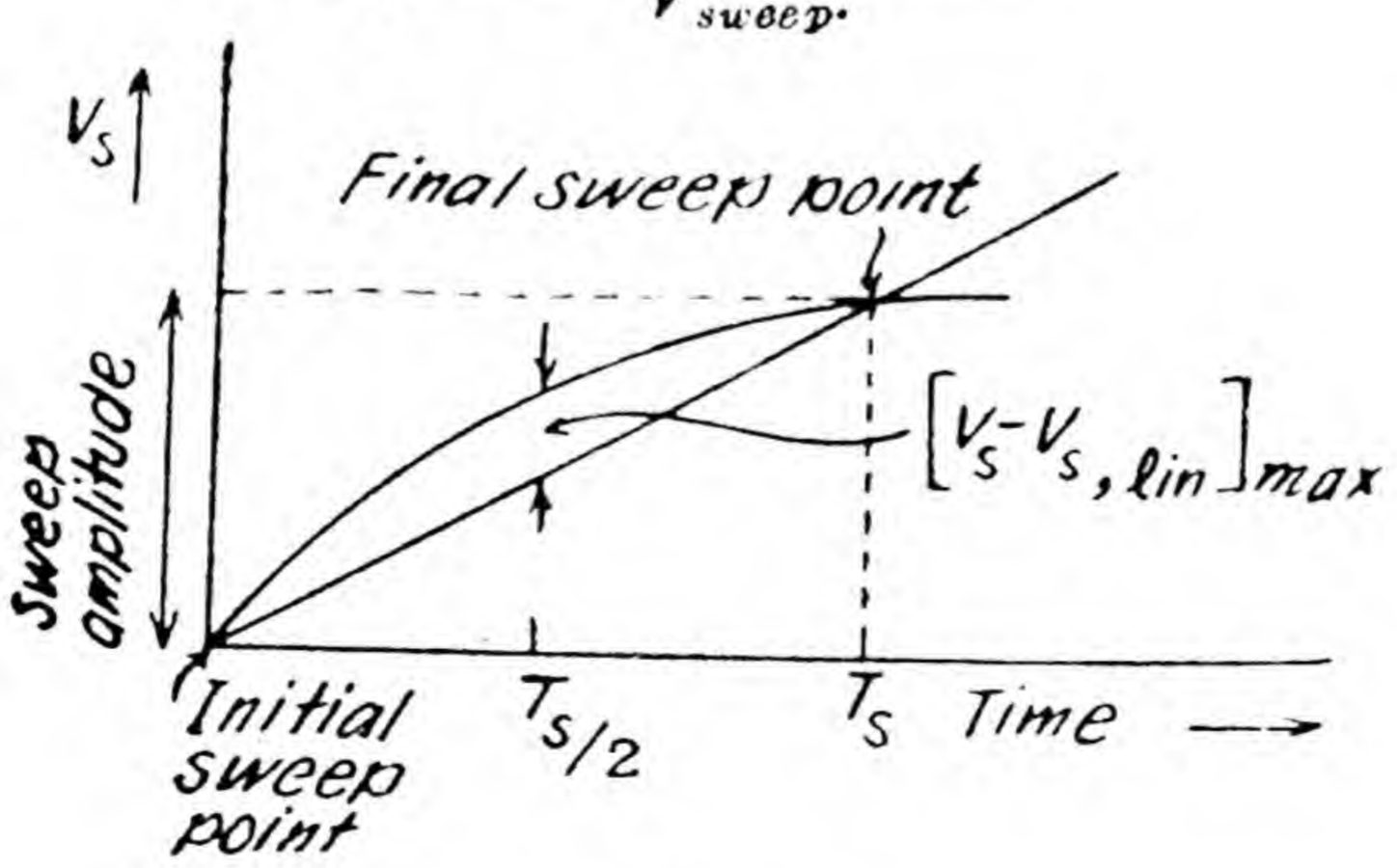


Fig. 4 (*b*).

Differential linearity : It is measured interms of the maximum deviation in rate of change of sweep voltage with time *i.e.* $\Delta\left(\frac{dV_S}{dt}\right)_{max.}$ from the average value defined over sweep duration *i.e.* $\left(\frac{dV_S}{dt}\right)_{av.}$. Thus

$$\text{Differential linearity} = \pm \frac{\Delta\left(\frac{dV_S}{dt}\right)_{max.}}{\left(\frac{dV_S}{dt}\right)_{av.}}$$

Average value can be found by

$$\left(\frac{dV_S}{dt}\right)_{av.} = \frac{1}{T_S}\int_0^{T_S}\left(\frac{dV_S}{dt}\right)dt.$$

Refer to fig. 5. The top part shows the actual sweep waveform and at the bottom is plotted rate at which sweep voltage is changing with time, (slope). Obviously at $T = T_S$ sweep voltage has become steady so that (dV_S/dt) is zero, while at the start, sweep voltage increases rapidly and (dV_S/dt) is maximum.

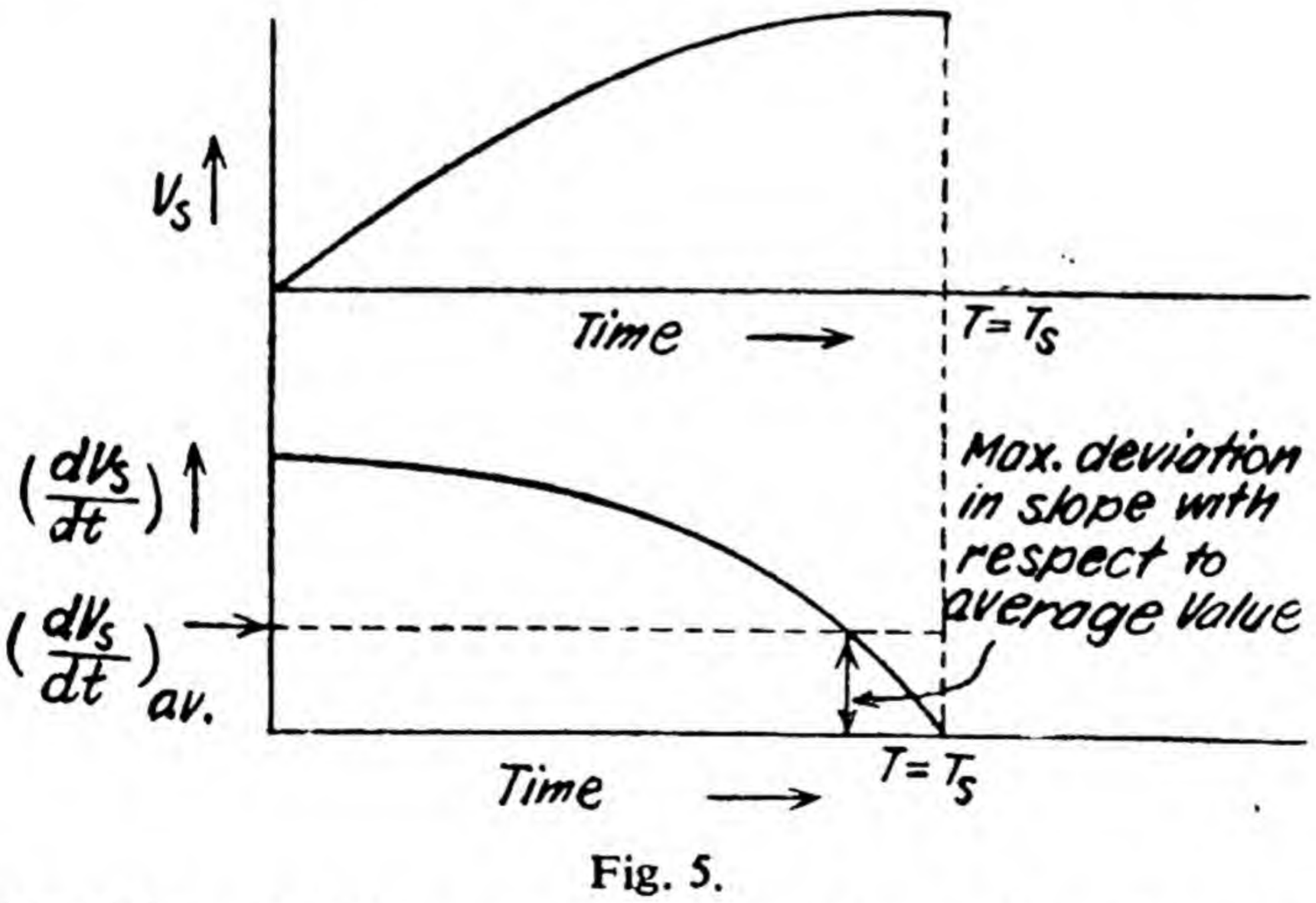


Fig. 5.

By differential linearity, we can account for the departure of voltage from linearty in a more better way.

Transmission error : Even if an ideal sweep voltage be transmitted through linear *R–C* networks, output sweep waveform departs from input sweep waveform. This departure arises due to the error that has crept in due to transmission through the network. This is called transmission error defined by

$$e_{trans.} = \frac{\text{Input sweep amplitude} \quad \text{Output sweep amplitude}}{\text{Input sweep amplitude}}$$

22·2. FUNDAMENTAL SWEEP-VOLTAGE GENERATORS :

Any method for generating sweep voltages utilizes the exponential charging and abrupt discharging of a capacitor through a voltage operated switch (*e g.* a UJT, neon bulb or transistor). Thus a fundamental sweep circuit should include : means of charging a capacitor and of discharging it when sweep voltage attains its amplitude at $t=T_S$. Further we require, in addition, circuit techniques to linearize the sweep waveform and a voltage controlled switch.

An exponential sweep circuit (*R–C* network with ideal voltage controlled switch) is shown in fig. 6. we assume that initially switch (*SW*) is closed, so $V_{out}=0$ At instant $t=0$, switch is open and the capacitor, *C*, begins to charge towards $+V_{cc}$ through resistor, *R*. The voltage across the capacitor *i.e.* sweep voltage, at any instant of time is given by

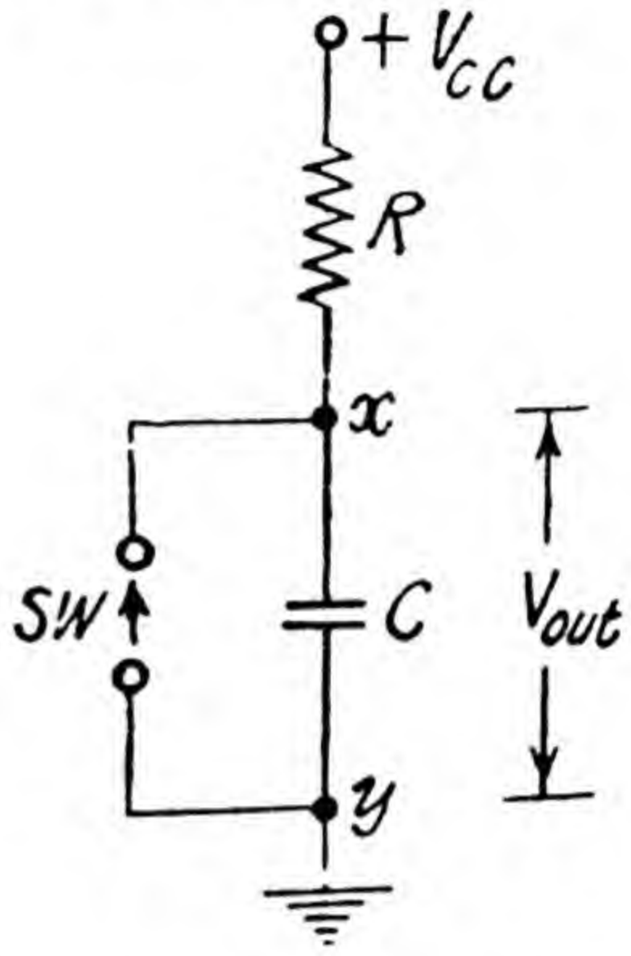


Fig. 6. An exponential sweep circuit.

$$V_{out}=V_S\Big|_t = V_{cc}\,(1-e^{-t/RC}) \quad (0<t<T_S) \qquad ...(1)$$

where *RC* is called time constant of the circuit.

Though the capacitor charge curve is extermely nonlinear, yet initial portion is almost linear. Say for first tenth of time constant (*i.e.*, $t=0{\cdot}1RC$), V_{out} is $0{\cdot}1\ V_{oo}$. This can be seen as follows: The initial charging rate of the capacitor is

$$\frac{dV_{out}}{dt}=\frac{I_{in}}{C}=\frac{V_{oc}/R}{C}=\frac{V_{cc}}{RC}$$

so that, if it continues charging at this rate, then at $t=0{\cdot}1RC$,

$$V_{out}=t\times\frac{V_{oc}}{RC}=0\ 1\ RC\times\frac{V_{cc}}{RC}=0{\cdot}1V_{cc}$$

This is the value at linear charging. But according to eq. (1), the exponential charging value will be

$$\begin{aligned}V_{out}&=V_{cc}\,(1-e^{-0{\cdot}1RC/RC})\\&=V_{cc}\,(1-e^{-0{\cdot}1}).\\&=0{\cdot}095\ V_{cc}\end{aligned}$$

which is the actual value as capacitor charges exponentially. Thus the two values are quite near, having a difference of $0{\cdot}005\ V_{cc}$ only (fig. 7). After $t=0\ 1\ RC$, the charge curve begins to depart radically from the linear curve.

Thus if the switch is left open for a time less than $0{\cdot}1\ RC$, the capacitor voltage, V_{out}, will be linear. When the switch is closed, V_{out} will abruptly drop to zero.

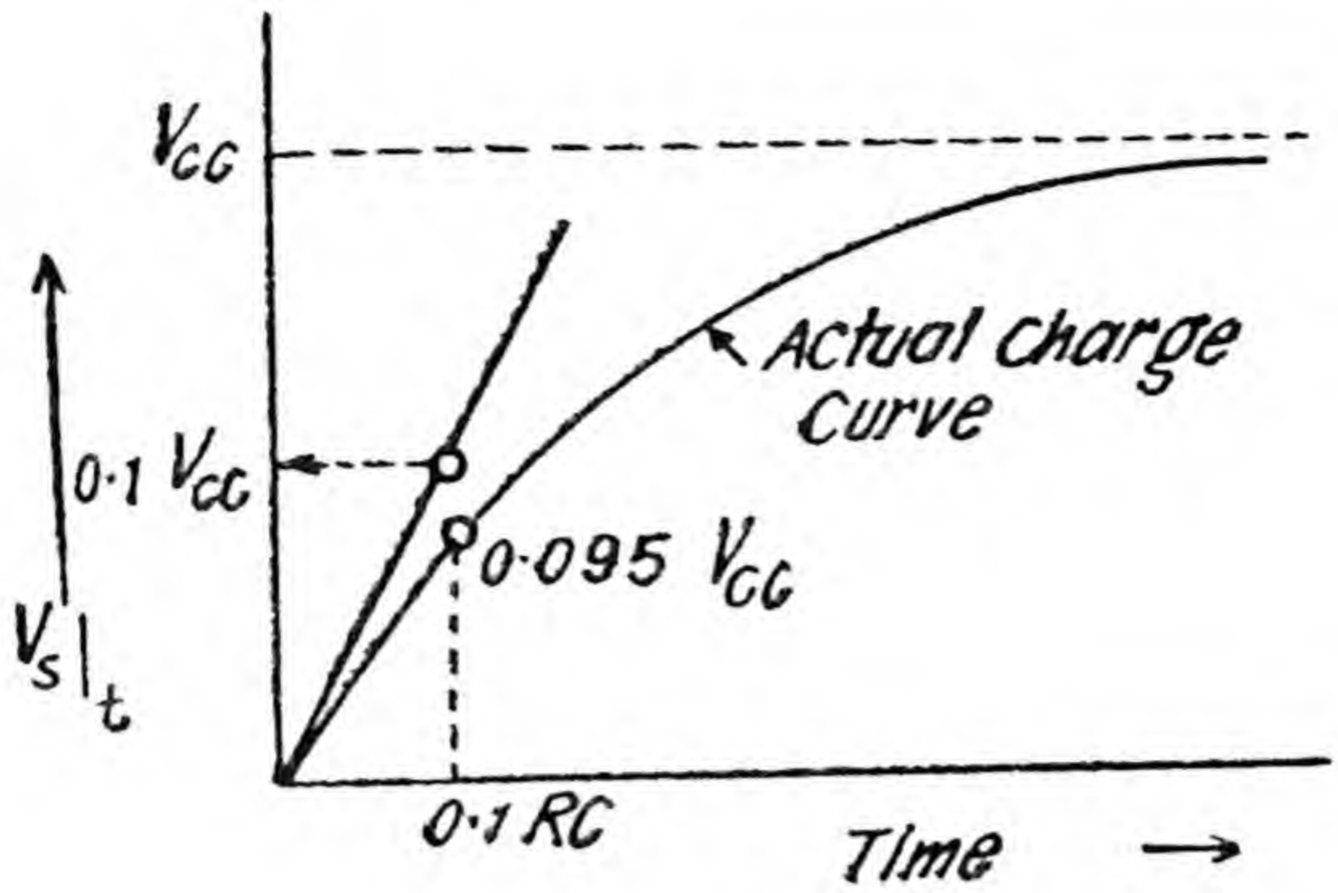


Fig. 7.

Thus sweep voltage is linear if the switch is alternately opened for 0·1 *RC* and closed for 0·1 *RC*. The sweep amplitude is 0·1 V_∞. Sweep waveform is shown in fig. 8. The sweep time is 0·1 *RC*, the fly back time is instantaneous and the OFF time is 0·1 *RC*.

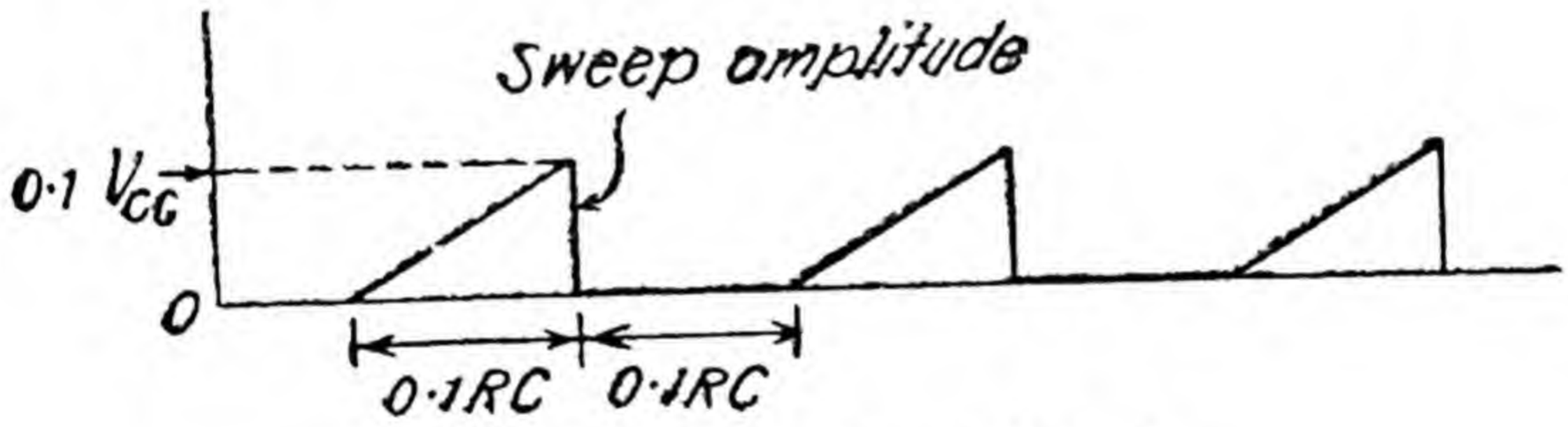


Fig. 8.

$$T_S = 0{\cdot}1RC$$

$$T_r \approx 0$$

$$E_{reset} = 0{\cdot}1\ RC.$$

Now we proceed to find errors.

(i) Sweep speed error : We have defined it earlier as

$$e_{speed} = \frac{\text{Initial sweep speed} - \text{Final sweep speed}}{\text{Initial sweep speed}}$$

$$= \frac{\left.\frac{dV_S}{dt}\right|_{t=0} - \left.\frac{dV_S}{dt}\right|_{t=T_S}}{\left.\frac{dV_S}{dt}\right|_{t=0}}$$

To find these terms, we have to differentiate eq. (1). That is,

$$\frac{dV_S}{dt} = \frac{V_{cc}}{RC} e^{-t/RC}$$

so that

$$\left.\frac{dV_S}{dt}\right|_{t=0} = \frac{V_{oo}}{RC} \qquad \ldots(2a)$$

and $$\left.\frac{dV_S}{dt}\right|_{t=T_s} = \frac{V_{cc}}{RC} e^{-T_S/RC} \qquad \ldots(2b)$$

But from eq. (1), at $t=Ts$, $V_S\big|_t = V_{sweep}$, we get

$$V_{sweep} = V_{cc}\ (1-e^{-T_S/RC})$$

or $$e^{-T_S/RC} = 1-\frac{V_{sweep}}{V_{co}}$$

Putting it in eq. (2*b*), we get

$$\left.\frac{dV_S}{dt}\right|_{t=T_S} = \frac{V_{cc}}{RC}\left(1-\frac{V_{sweep}}{V_{cc}}\right) \qquad \ldots(3)$$

From eqs. (2*a*) and (3), we arrive at

$$e_{speed} = \frac{\frac{V_{cc}}{RC} - \left(1-\frac{V_{sweep}}{V_{cc}}\right)\frac{V_{co}}{RC}}{\frac{V_{cc}}{RC}}$$

$$= \frac{V_{sweep}}{V_{cc}} \qquad \ldots(4)$$

Which predicts that, for error to be small, supply voltage should be large and sweep amplitude should be small. If sweep speed error is less, more linear will be the sweep or ramp waveform. For an ideal sweep waveform, sweep speed is constant and sweep speed error is zero. Let us see how much time constant *RC* compared to sweep duration, *Ts*, be kept so that we have a fairly linear sweep waveform. From eq. (1)

$$V_S = V_{cc}\ (1-e^{-t/RC})$$

$$= V_{cc}\left[1-\left\{1-\frac{t}{RC}+\frac{t^2}{21R^2C^2}-\frac{t^3}{31R^3C^3}+\cdots\right\}\right]$$

$$= V_{cc}.\ \frac{t}{RC}\left[1-\frac{t}{2RC}\right] \qquad \ldots(5)$$

$$= \frac{V_{cc}.t}{RC}, \text{ neglecting second order terms.}$$

Thus at $t=T_S$, we have

$$V_{sweep} = V_S\Big|_{t=T_S} = \frac{V_{co}.T_S}{RC}$$

so that from eq. (4), we get

$$e_{speed} = \frac{V_{sweep}}{V_{cc}} = \frac{T_S}{RC} \qquad \ldots(6)$$

We have already shown that sweep is almost linear for $t=0{\cdot}1$ *RC* (one-tenth of time constant). Suppose sweep duration is $0{\cdot}1$ *RC* then

$$e_{speed} = \frac{0{\cdot}1RC}{RC} = 0{\cdot}1$$

so that from eq. (4), for sweep amplitude of 1 volt, we shall require V_{cc} of value

$$0{\cdot}1=\frac{1}{V_{cc}}$$

or $$V_{cc}=10 \text{ volt.}$$

Transmission error : From eq. (5), output sweep voltage is

$$V_S\Big|_{t=T_S}=\left(V_{sweep}\right)_{out}=\frac{V_{cc}\,T_S}{RC}\left(1-\frac{T_S}{2RC}\right)$$

Sweep amplitude for linear sweep is given by

$$(V_{sweep})_{lin}=\frac{V_{cc}\,T_S}{RC}$$

Therefore transmission error is given by

$$e_{trans.}=\frac{(V_{sweep})_{lin}-(V_{sweep})_{out}}{(V_{sweep})_{lin}}$$

$$=\frac{\dfrac{V_{cc}\,T_S}{RC}-\dfrac{V_{cc}\,T_S}{RC}\left(1-\dfrac{T_S}{2RC}\right)}{\dfrac{V_{cc}\,T_S}{RC}}$$

$$=\frac{T_S}{2RC}=\frac{e_{sweep}}{2}$$

Therefore, if time constant is high, transmission error will be minimum and waveform will be nearly linear. And for a sweep to be transmitted through high pass circuit without distortion, a very high value of capacitor, C, is needed.

Displacement error : Refer to fig. 4 (b). The maximum deviation from linear waveform occurs at $t=T_S/2$. So that

$$(\triangle V_S)_{max}=(V_S)_{t=T_S/2,\text{ actual}}-(V_S)_{t=T/2,\text{ linear}}$$

$$=V_{cc}\frac{T_S}{2RC}\left(1-\frac{T_S}{4RC}\right)-\left(\frac{V_{cc}\,T_S}{2RC}\right)$$

$$=\frac{V_{cc}}{8}\left(\frac{T_S}{RC}\right)^2$$

so that

$$e_{disp.}=\frac{(\triangle V_S)_{max}}{V_{sweep}}=\frac{\dfrac{V_{cc}}{8}\left(\dfrac{T_S}{RC}\right)^2}{V_{cc}\dfrac{T_S}{RC}}=\frac{T_S}{8RC}=\frac{e_{trans.}}{4}$$

Sweep circuits : Now we shall list here the type of the circuits used to generate a linear sweep voltage. They are

(*i*) Constant current charging methods

(*ii*) A negative feed back technique to minimise the sweep

speed error by enhancing supply voltage V_{cc} and the time constant RC. (Miller circuit)

(*iii*) A positive feed back technique (Boot strap method) in which the error is first detected and then charging current amplitude is controlled by a suitable positive voltage feedback.

(*iv*) Integration of an asymmetrical square wave.

22·3. TRANSISTOR CONSTANT CURRENT SWEEP GENERATOR :

We have shown in the last section that a generator, utilising a capacitor charging through a resistor, produces a sweep waveform with poor linearity if more than 10 percent of the charge curve ($t = 0{\cdot}1\ RC$) is used. And if 10 percent of charge curve is used, sweep amplitude is small. These limitations can be overcome if a method is used in which the capacitor is charged by a constant current because if the current charging the capacitor is *held constant during sweep* time the voltage across capacitor will rise linearly. Thus it will result in linear sweep.

A circuit using a *PNP* transistor in common base configuration is shown in fig. 9 (*a*). The transistor is biased to operate in the active region ($E-B$ junction is forward biassed). Let the base be biased at $+10V$ above ground so that $C-B$ junction is reverse biased, while $E-B$ junction is forward biased. Let the capacitor be fully discharged initially so that collector is at ground potential ($V_{CC}=0$) and $V_{CB}=-10\ V$. Further as $E-B$ junction is forward biased, it will conduct and maintain a forward voltage drop of about $0{\cdot}7V$ (silicon) so that $V_{BE}=10{\cdot}7V$. The emitter current is then

$$I_E=\frac{20V-10{\cdot}7V}{4{\cdot}7\ k\Omega}\simeq 2\ mA$$

The collector current, I_C, at the initial operation point is also approximately 2 *mA*. It will charge the capacitor positively at the rate

$$\frac{dV}{dt}=\frac{I_C}{C}=\frac{2mA}{1\mu F}=2000\ V/s \qquad (\text{if } C=1\mu F)$$

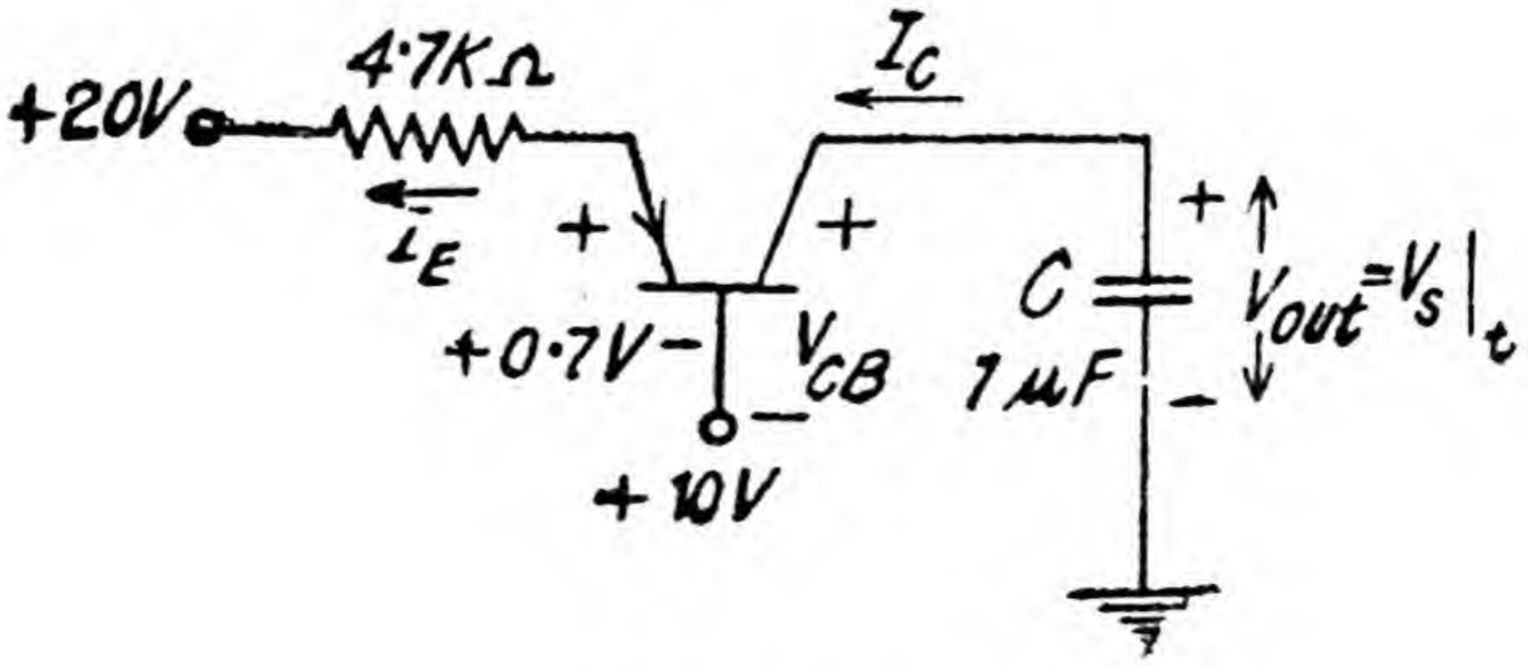


Fig. 9 (*a*)

In 4 milliseconds, it will be charged to a voltage

$$= 2000\ V/s \times 4ms$$

$$= 2000 \times 4 \times 10^{-3} = 8V$$

so that net reverse bias is now only $-10V + 8V = -2V$, *i.e.*, $V_{CB} = -2V$. Thus reverse bias V_{CB} goes on decreasing but the collector current, I_C, is still essentially constant at $2mA$ and continues to be constant upto 5 *ms* when capacitor has reached a voltage $10V$ and $V_{CB} = 0$, (fig. 9*b*). Thus *rise of voltage across the capacitor* (due to constant charging current, I_C, upto $V_{CB} = 0$) is *linear until it* (capacitor voltage) *reaches a value equal to the base bias voltage*.

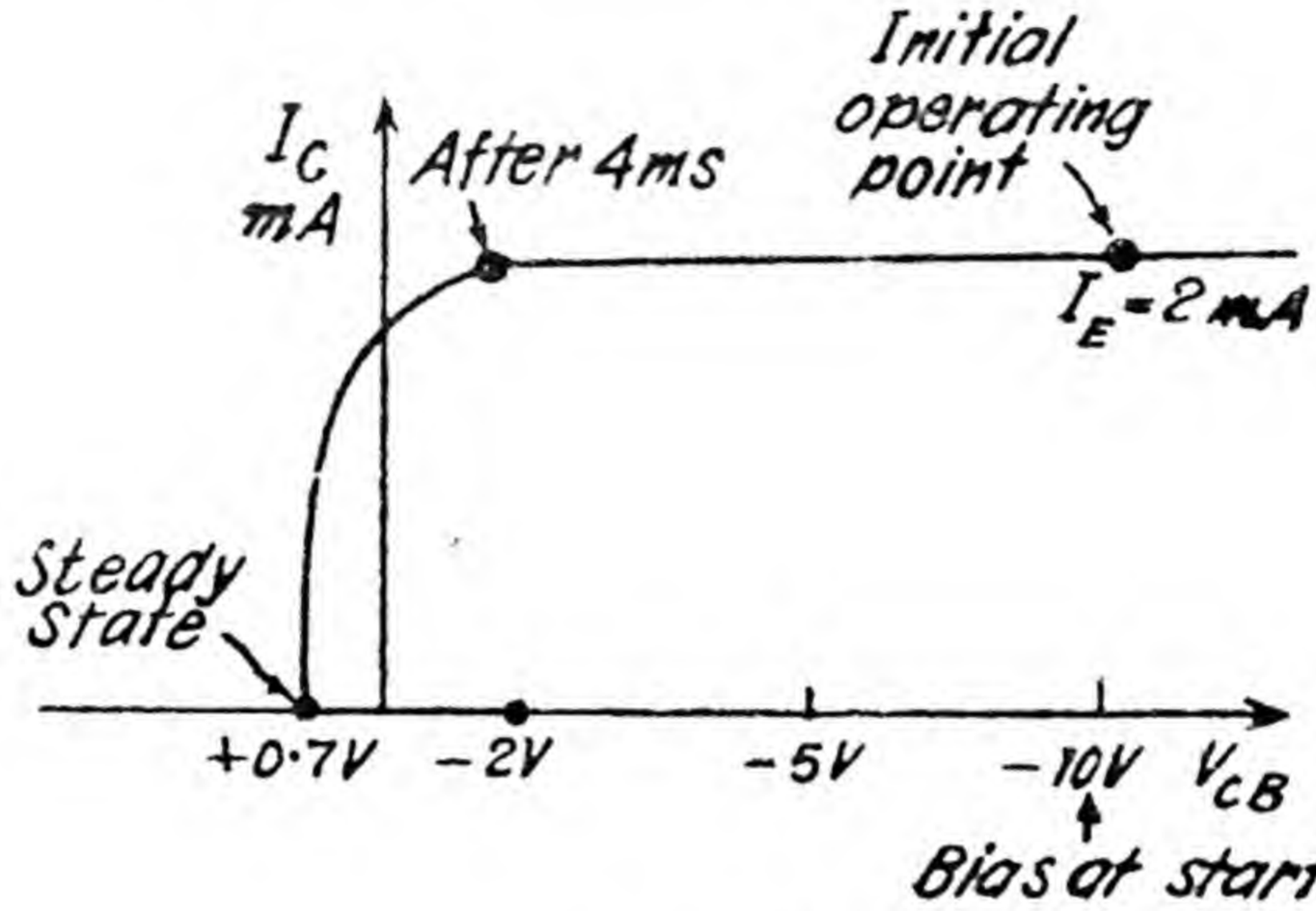


Fig. 9 (*b*)

After capacitor voltage has reached base bias voltage, $V_{CB} = 0$ and after this point, V_{CB} starts going positive, causing the co lector current to decrease rapidly. Consequently capacitor will charge at a slower rate until I_C becomes zero. When I_C reaches zero, the capacitor stops charging (steady state). Thus rise of voltage across C is no more linear after V_{CB} starts going positive. *It means if the capacitor is discharged on or before the time it reaches* $10V$, *the output sweep voltage* $V_{out}\left(= V_S\big|_t\right)$ *will be linear. That is, sweep voltage waveform obtained will be linear.*

Fig. 10, shows a method for discharging the capacitor to produce flyback with the help of a transistor switch. The input to Q_2 is normally zero so Q_2 is OFF, and allows C to be charged

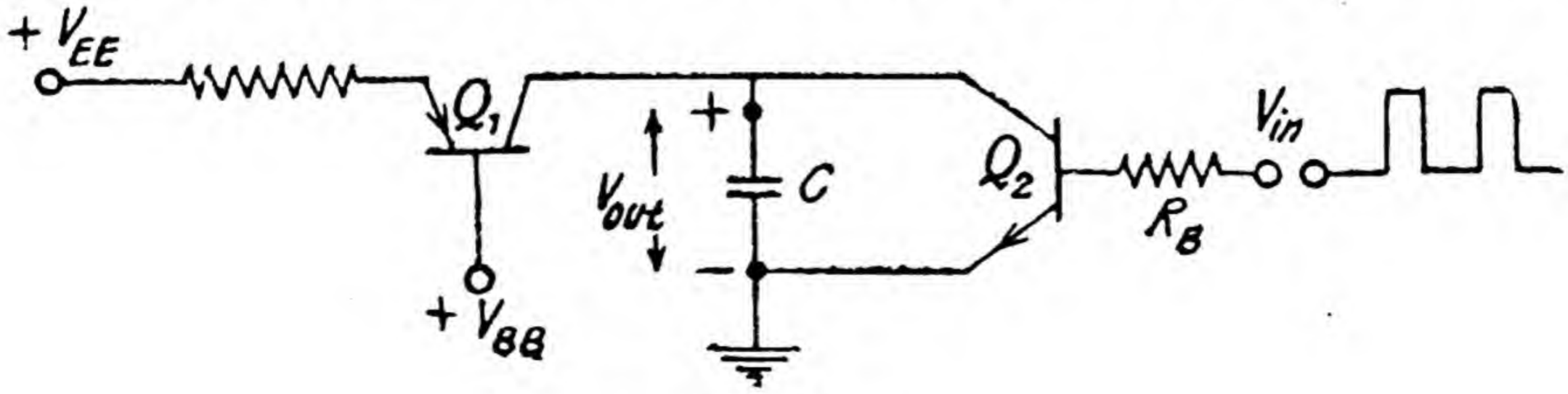


Fig. 10.

linearly by the collector current of Q_1. When Q_2 is pulsed, it turns ON and discharges C to $0V$.

22·4. UJT SWEEP WAVEFORM :

First we shall explain *UJT* operation and then production of sweep waveform.

U.IT operation : *UJT* is a three terminal device. It usually consists of an *N*-type silicon bar (fig. 11), along one side of which a *PN* junction (near mid point) is formed. Connection made to *P*-type material of the junction is called emitter *E*. Connections to the two ends of the bar B_1, and B_2 are called bases. Thus this structure has two bases and unijunction.

Between B_1 and B_2, structure is simply a resistor, total resistance is $(R_{B_1}+R_{B_2})$ where R_{B_1} is the resistance of the portion between B_1 and the point, J (cathode of *PN* junction) and R_{B_2} is the resistance of the portion between B_2 and the point J. Obviously if V_{BB} be the voltage applied between B_1 and B_2 then voltage at point J will be

$$V_J=\frac{R_{B_1}}{R_{B_1}+R_{B_2}}\times V_{BB}$$
$$=\eta V_{BB}.$$

when η is called internal voltage divider ratio and is supplied by the manufacturer. R_{B_1} is variable as it changes considerably when *UJT* is ON.

When we apply a voltage V_{BB} to the cathode then if it is less

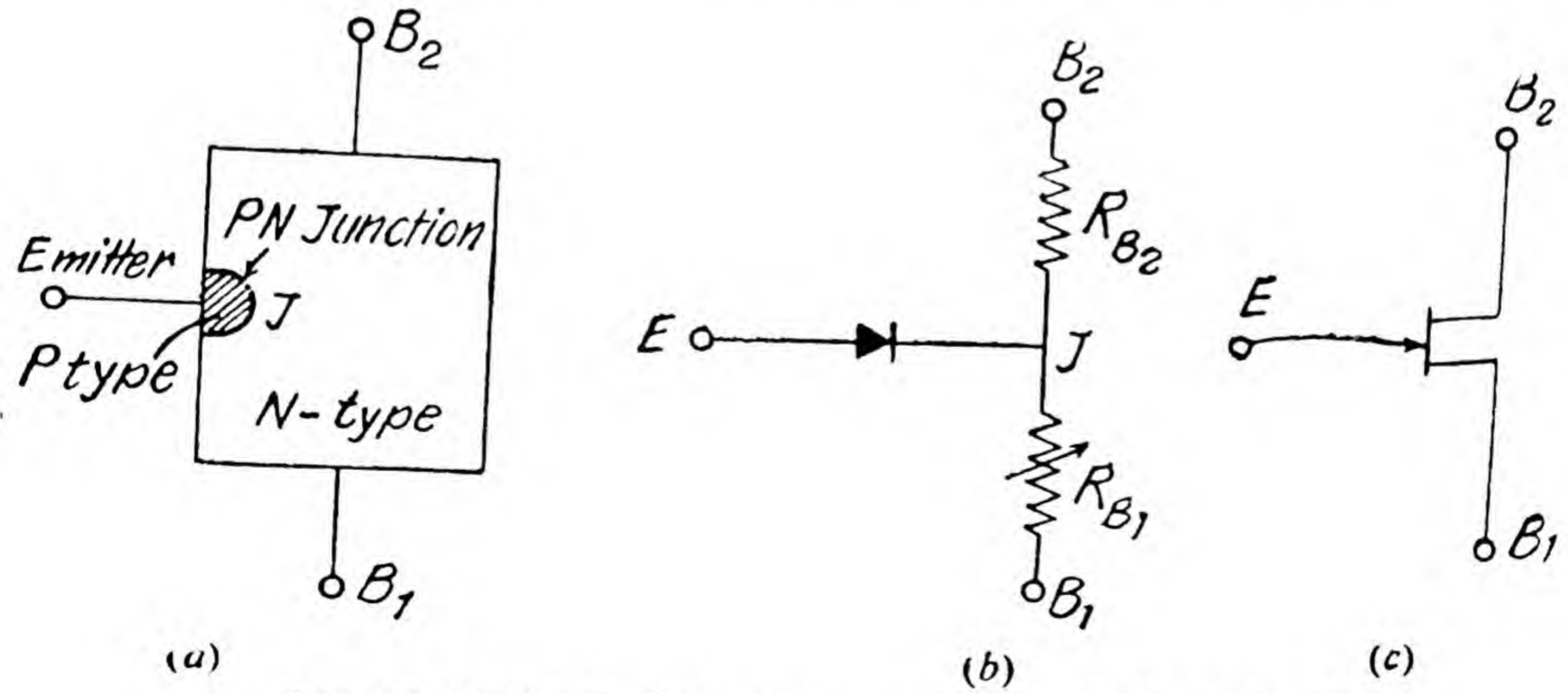


Fig. 11. *UJT* (*a*) Structure, (*b*) equivalent circuit, (*c*) electronic symbol

than V_J then, as the diode will remain reverse biased, no emitter current, I_E, flows (though current will flow between B_1 and B_2).

This is called OFF state of *UJT*. Resistance R_{B_1} is very high in this state (fig. 12).

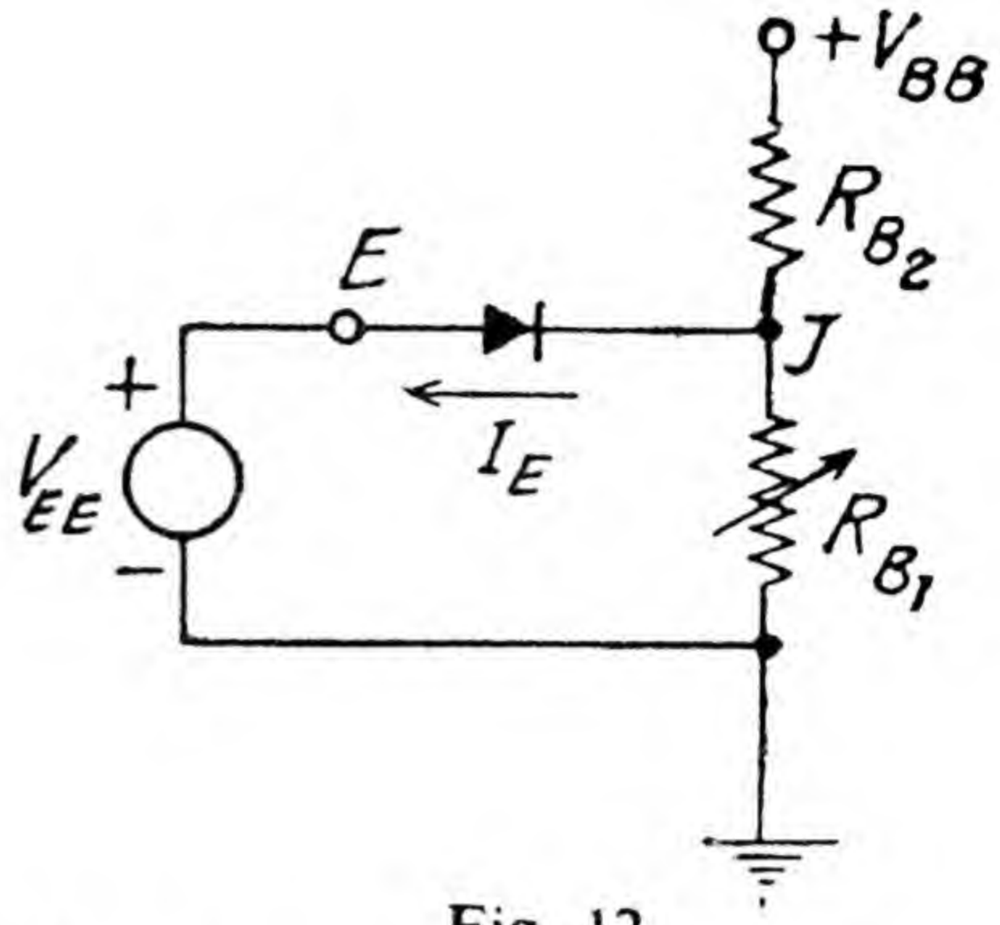


Fig. 12.

Turning **ON** *of UJT* : When emitter voltage is increased to a value such that the diode begins to turn ON, the emitter current will start flowing. Such a value of V_{EE} is called peak point voltage V_P. Any increase in emitter voltage, V_{EE}, increases I_E and when I_E reaches a value called peak point current, I_P, then *UJT* switches (atonce, within $1\mu s$) to its ON state. In ON state R_{B_1} drops to a low value.

Turning **OFF** *of UJT* : Once *UJT* is ON, V_{EE} can be reduced below V_P and *UJT* will remain ON but I_E will decrease. When I_E decreases below a value, I_V, called valley current, *UJT* will revert back to its OFF state.

UJT as Sweep Generator : Circuit is shown in fig. 13 (*a*). When a supply voltage V_{BB}, say $20V$, is initially applied to the circuit, *UJT* is momentarily OFF because capacitor grounds the emitter to 0, volt and no emitter current flows. Consequently, *capacitor C, will begin to charge towards* $20V$ *through resistor R*, (fig. 13 *b*). But as *C* charges towards $20V$, emitter voltage increases and as it reaches peak point voltage, V_P, *UJT* turns ON, and the resis'ance

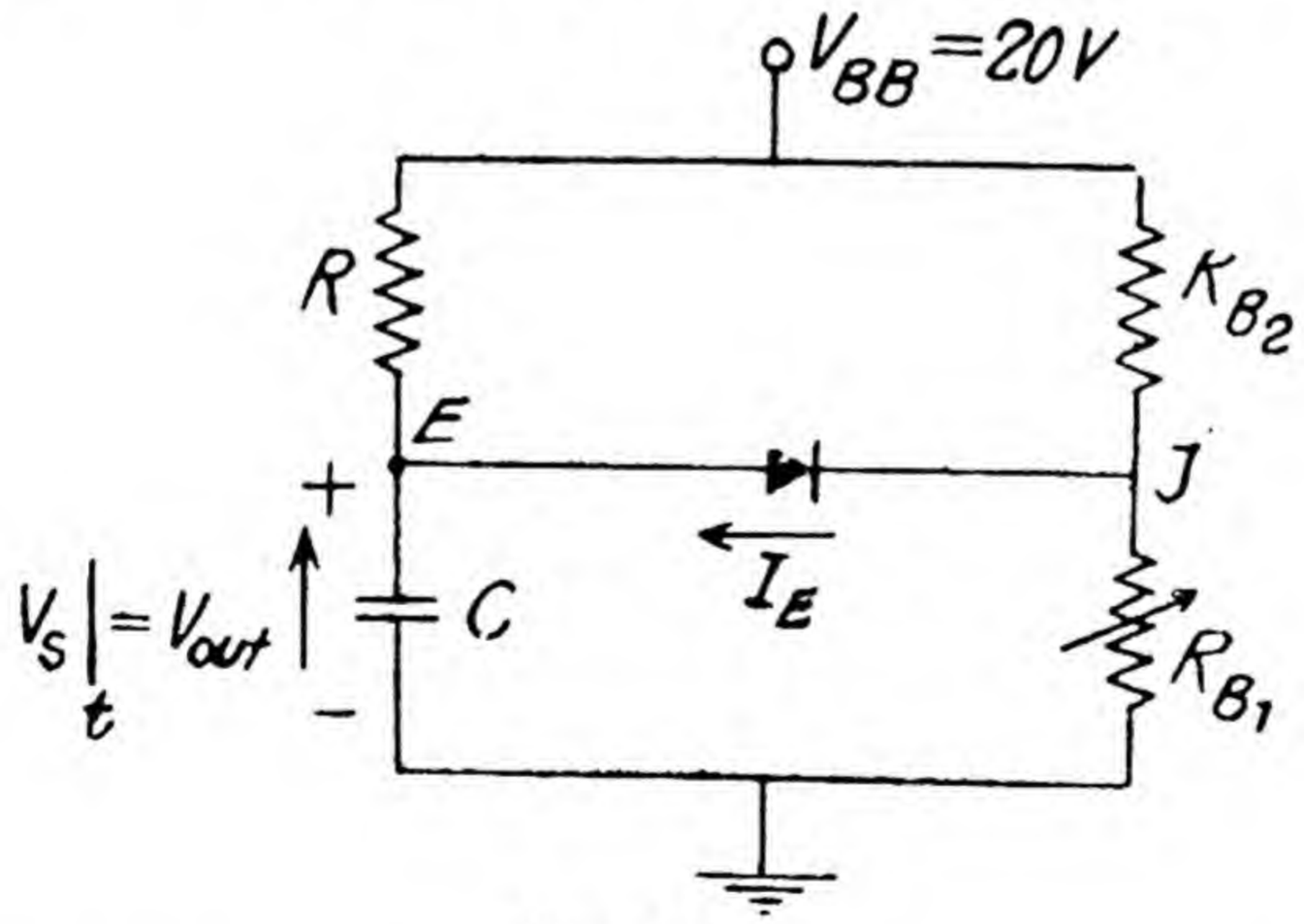


Fig. 13. (*a*).

R_{B_1}, decreases appreciably. But such a low R_{B_1} (about 50 ohm) provides a path for the condenser, *C*, to discharge. Thus *C* discharges and its discharge current provides the emitter current needed to hold *UJT*, ON. The capacitor will continue to discharge until emitter current, I_E, drops below peak point current, I_P and then *UJT* is tuned OFF. In OFF state, *C*, again starts charging towards $20V$ and the whole process is repeated.

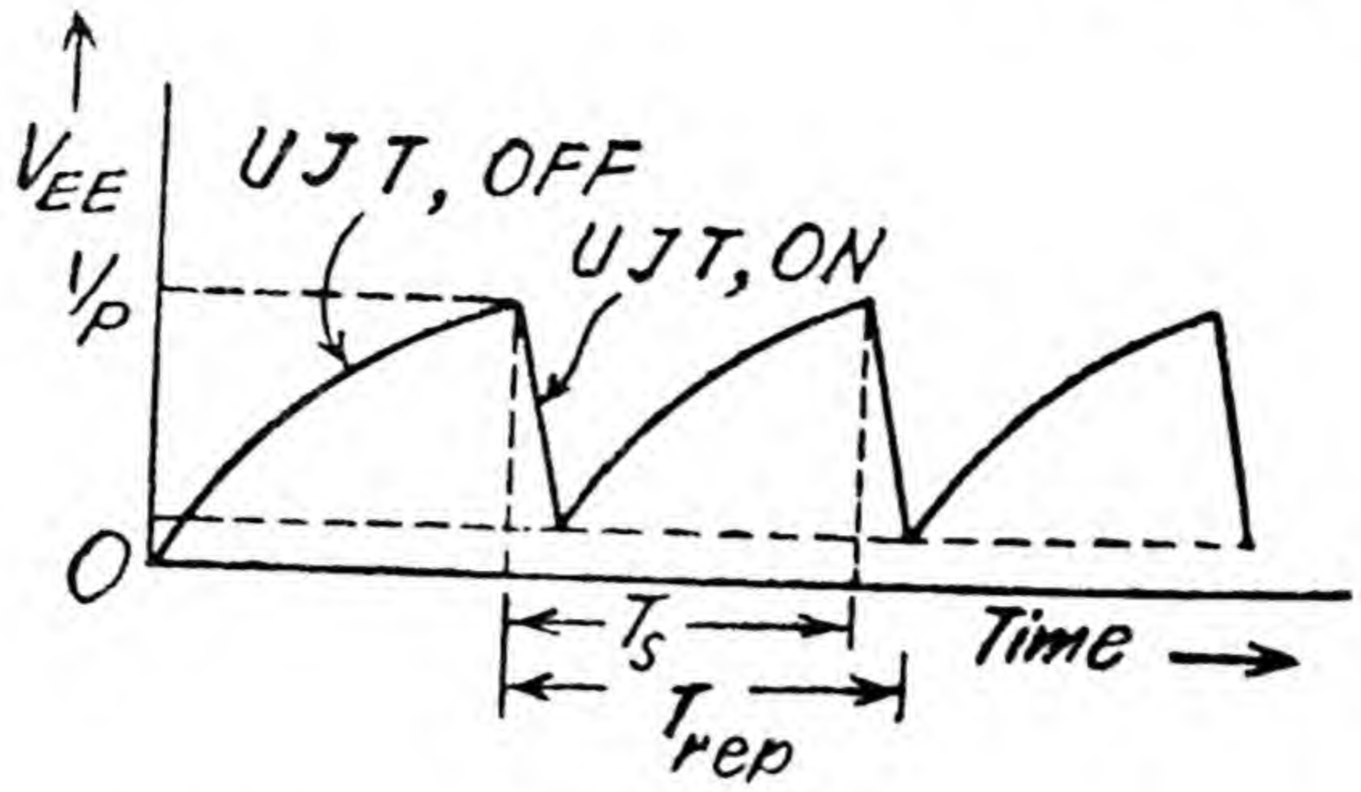


Fig. 13. (*b*)

22·5. MILLER INTEGRATOR SWEEP CIRCUIT :

By using a negative voltage or current feedback, this circuit produces a linear sweep voltage.

A basic sweep circuit is shown in fig. 14 (*a*). If switch, *S*, is open, voltage across *C*, will rise towards *V* to give sweep voltage. The sweep waveform will not be linear if the charging current is not constant. *Suppose if the voltage drop across C, be supplied by a voltage source, say, a fictitious variable voltage generator v, (fig. 14b) then charging current will remain constant at i=V/R, and consequently a linear sweep will be obtained.* If terminal, *Z*, is grounded (fig. 14*c*) then this linear sweep will appear between point *Y* and ground and will increase in the negative direction. We now replace the fictitious generator by an amplifier with output terminals *YZ* and input terminals *XZ* (fig. 14*d*). Since voltage across *C* is supplied by the generator voltage, *v*, at every instant of time, the

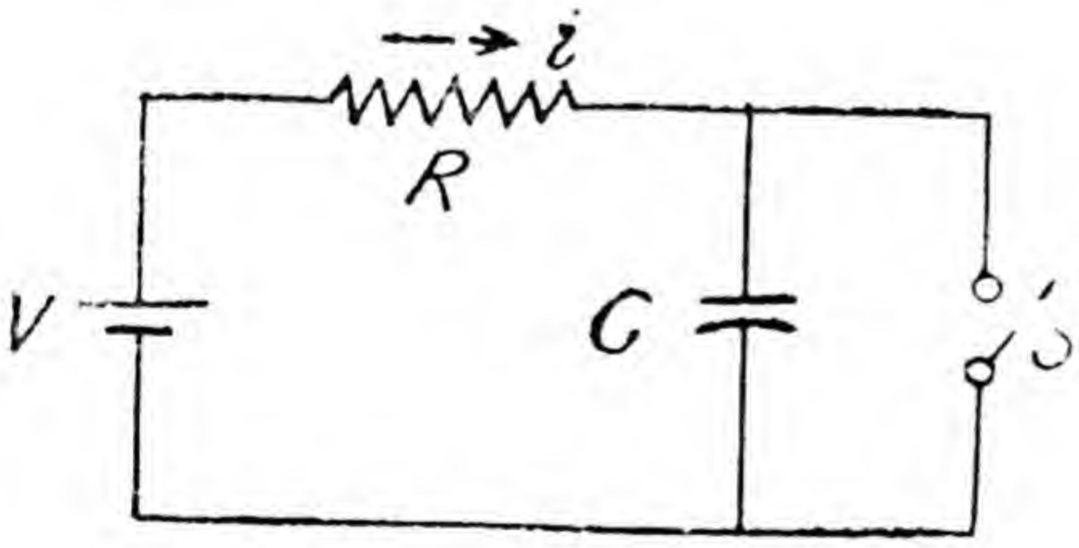


Fig. 14 (*a*)

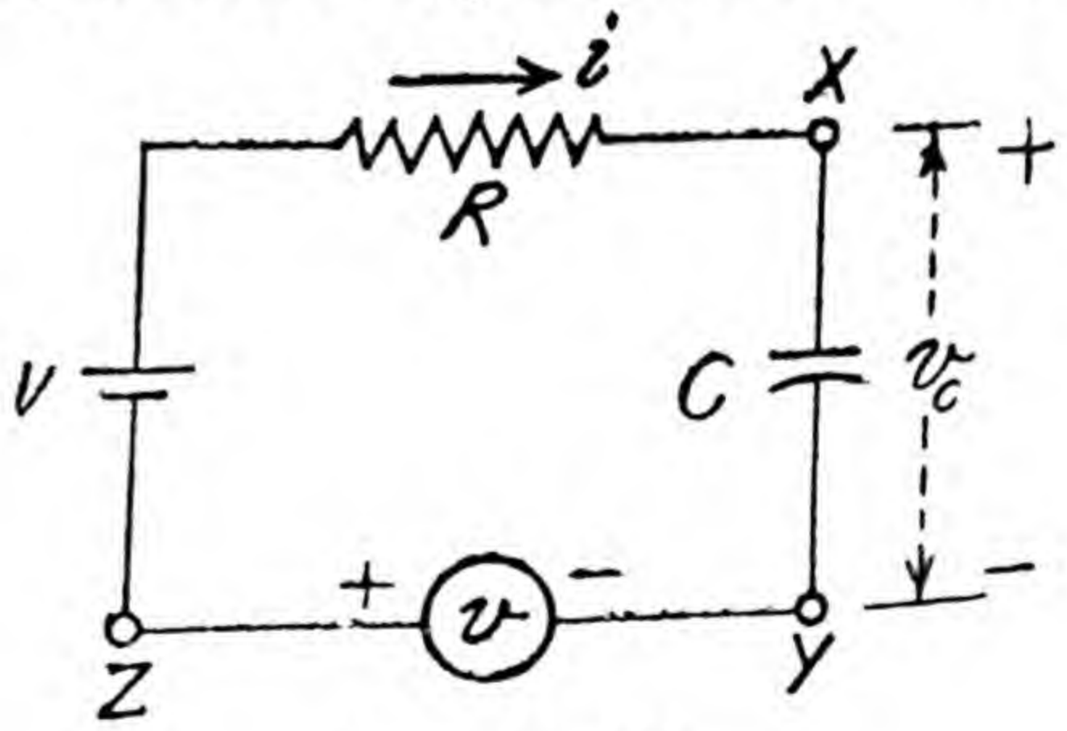


Fig. 14 (*b*)

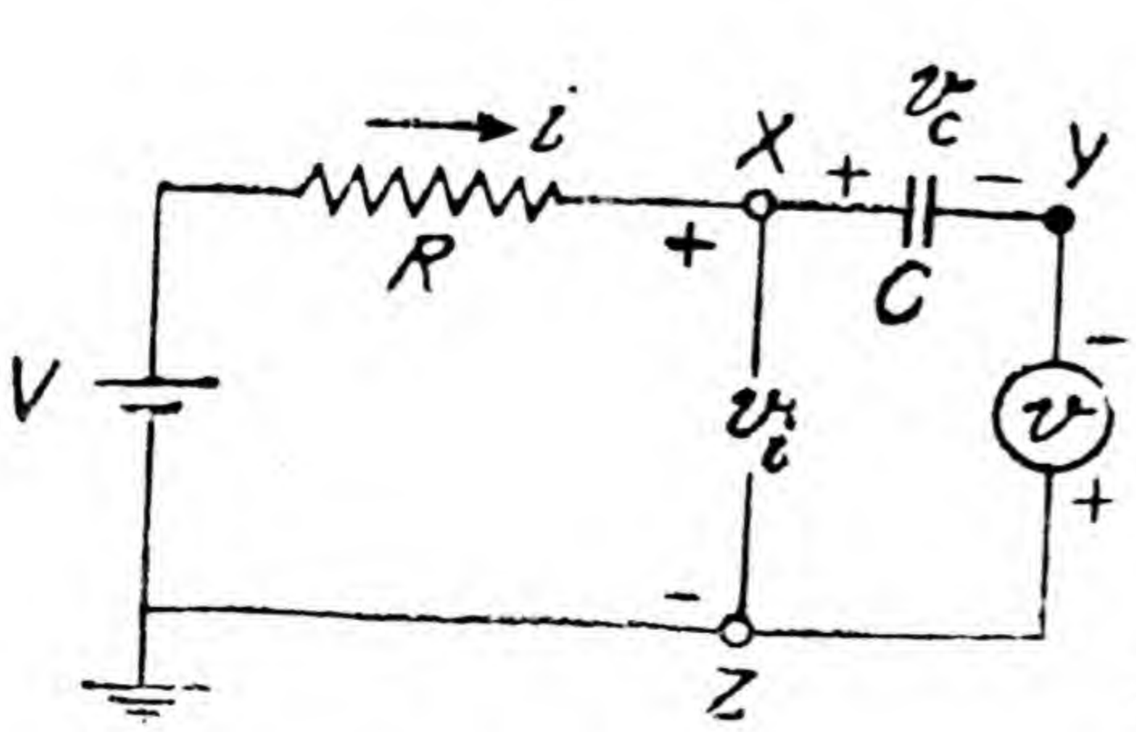


Fig. 14. (*c*)

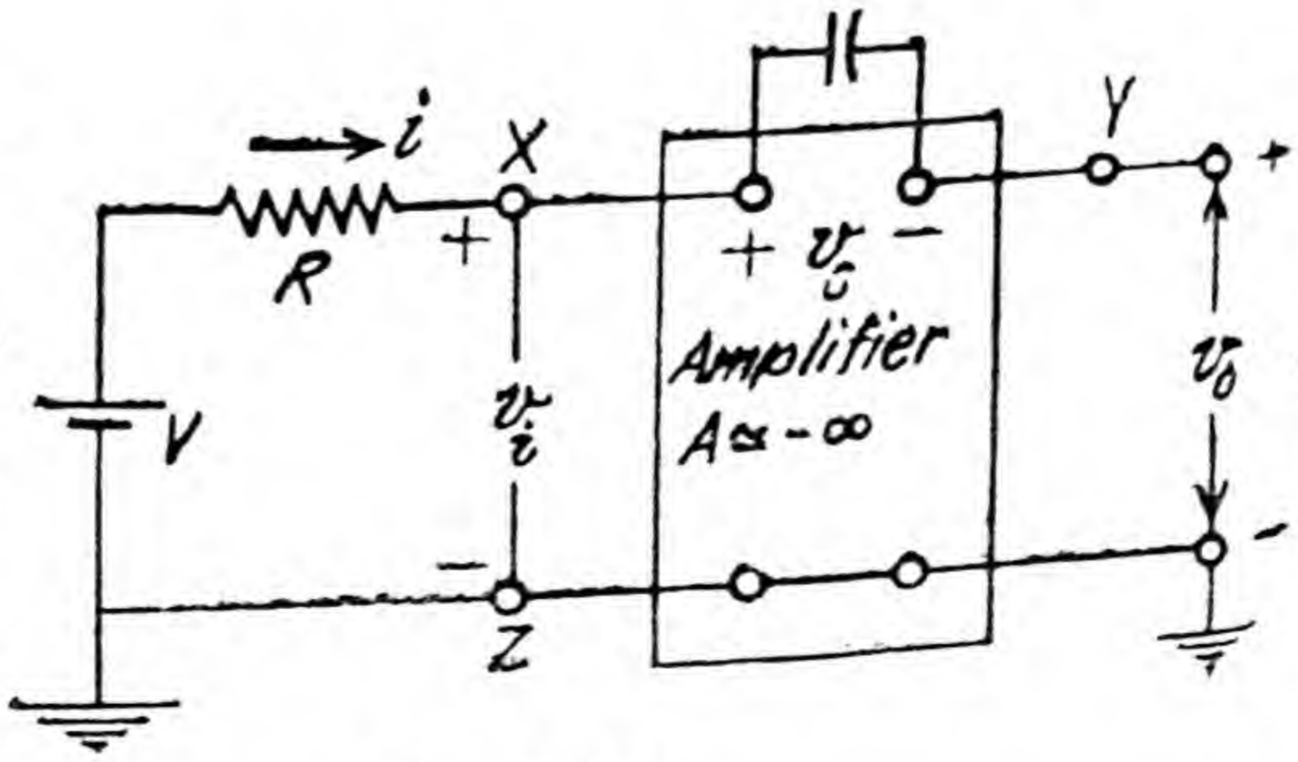


Fig. 14 (*d*)

input, v_i, to the amplifier is zero. It means X is at ground potential (like Z) *i.e.* X behaves as virtual ground and yet with zero input, if we want to obtain a finite output, the gain of the amplifier, A, should ideally be infinite (a finite output divided by zero input gives infinite gain) but it is a characteristic of an ideal operational amplifier acting in an inverting mode. If feedback resistor of such an *PO-AMP* in inverting mode be replaced by a capacitor (fig. 14*e*) then it becomes an operational integrating amplifier. Such a circuit (fig. 14*e*) is what we have described in (fig. 14*d*). The feedback through C, forces the terminal X to exist at virtual ground. This feedback capacitor, C, appears as $(1+A)C$ across the input terminals (fig. 14*f*). We may remember that in Miller effect also

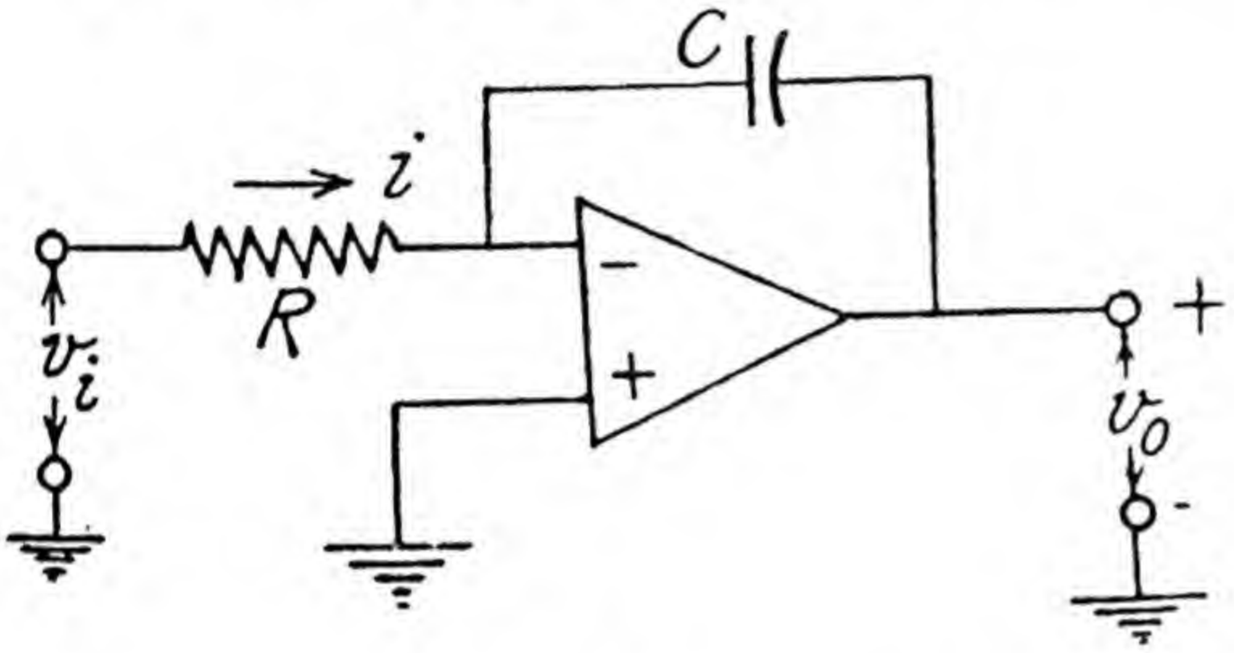


Fig. 14 (*e*)

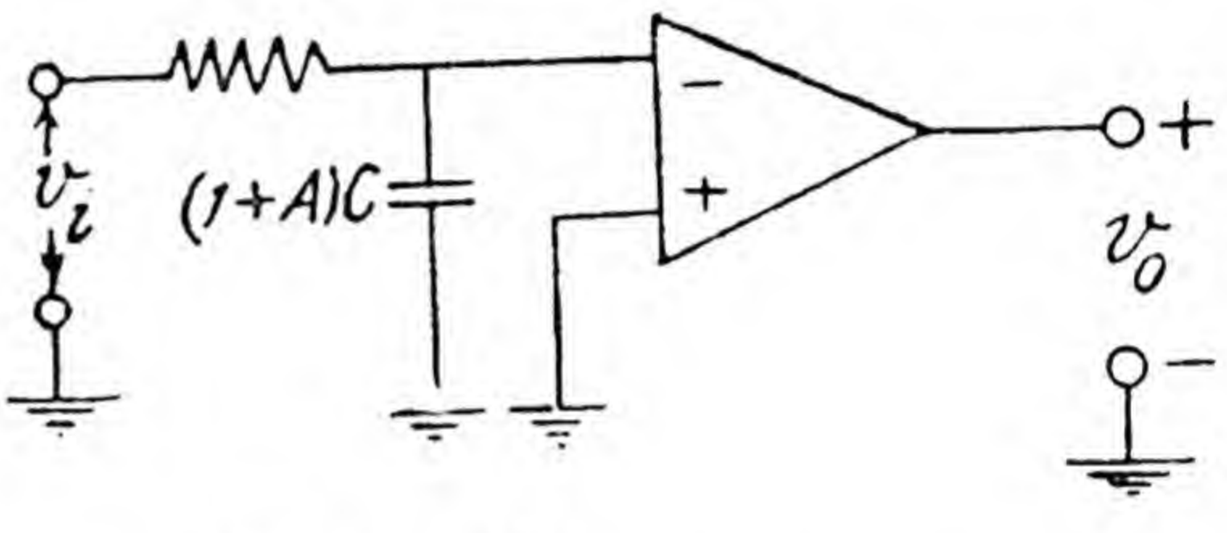


Fig. 14 (*f*)

a capacitance connected between input and output terminals (*i.e.* grid and plate circuit, *i.e.* C_{gp}) effectively raises the input capacitance by a factor $(1+A)$. Hence such an integrator is called *Miller integrator* or Miller sweep.

The Miller Sweep : Fig. (14*d*) has been redrawn in fig. (14*g*) showing a starting switch S, [note that here this switch is connected such that the sweep starts to form when we *close* it, unlike in fig. (14*a*)], amplifier input resistance (R_i) and a Thevenin's circuit for the base amplifier. We take that the output resistance of the base amplifier is negligible. Let A be the *open circuit voltage gain*

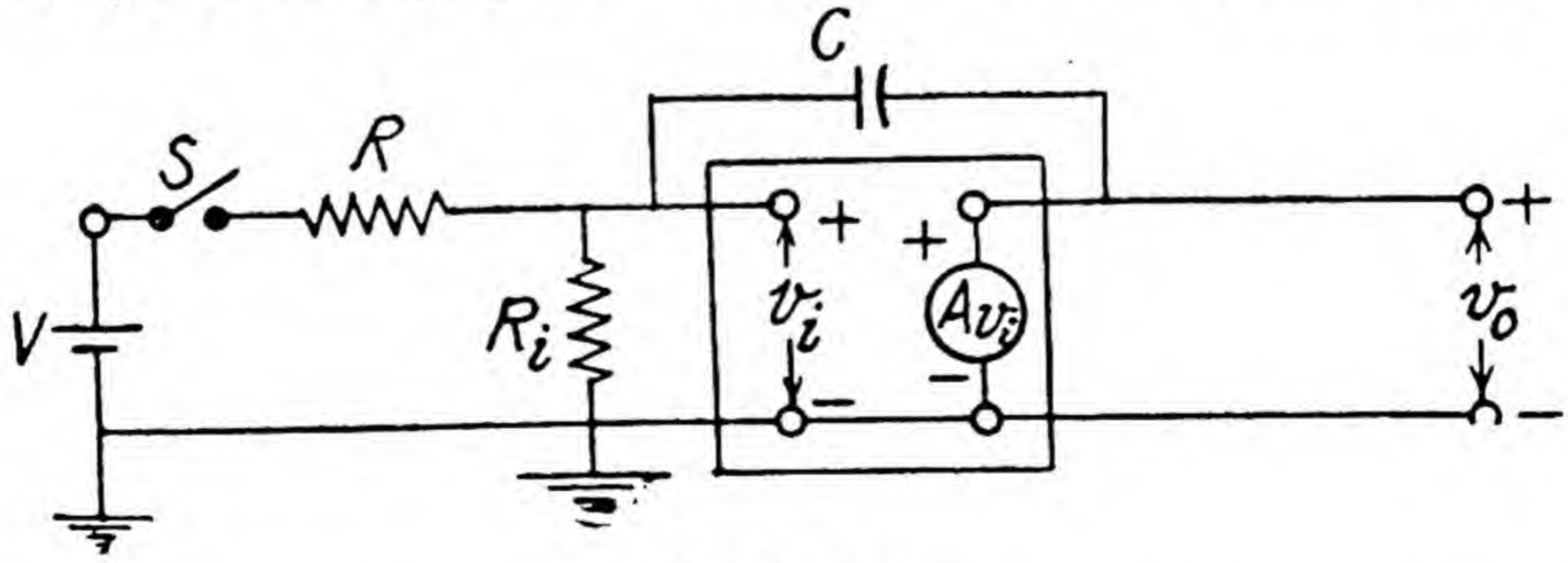


Fig. 14 (*g*)

of the base amplifier. We have further replaced the input circuit of fig. (14*g*) by a Thevenin's equivalent (fig. 14*h*) in which

$$V' = \left(\frac{V}{R_i + R}\right) \quad R_i = \frac{V}{1 + R/R_i}$$

and $$R' = \frac{R_i R}{R_i + R} \qquad \left(\text{i.e. } \frac{1}{R'} = \frac{1}{R_i} + \frac{1}{R}\right)$$

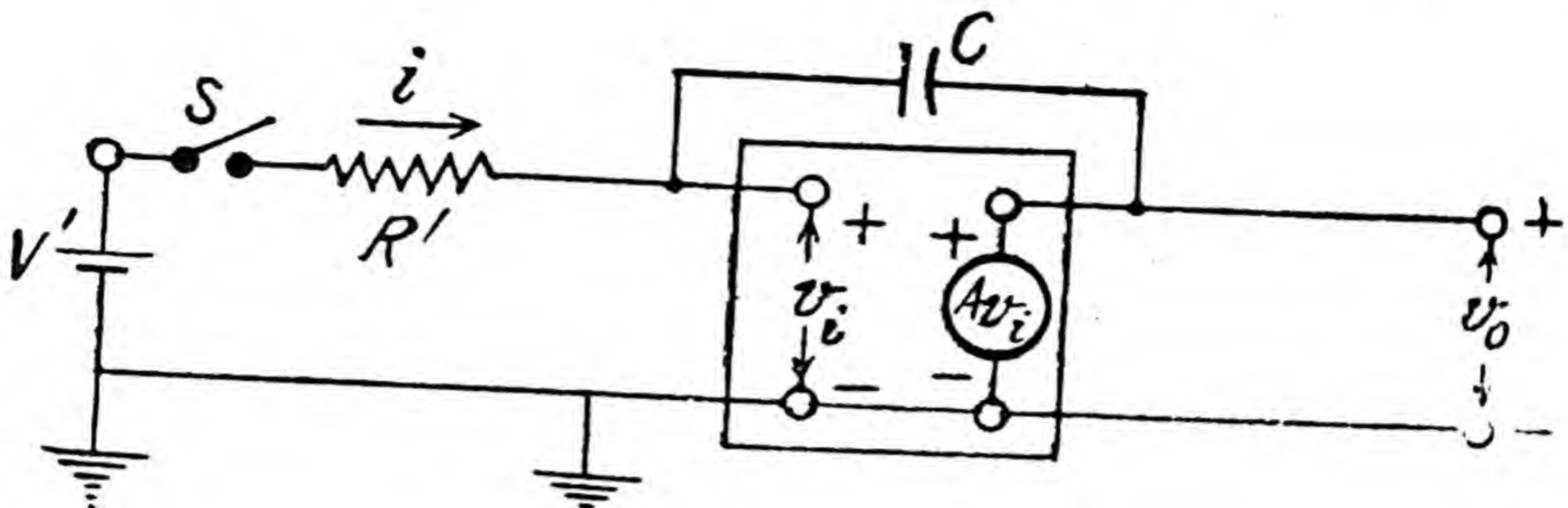


Fig. 14 (*h*)

(i) At start : When switch, *S*, is closed at $t=0$, we consider that the capacitor voltage is zero. Since the capacitor voltage was zero at $t=0^-$ (meaning $t<0$ but close to 0 *i.e.*, immediately before closing the switch), it must be zero at $t=0^+$ (meaning $t>0$ but close to 0, *i.e.* immediately after closing the switch). The voltage across the capacitor at $t=0^+$ is

$$v_i - Av_i = (1-A)\, v_i$$

but it should be zero so that

$$(1-A)\, v_i = 0$$

or

$$v_i = A\, v_i = v_0 = 0, \qquad \ldots(1)$$

i.e. output voltage is zero.

(ii) At $t=\infty$: The capacitor is fully charged and no current flows through it so that

$$v_i = V'$$

and

$$v_0 = Av_i = AV' \qquad \ldots(2)$$

Thus output voltage, v_0, waveform is a simple exponential because the circuit involves a single capacitor. It will reach a steady state value AV' starting from zero. The sweep is negative going because A is a negative number.

Slope error is given by

$$e_{slope} = \frac{\text{Sweep amplitude}}{\text{Steady state value}} = \frac{V_S}{|A|\,V'}$$

$$= \frac{V_S}{V'}\,\frac{1+R/R_i}{|A|} \qquad \ldots(3)$$

where as in simple RC circuit (shown in fig. 14*a*) it is

$$e_{slope} = \frac{V_S}{V'}$$

Thus in present case (Miller sweep) deviation from linearity is $\left(1+\frac{R}{R_i}\right)\Big/ A$ times that of a capacitor charging through R directly from a source, V.

The *sweep speed* is given by i/c for the circuit, where virtual short circuit exists at the input terminals to the base amplifier, we have $i \approx V'/R'$ so that

$$\text{Sweep speed} = \frac{V'/R'}{C}$$

$$= \frac{V'}{R'C} \cdot \qquad \ldots(4)$$

and it is same as in the case of a capacitor charging through a resistance directly from the source.

The deviation from linearity can also be expressed interms of sweep period as follows :

As gain, A, is negative, we can take the shunting capacitance

$$C_i = C\,(1-A)$$

so that input voltage, v_i, is

$$v_i = V'\,(1-e^{-t/\tau_1}) \qquad \ldots(5)$$

where $\tau_1 = (1-A)\,R'C.$

We note that the effective input time constant τ_1. is $R'C$ multiplied by the voltage gain of the amplifier and will normally be large. Thus time constant get *augmented* by the feedback and linearity of the sweep is improved.

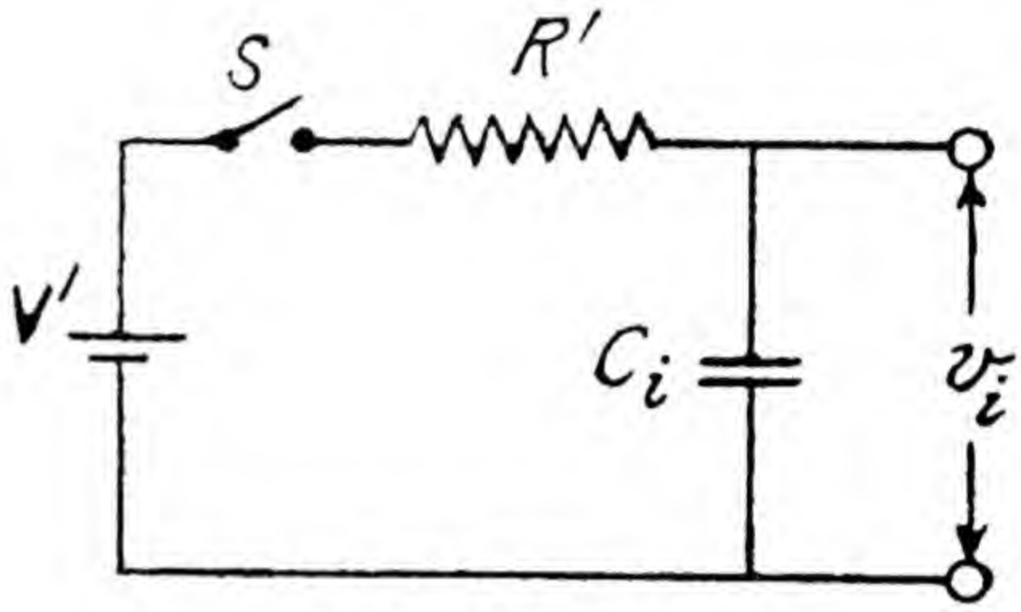


Fig. 14(*i*)

The output voltage is given by

$$v_0 = A v_i\,(1-e^{-t/\tau_2}) \qquad \ldots(6)$$

where $\tau_2 = R_0 C$

as we have already taken R_0, the output resistance of base amplifier as negligible, *i.e.*, $r_2 \ll \tau_1$. Thus t/τ_2 is quite large so that we can write

$$v_0 = A\,v_i$$

$$= AV'\,(1-e^{-t/\tau_1}, *$$

$$= AV'\left[1-\left(1-\frac{1}{\tau_1}+\frac{t^2}{2\tau_1^2}-\cdots\right)\right]$$

$$= AV'\left(\frac{t}{\tau_1}-\frac{t^2}{2\tau_1^2}\right)$$

$$= \frac{AV't}{\tau_1}\left(1-\frac{t}{2\tau_1}\right) \qquad \ldots(7)$$

and at the end of the sweep period

$$v_0(T) = \frac{AV'T}{\tau_1}\left(1-\frac{T}{2\tau_1}\right)$$

where T is time at the end of sweep period. The second term in eq. (8) represents deviation from linearity. Thus non-linearity is

*Note for $t=\infty$,

$$e^{-t/\tau_1} = e^{-\infty} = 0$$

so that $v_0 = AV'$ as deduced earlier in eq. (2).

within $\left(\frac{T}{2\tau_1}\right)\times 100$ percent. As voltage gain, A, is sufficiently high in the case of an operational amplifier, long period sweep voltages can be generated.

To sketch and label the output waveform : Let us take an example. Specifications are indicated in the figure (14*j*). OP-AMP

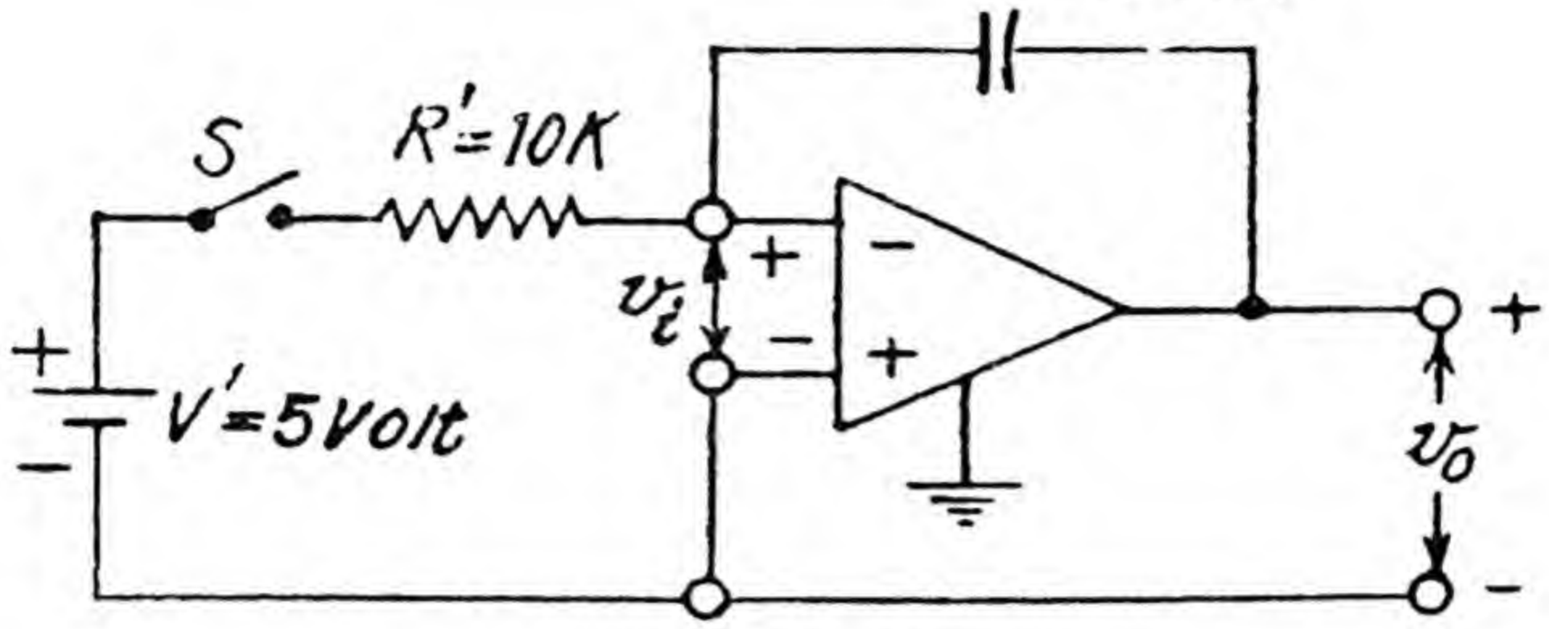


Fig. 14 (*j*)

has a maximum voltage gain of -5000, a very large input impedance, an output voltage limit of $\pm 10V$, and the switch is closed for 100 ms. We shall aiso determine the percent deviation from linearity.

From $\tau = (1 - A)\, R'C$

$\approx -AR'C$

$\simeq 5000 \times 10^4 \times 0{\cdot}1 \times 10^{-6}$

$= 5s.$ (as $R' = 10K\Omega$, $C = 0{\cdot}1\mu F$)

From equation (7), we have output voltage for $t = 100\ ms$ for which switch is closed as

$$v_0\,(0{\cdot}1) = \frac{(-5000)\times 5\times 0{\cdot}1}{5}\left(1 - \frac{0{\cdot}1}{2\times 5}\right) \quad (\text{as } t = 0{\cdot}1s)$$

$$\simeq -500V.$$

But output can reach only $-10V$. Therefore time T, in which output reaches this value is

$$v_0(T) = \frac{AV'T}{\tau_1}\left(1 - \frac{T}{2\tau_1}\right)$$

$$\simeq \frac{AV'T}{\tau_1}$$

or $$-10 = \frac{-5000\times 5\times T}{5}$$

or $$T = 0{\cdot}002s \qquad (= 2ms).$$

This is the sweep deviation. For this sweep deviation, percent of deviation from linearity is given by

$$= \frac{T}{2\tau_1}\times 100$$

$$= \frac{0{\cdot}002}{2\times 5}\times 100$$

$$= 0{\cdot}02\%.$$

Output waveform is shown in fig. (14*k*). Sweep waveform is only for duration 2*ms*.

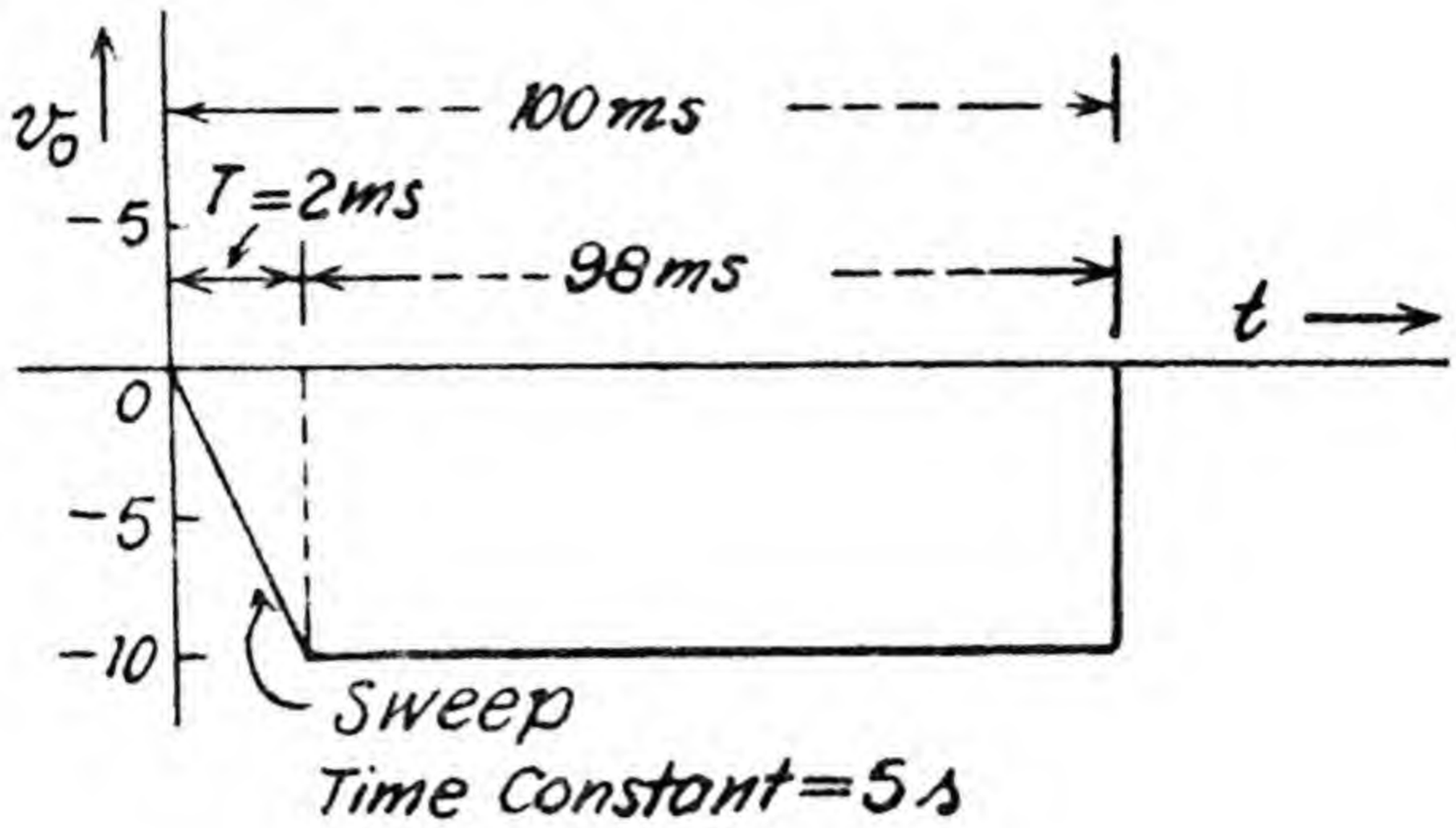


Fig. 14(*k*)

For vacuum tube Miller circuit refer to chapter on 'Pulse Generation and Scaling Circuits'.

22·6. BOOT STRAP SWEEP GENERATOR :

In this type of generator, we use positive feedback for achieving linearity in sweep. A general capacitor charging circuit is shown in fig 15 (*a*). Current, *i*, is given by

$$i=\frac{V-v_c}{R},$$

where v_c is voltage across the capacitor, at any time t, after the switch is off. This current is not constant due to nonlinearity in v_c. If V could be replaced by $(V+v_c)$ then the desired charging current will become constant at

$$i=\frac{(V+v_c)-v_c}{R}=\frac{V}{R}$$

It means we have to employ means of positive feedback of a voltage equa l to v_c. This can be done by amplifying v_o with unity gain and introducing the output in *series* with V. A basic boot strap circuit employing positive feed back for sweep linearization is shown in fig. (15*b*). The gain of the OP-AMP is A so that at $t=0^+$, we have

$$i_c=\frac{V-v_c+Av_c}{R},$$

assuming input impedance of the amplifier to be very large. Thus

$$i_c=\frac{V-(1-A)\,v_c}{R}, \qquad \ldots(1)$$

Obviously if $A=1$. unity gain, then

$$i_c=\frac{V}{R}=\text{constant}=I_f \qquad \ldots(2)$$

and the output is

$$v_0=Av_c$$

$$= A\frac{q}{C} = \frac{AI_c t}{C} = \frac{A}{C} \cdot \frac{V}{R} \cdot t$$

$$= \frac{V}{RC} \cdot t \qquad \text{for unity gain} \qquad \ldots(3)$$

We note that the input of fig. 15 (*b*) is lifted by its own output or by its 'boot straps'.

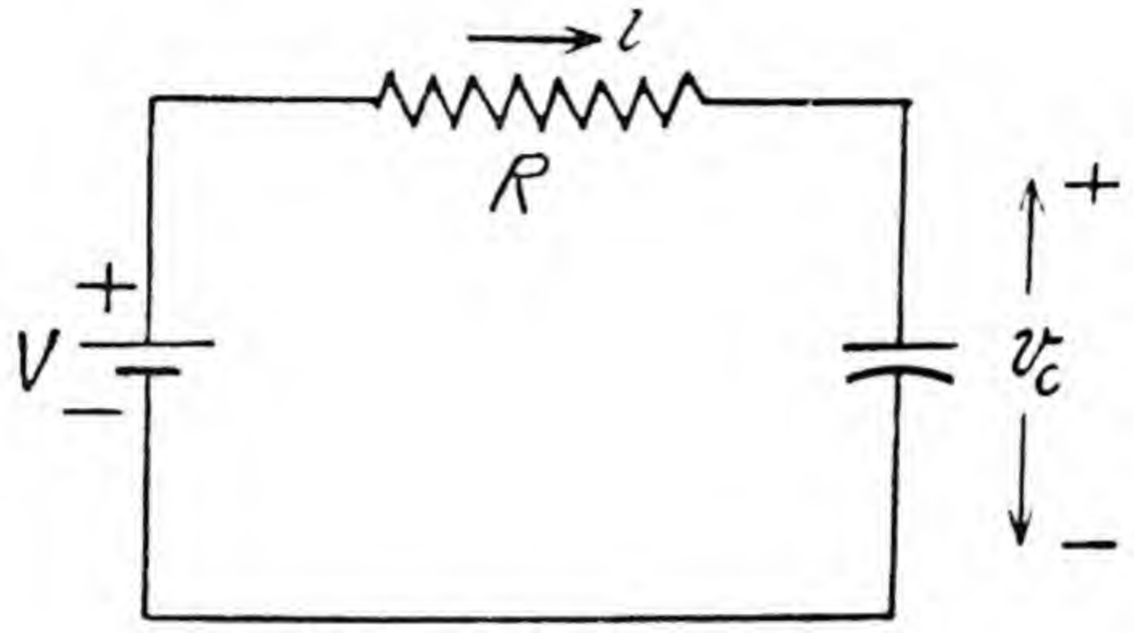


Fig. 15 (*a*)

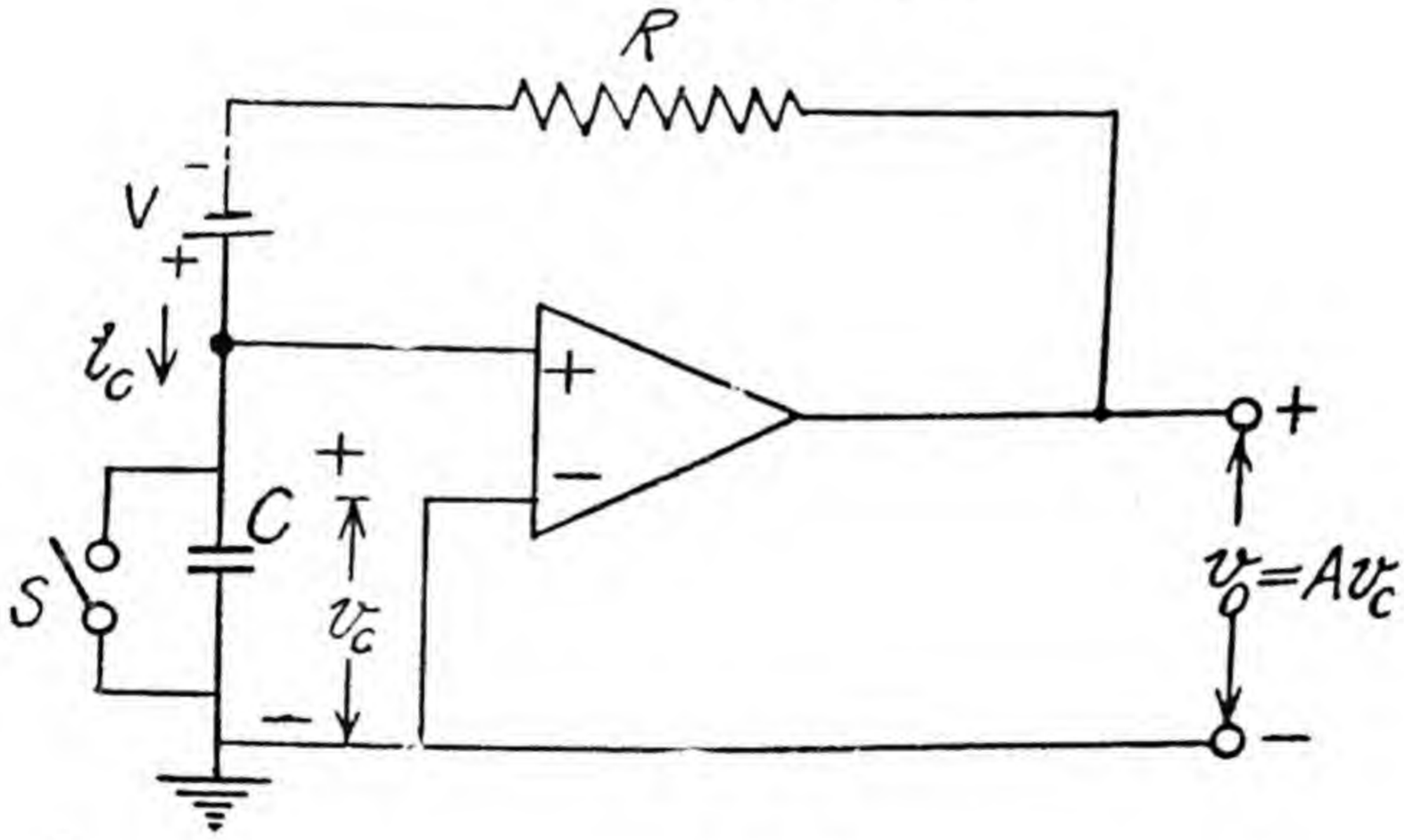


Fig. 15(*b*)

At any instant of time, voltage, v_c, across the capacitor, C, is given by

$$v_c\,{}_t = \frac{1}{C}\int_0^t i_c\, dt$$

$$= \frac{1}{C}\int_0^t \left[\frac{V}{R} - \frac{(1-A)\,v_c}{R}\right] dt \qquad \ldots(4)$$

Differentiating it with respect to t, we get

$$\frac{dv_c}{dt} + \frac{(1-A)v_c}{CR} = \frac{1}{CR}\, V \qquad \ldots(5)$$

solution of which is

$$v_c\big|_t = \frac{V}{(1-A)}\left[\, 1 - e^{-(1-A)t/RC}\right] \qquad \ldots(6)$$

This shows that time constant is $RC/(1-A)$. *Obviously, as* $A \to 1$, *time constant tends towards infinite value. Thus positive feed*

back with unity gain greatly enhances the time constant and improves the sweep linearity.

In practice it is not possible to keep amplifier gain, A, close to unity Therefore it is desirable to establish a relation between the time duration in which the voltage ramp reaches the sweep amplitude, V_S as a function of gain, A.

We can write, using eq. (6), that

$$V_S = \frac{V}{(1-A)}[1-e^{-(1-A)t/RC}]$$

or $$V_S(1-A) = V[1-e^{-(1-A)t/RC}]$$

or $$V\,e^{-(1-A)t/RC} = V-(1-A)\,V_S$$

or $$e^{(1-A)t/RC} = \frac{V}{V-(1-A)\,V_S}$$

or $$t = \frac{RC}{(1-A)}\log_e\left[\frac{V}{V-(1-A)V_S}\right] \quad \ldots(7)$$

If we put $V=V_S$, than $t=T$, the sweep duration

$$T \quad \frac{RC}{(1-A)}\log_e\left(\frac{1}{A}\right) \quad \ldots(8)$$

From eq. (3), we can infer that, as $A=1$, we get linear ramp (sweep). The sweep duration, T_{lin}, for linear ramp is

$$T_{lin} = CR$$

so that from eq. (8), we arrive at

$$T = \frac{T_{lin}}{(1-A)}\log_e\left(\frac{1}{A}\right)$$

or $$\frac{T}{T_{lin}} = \frac{1}{(1-A)}\log_e\left(\frac{1}{A}\right) \quad \ldots(9)$$

Which for $A=0{\cdot}98$ is

$$\frac{T}{T_{lin}} = 1{\cdot}08.$$

From eq. (9), it is clear that the sweep duration is very sensitive to amplifier gain variation. For better lineraity, amplifier gain should be stable and have unity magnitude.

Boot Strap Generator using Emitter Follower :

Circuit is shown in fig. 15 (*c*). In it a transistor with high β is used so that the emitter follower may have a gain close to unity required for linearization of sweep

V is a d.c. source which will charge the capacitor once the switch, S, is open. Normally switch, S, is closed and resistor, R, controls the current through diode. Initially voltage across the capacitor equals the forward voltage drop across the diode.

Sweep starts when switch, S, is opened at $t=0$. Diode, D,

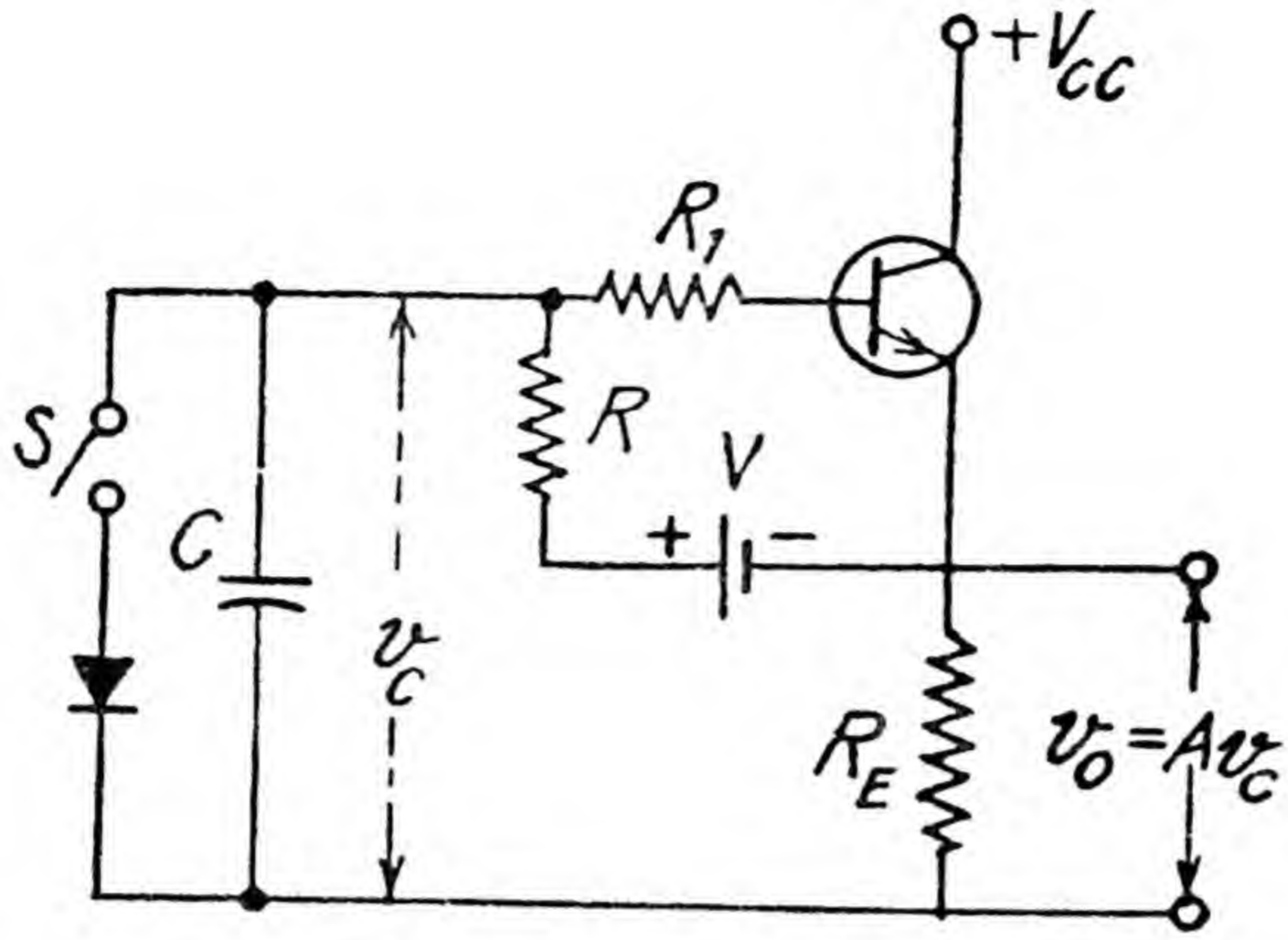


Fig. 15 (*c*)

through its forward voltage drop, helps in turning ON the transistor so that any delay in starting the sweep is minimised. Switch is opened. v_c rises towards steady state value,

$$v_o\Big|_{t=\infty} = \frac{V}{(1-A)} \text{ (form eq. 6 on putting } t=\infty)$$

with charging time eonstant

$$\tau = \frac{CR}{(1-A)}$$

As the sweep voltage begins to increase, the amplified sweep voltage adds in series with the d.c. voltage, V, to provide the boot strap action needed for linearity of the sweep. The output voltage, v_0, continues to increase till the transistor saturates at time $t \approx T_{sweep}$.

For vacuum tube bootstrap circuit, refer to chapter on 'Pulse Generation and Scaling Vacuum Tube Circuits'.

22·7. CURRENT TIME BASE GENERATOR :

Such generators are used to obtain a current waveform which increases linearly with time. They are used in television deflection systems where, by employing magnetic deflection, a greater deflection sensitivity is achieved.

Linear sweep voltage. methods which have been described in earlier articles, if applied to a resistor will generate a current that increases linearly with time. But we shall discuss here a method that will cause a linearly time varying current to flow through an inductor.

Behaviour of an Ideal Inductor with Ideal Voltage source : Ideal inductor means that its self capacitance and resistance, R_L, are negligible. Ideal voltage source means that its source resistance

R_S, is zero. If, at time $t=0$, (fig. 16*a*) a voltage, V_S, is applied to a coil of inductance, L, in which the current is initially zero, then the current in ideal inductor, i_L, is given by

$$i_L\Big|_{t>0}=\frac{1}{L}\int V_S\,dt=\frac{V_S}{L}\,t \qquad \ldots(1)$$

because, V_S, being derived from an ideal voltage source, is constant. We note thet current, i_L, will increase linearly with time (fig. 16*b*).

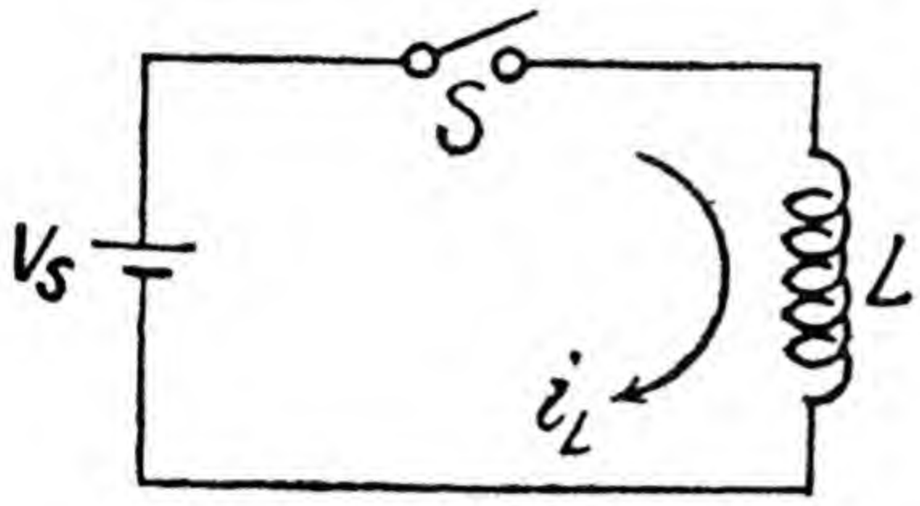


Fig. 16 (*a*)

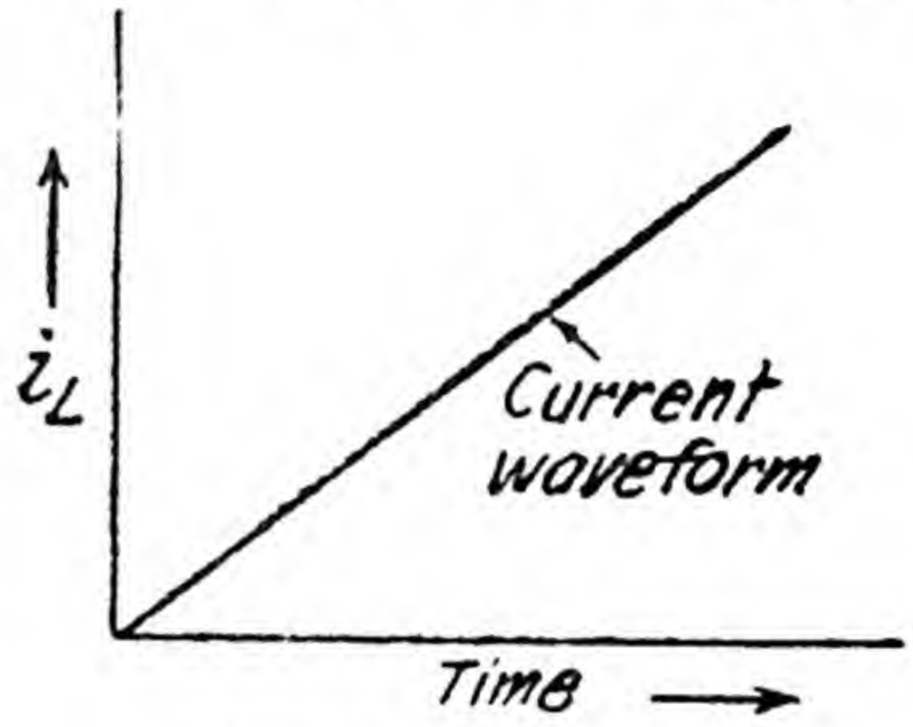


Fig. 16 (*b*)

Behaviour of Practical Inductor with Practical Voltage Source. A practical inductor will have a self capacitance and a self resistance R_L. And a practical voltage source will consist of a source resistance R_S. Taking initially $t=0$, current, i_L, to be zero, we can write for such a circuit (fig. 16*c*) that

$$V_S=i_L\,(R_S+R_L)+L\,\frac{di_L}{dt}$$

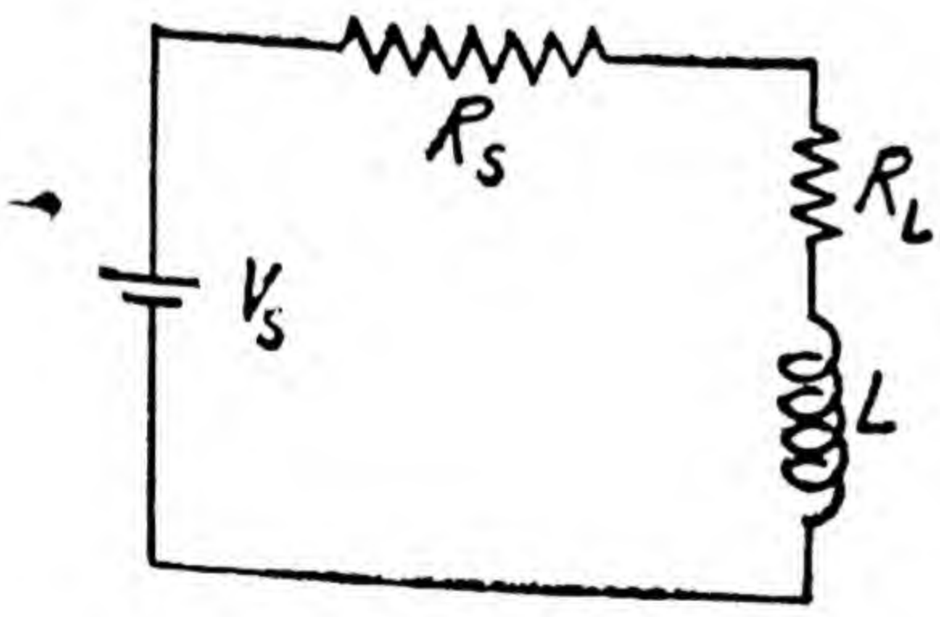


Fig. 16 (*c*)

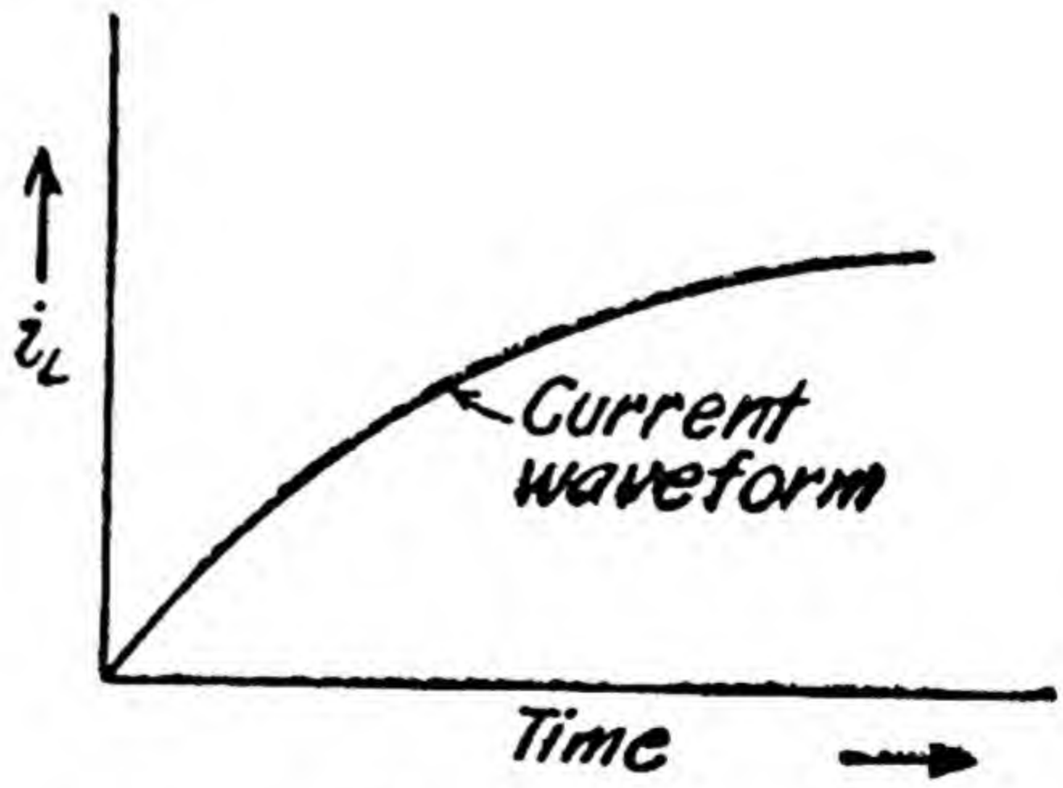


Fig. 16 (*d*)

Which can be solved to give an inductor current of value

$$i_L\Big|_{t>0}=\frac{V_S}{R_S+R_L}\left[1-e^{-t\,(R_S+R_L)/L}\right] \qquad \ldots(2)$$

Which shows a variation depicted in fig. (16 *d*), predicting that current rises exponentially to a steady state value

$$i_L\Big|_{t=\infty}=\frac{V_S}{R_S+R_L}=I_L \qquad \ldots(3)$$

From eq. (3) we can then write

$$i_L\Big|_{t>0} = i_L\Big|_{t\to\infty}\left[1-e^{-\frac{t}{L/(R_S+R_L)}}\right] \qquad \ldots(4)$$

The time constant of the circuit is

$$\tau = \frac{L}{R_S+R_L} \qquad \ldots(5)$$

For sweep duration (T_S) much less than the time constant, current i_L, increases linearly with time. Voltage across inductor is

$$\begin{aligned} V_L\Big|_{t>0} &= L\,\frac{di_L}{dt} \\ &= L\,\frac{d}{dt}\left[\frac{V_S}{R_S+R_L}\left(1-e^{-t\,(R_S+R_L)/L}\right)\right] \\ &= L\cdot\frac{V_S}{R_S+R_L}\cdot\frac{R_S+R_L}{L}\cdot e^{-t\,(R_S+R_L)/L} \\ &= V_S\ e^{-t\,(R_S+R_L)/L} \qquad \ldots(6) \end{aligned}$$

so that $V_L\,|_{t=0} = V_S$ .. (7)
$V_L\,|_{t=\infty} = 0.$

A simple current sweep : A simple current sweep circuit biased on basic principle is shown in fig. (16 *e*). Transistor acts as a switch as it switches ON or OFF the inductive load. As the energy is stored in the inductor when a current is flowing through it, some means should be provided to dissipate this energy when the transistor turns off otherwise it will damage the transistor. To provide a path for it, a diode is shunted across the inductor. The cathode of the diode is connected to positive terminal of the supply while anode is connected to the collector.

The input waveform has two levels. The lower level keeps the transistor cut off while upper level drives the transistor into saturation. When the transistor is turned ON, then current i_L

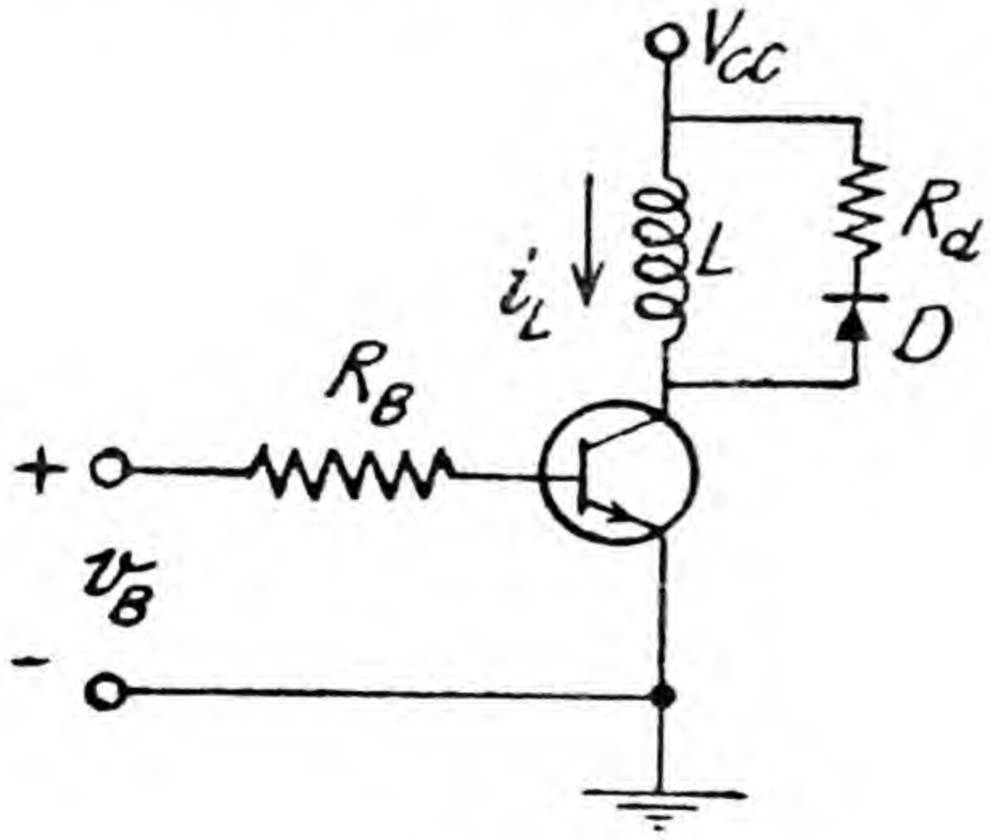


Fig. 16 (*e*).

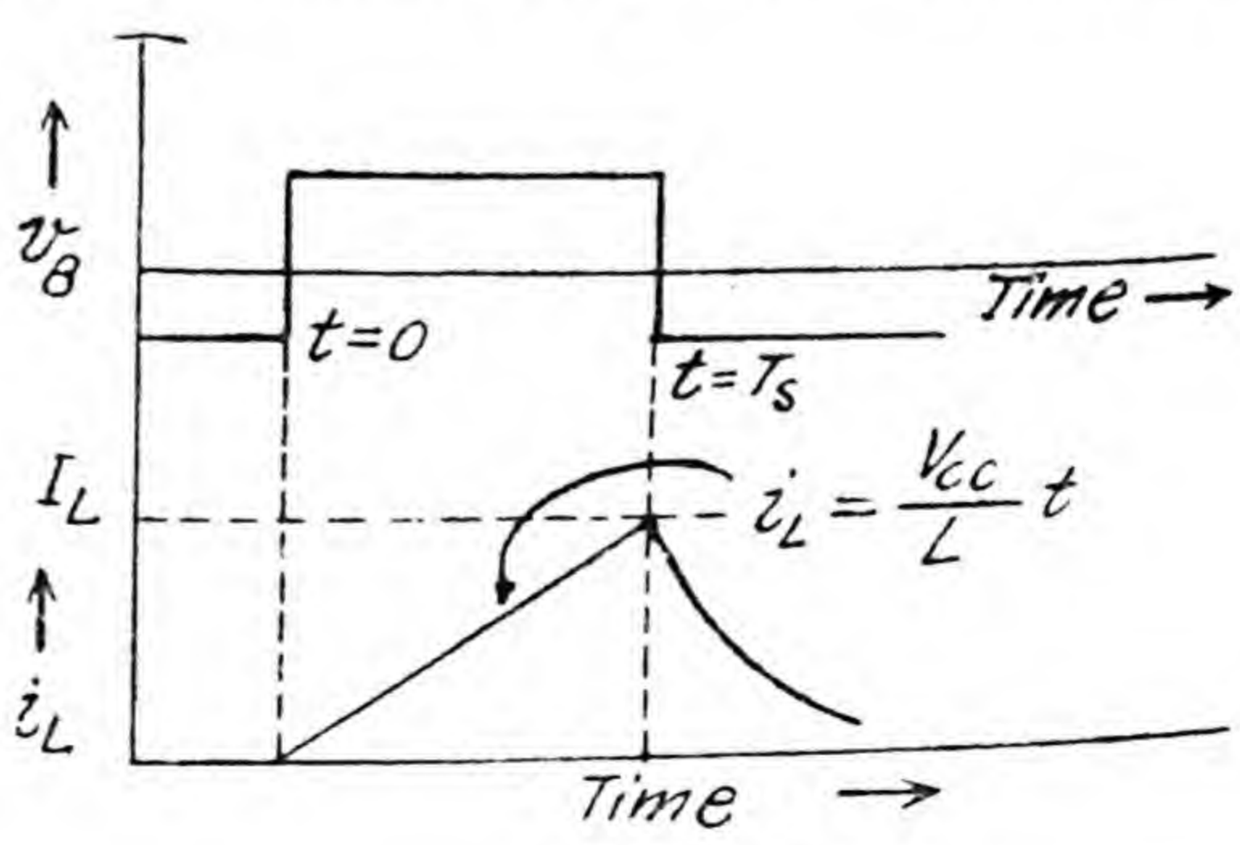


Fig. 16 (*f*)

increases* linearly with time. The diode D does not conduct due to reverse bias across it. The sweep terminates at $t=T_S$, when the input signal drives the transistor to cut-off. The stored inductive energy ($E_L=\frac{1}{2}\ Li_L^2$) makes the anode positive with respect to cathode and the diode conducts. The inductor current then continues to flow through the diode, D and resistor R_d until it decays to zero. This decay is exponential with time constant $\tau=\frac{L}{R_d}$ where R_d is sum of damping resistance and diode forward resistance.

If collector saturation resistance, R_{CS}, of the transistor and resistance associated with inductor, R_L, is considered then eq. (3) (taking $V_S=V_{CC}$) can be written as

$$\begin{aligned} i_L &= \frac{V_{cc}}{(V_L+R_{CS})}\left[1-e^{-(R_L+R_{CS})\,t/L}\right] \\ &= \frac{V_{cc}}{(R_L+R_{CS})}\left[\frac{(R_L+R_{CS})\,t}{L}-\frac{(R_L+R_{CS})\,t^2}{2L^2}\right] \\ &= \frac{V_{cc}.t}{L}\left[1-\tfrac{1}{2}\frac{(R_L+R_{CS})\ t}{L}\right] \\ &= \frac{V_{cc}.t}{(R_L+R_{CS})}\cdot\frac{1}{L/(R_L+R_{CS})}\left[1-\tfrac{1}{2}\frac{(R_L+R_{CS})\,t}{L}\right] \\ &= \frac{I_L\,t}{\tau}\left[1-\frac{t}{2\tau}\right] \end{aligned} \quad \ldots (8)$$

Using eq. (3) and (5) with $V_S=V_{cc}$. This equation is analogous to eq. (8) of art. 22·5. The current then departs from a linear increase with time just as in case of voltage sweep For sweep period, T_S, the non linearity is represented by the term $\frac{T_S}{2\tau}$. To keep it small, τ should be large or (R_L+R_{CS}) should be small or in otherwords, if voltage $(R_L+R_{CS})\,I_L$ across the total circuit resistance is kept small compared to supply voltage V_{cc}, then linearity in inductor current waveform will be maintained.

EXERCISES

1. With reference to voltage sweep generators define ideal voltage ramp, slope error, differential linearity, and transmission error.
2. Describe a fundamental sweep voltage generator. Show that sweep is linear if time constant is about ten times larger than the sweep duration. Find slope and transmission error in this case.
3. Describe a transistor constant current sweep generator. Show that sweep is linear until capacitor voltage reaches a value equal to the base bias voltage in a common base configuration.
4. Describe a UJT sweep waveform generator. Sketch the output waveform.

*Neglecting the effect of collector saturation resistance (R_{CS}) of the transistor.

5. Discuss a Miller integrator sweep circuit. Show that nonlineartiy of sweep waveform is within $\frac{T}{2\tau} \times 100$ percent where T is time at the end of sweep period and τ is the effective input time constant.
6. Discuss how sweep linearization is achieved in a bootstrap sweep generator. Show that positive feedback with unity gain enhances the time constant of the circuit. Describe an emitter follower circuit.
7. For what purpose current time base generators are put to ? Describe a simple current sweep circuit, with theory indicating clearly the effect of associated resistance of the inductor on the linearity of the sweep.

23 INTEGRATED CIRCUIT FABRICATION

For more than half a century, electronic circuits have been assembled from individual components. The requirements of a good electronic circuit are, compactness, reliability and cost. The size of the electronic circuits has been considerably reduced by the introduction of transistors in the field of electronics. The size of the electronic circuit has been further made compact by the use of printed circuits. In printed circuits the inter connections between various com ponents are made through thin metal strips formed on an insulating sheet. This method brought a certain degree of compactness but still there is no increase in reliability because the active and passive components are to be soldered. of course connecting wires are eliminated. The reliability of the circuit can be increased provided the inter connection between active and passive components are eliminated. Actually this is made possible in integrated circuits. In integrated circuits, the active and passive components are built in the same crystal. An integrated circuit consists of a single crystal chip of silicon, typically 50 by 50 mils* in cross section containing both active and passive elements and their interconnections.

23·1. INTEGRATED CIRCUIT TECHNOLOGY :

The basic structure of a integrated circuit is shown in fig. 1 (b). This consists of the following four layer of different materials :

(1) The bottom layer is a p-type silicon layer of 0·1524 mm thickness. This layer serves as a body or substrate upon which integrated circuit is to be built.

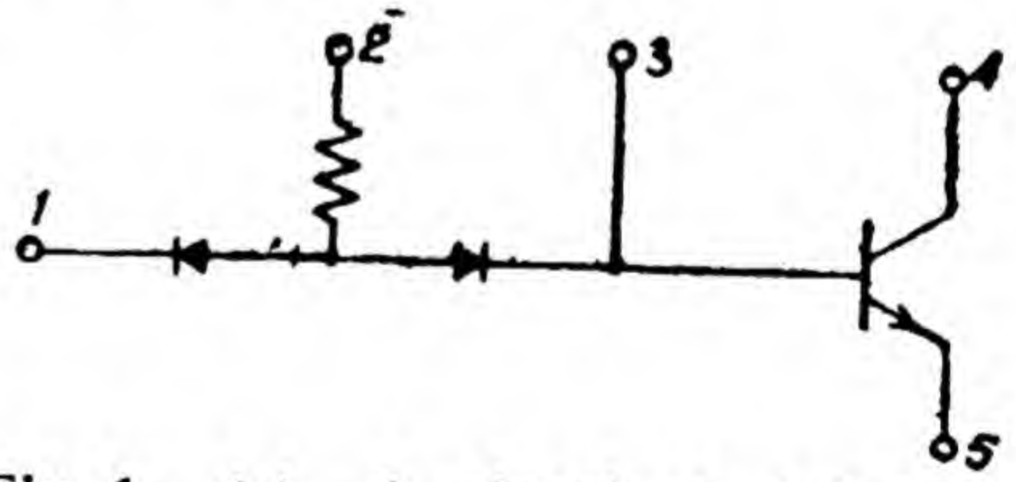


Fig. 1. (*a*). A circuit containing a resistor, two diodes and a transistor.

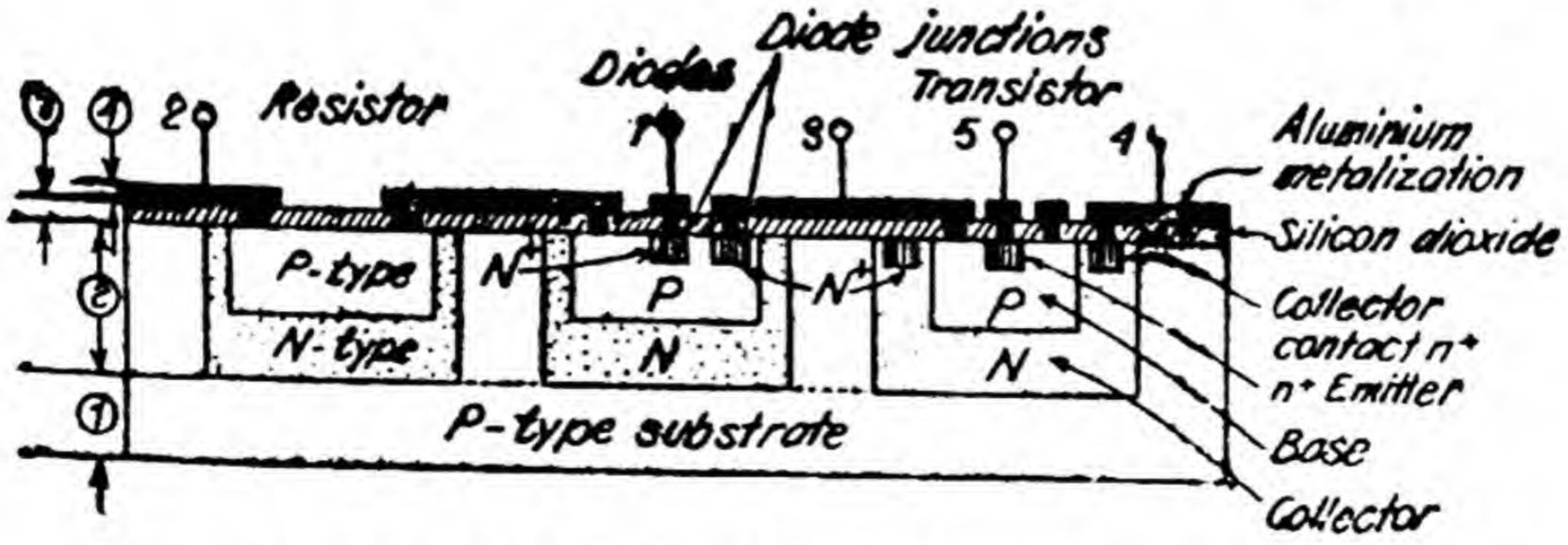


Fig. 1. (*b*) Showing the basic structure.

*1 ml=0·0254 m.m.

(2) The second layer (which is nearly 0·025 mm thick) is of N-type material which is grown as a single crystal extension of the body or substrate Using a number of diffusion steps, all the active and passive components are fabricated or built within this layer. These components are transistors, diodes, capacitors and resistors and are made by diffusing P-type and N-type impurities. Here it is necessary to diffuse impurities in certain precisely defined regions within this layer. The fabrication of a transistor is most complicated.

(3) The third layer is of silicon diode (SiO_2) material. The purpose of this layer is to protect the semiconductor surface against contamination. The layer acts as a barrier in the selective diffusion of impurities in the second layer. The layer also protects portions of the wafer against impurity penetration. The SiO_2 layer is etched away in the regions where diffusion is to take place, leaving the rest of the wafer protected against diffusion For selective etching, the SiO_2 layer is subjected to a photolithographic process which is discussed in later article.

(4) The fourth layer is a metallic layer of aluminium. This is added to supply the necessary inter-connections between different fabricated components.

23·2 ADVANTAGES AND LIMITATIONS OF INTEGRATED CIRCUITS :

Advantages : Following are the advantages offered by integrated circuits :

(i) Due to the processing of large quantities of the components, the cost is low.

(ii) The size is small.

(iii) They have high reliability because all the components are fabricated simultaneously and there are no soldered joints.

(iv) To obtain better functional characteristics, more complex integrated circuits may be used and hence the performance is improved.

(v) There are no inter-connection errors.

(vi) Temperature differences between parts of a circuit are small.

(vii) A close matching of components and temperature coefficients is possible because they are fabricated simultaneously by the same processes.

Limitations : Following are the limitations of the integrated circuits :

(i) The inductors can not be integrated directly.

(ii) Capacitors and resistors are limited in maximum value.

(iii) Resistance and capacitance values are often dependent on voltage.

(iv) High grade PNP unit is not possible easily.

(v) Power dissipation is limited.

(vi) Low noise and high voltage operation are not easily obtained.

(vii) High frequency response is limited by parasitic capacitance.

23 3. BASIC MONOLITHIC INTEGRATED CIRCUIT TECHNOLOGY :

The word "*monolithic*" is derived from Greek *monos* meaning "single" and *lithos* meaning "stone". Thus a monolithic circuit is fabricated into a single stone *i.e.* a single crystal. The monolithic circuit is formed by the following steps :

Step 1 Epitaxial growth : An N–type epitaxial layer, typically 5 to 25 μm thick, is grown on a P–type substrate which has resistivity approximately 10 ohm-cm. Epitaxial process is one in which a thin layer of high resistivity silicon is grown on a low resistivity substrate The resistivity of N-type epitaxial layer can be chosen independently of that of the substrate. Generally values from 0·1 to 0·5 ohm-cm. are chosen for N–type layer. An oxide thin layer (0·5 micron = 5·000Å) of SiO_2 is formed over the entire epitaxial layer after polishing and cleaning as shown in fig 2. The SiO_2 is grown by exposing the epitaxial layer to oxygen atmosphere and heating at obout 1000°C. SiO_2 has the property of preventing the diffusion of impurities through it.

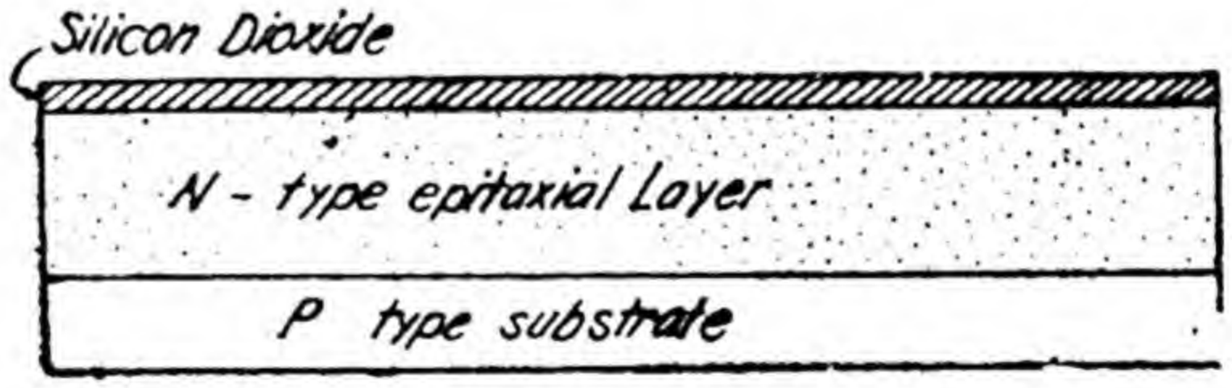


Fig. 2 Epitaxial growth

Step. 2. Isolation diffusion : By means of a photolithographic etching process (discussed in later article), the silicon dioxide is removed from four different places from the wafer as shown in fig. (3). The remaining SiO_2 serves as mask for the diffusion of acceptor impurities (boron, in this case). The wafer is now ready for the subjection of *isolation diffusion*. The P-type impurities are diffused to penetrate the N–type epitaxial layer and reach the P–type substrate, The diffused regions connect up with underlying P–region (the substrate) and form isolation pockets in the epitaxial layer. In this way we have the N–type shaded regions as shown in figure. These regions or sections are called *isolation islands* or *isolated regions*. This is due to the fact because they are separated by two back to back P–N junctions. The purpose of isolated regions is to

allow electrical isolation between different circuit components. Here it should be noted that (i) the P-type substrate is always held at a negative potential with respect to the isolation islands so that the P-N junction is reversed biased and (ii) the concentration of acceptor atoms in the region between isolated islands is generally much higher than in P-type substrate to prevent the depletion region of the reversed-biased isolation to substrate junction from extending into P material and possibly connecting two isolation islands.

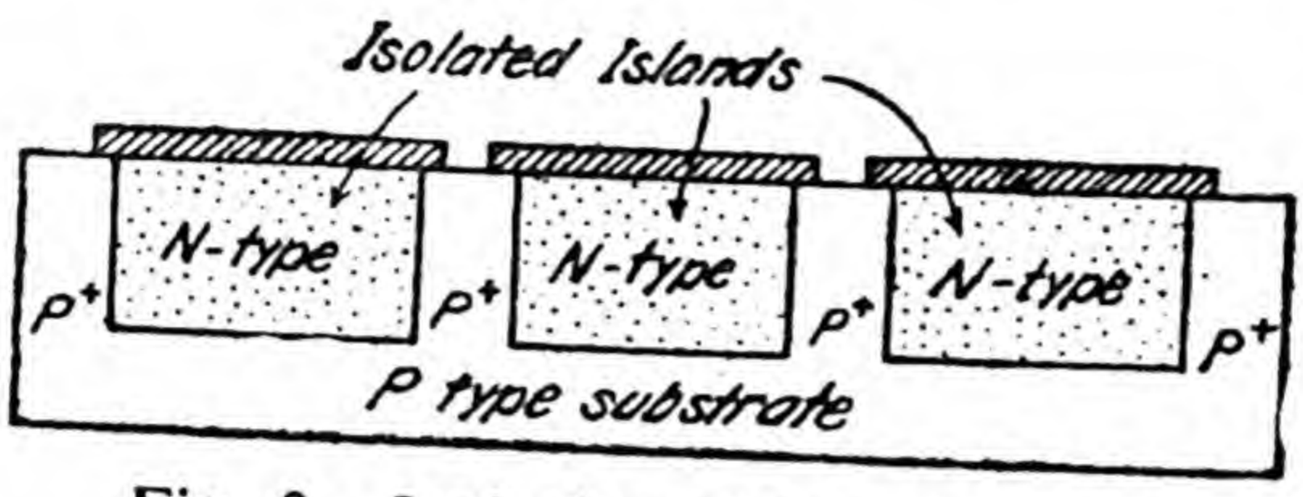


Fig. 3. Isolation Diffusion.

Step. 3. Base diffusion : In the base diffusion process, a new layer of oxide is again formed over the wafer. Now using photolithographic process new opening are created as shown in fig (4 . The P-type impurities (boron) are diffused through these openings. In this way, transistor base regions as well as resistors, the anode of diodes, and junction capacitors (if any) are formed. Here the depth of this diffusion is controlled in such a way that it does not penetrate to the substrate.

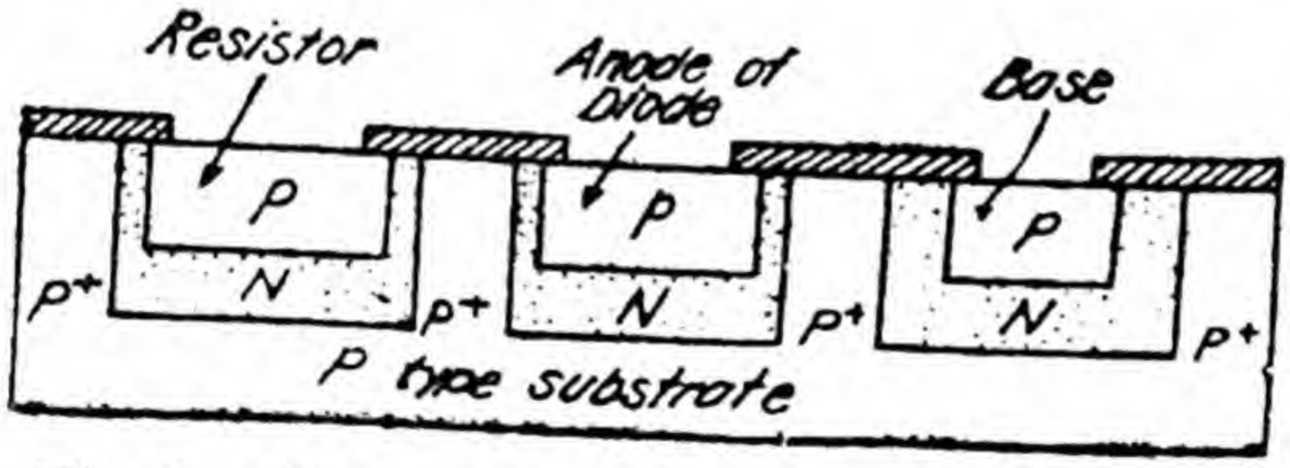


Fig 4. Base diffusion process.

Step. 4. Emitter diffusion : Again a layer of oxide is formed over the entire surface. By making and etching processes some windows are open in P-type regions as shown in fig. (5). Now N-type impurities (phosphorous) are diffused through these openings for the formation of transistor emitters, the cathode regions for diodes and junction capacitors. Some additional windows such as W_1 and W_2 are often made into N regions to which a lead is to be connected using aluminium as the interconnecting metal. In the

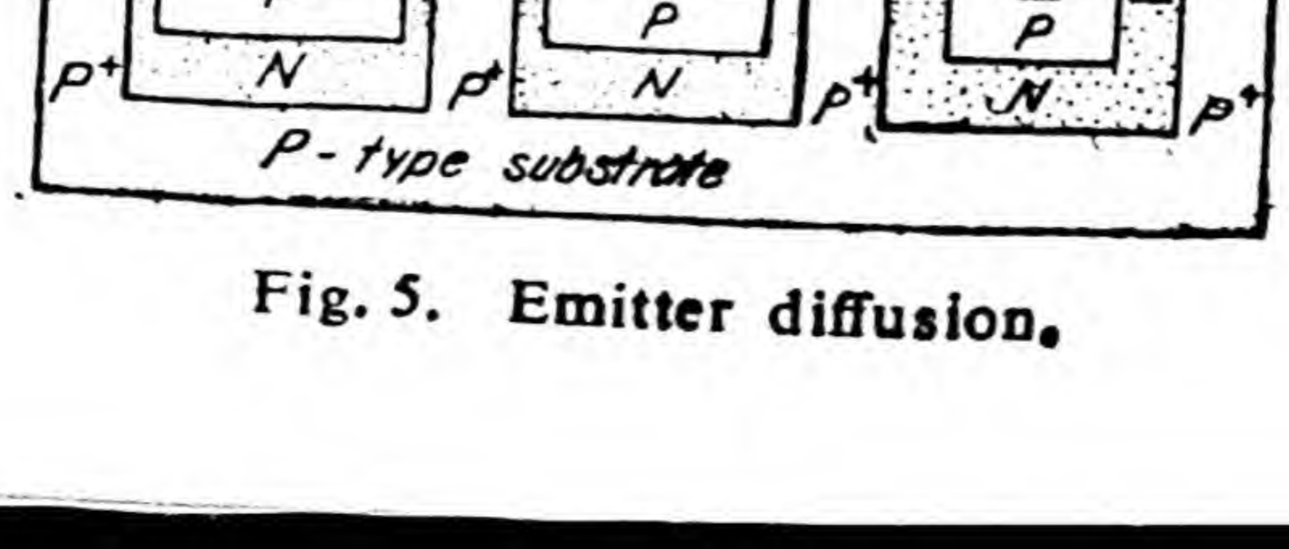


Fig. 5. Emitter diffusion.

phosphorous diffusion a heavy concentration (called N^+) is formed at the points where contact with aluminium is to be made.

Step. 5. Aluminium metalization : Using above mentioned steps, all P-N junctions and resistors have been formed. Now the only problem is to interconnect the various components of the integrated circuit as desired. To make these connections, the diffusion windows are closed with silicon dioxide layer and a fourth set of windows is opened at the points where contacts are to be made. This is shown in fig. (6). Now first of all the interconnections are made using a vacuum deposition of a thin even coating of aluminium over the entire wafer and then photoresist technique is used to etch away all undesired aluminium areas. In this way the desired interconnections are made as shown in fig. (6).

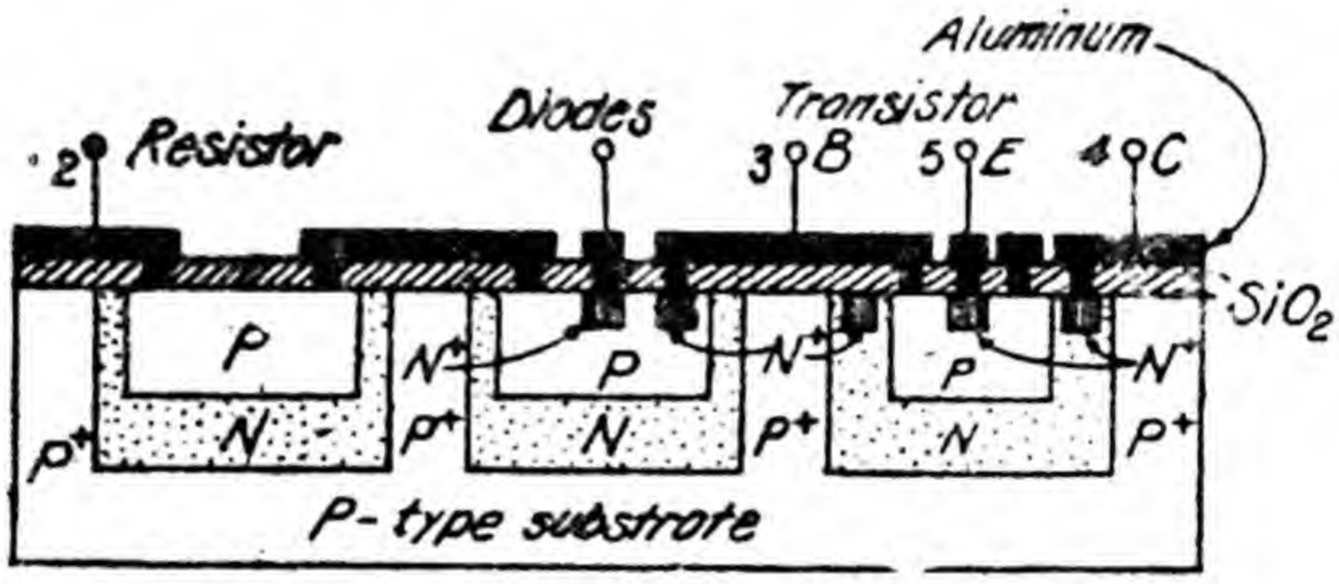


Fig. 6. Aluminium metallization

23·4. BASIC PROCESSES USED IN MONOLITHIC TECHNOLOGY :

The aim of this article is to discuss the different processes used in monolithic technology. Here we shall discuss these processes one by one.

(1) Epitaxial growth : The word epitaxial growth is derived from greek ' epi'' means *upon* and "teinein" means *arrange*. Epitaxial growth is a process or chemical reaction to form a thin film of a single crystal (monocrystalline) silicon (from a gaseous solution or vapour phase) with certain conduction properties on the surface of another silicon wafer or slice. The epitaxial laver may be either P-type or N-tvpe. The basic chemical reaction used to describe the epitaxial growth of pure silicon is the hydrogen reduction of tetrachloride :

$$SiCl_4 + 2H_2 \overset{1200°C}{\rightleftharpoons} Si + 4HCl.$$

As the epitaxial film of specific impurity concentrations is required and hence it is necessary to introduce impurities such as phosphine (PH_3) for N type doping or biborane (B_2H_6) for P-type doping into silicon tetrachloride. Hydrogen is bubbled through the volatile silicon compound, causing it to vaporize.

Fig. (7) shows the diagrammatic representation of a system for production growth of silicon epitaxial layers. The apparatus

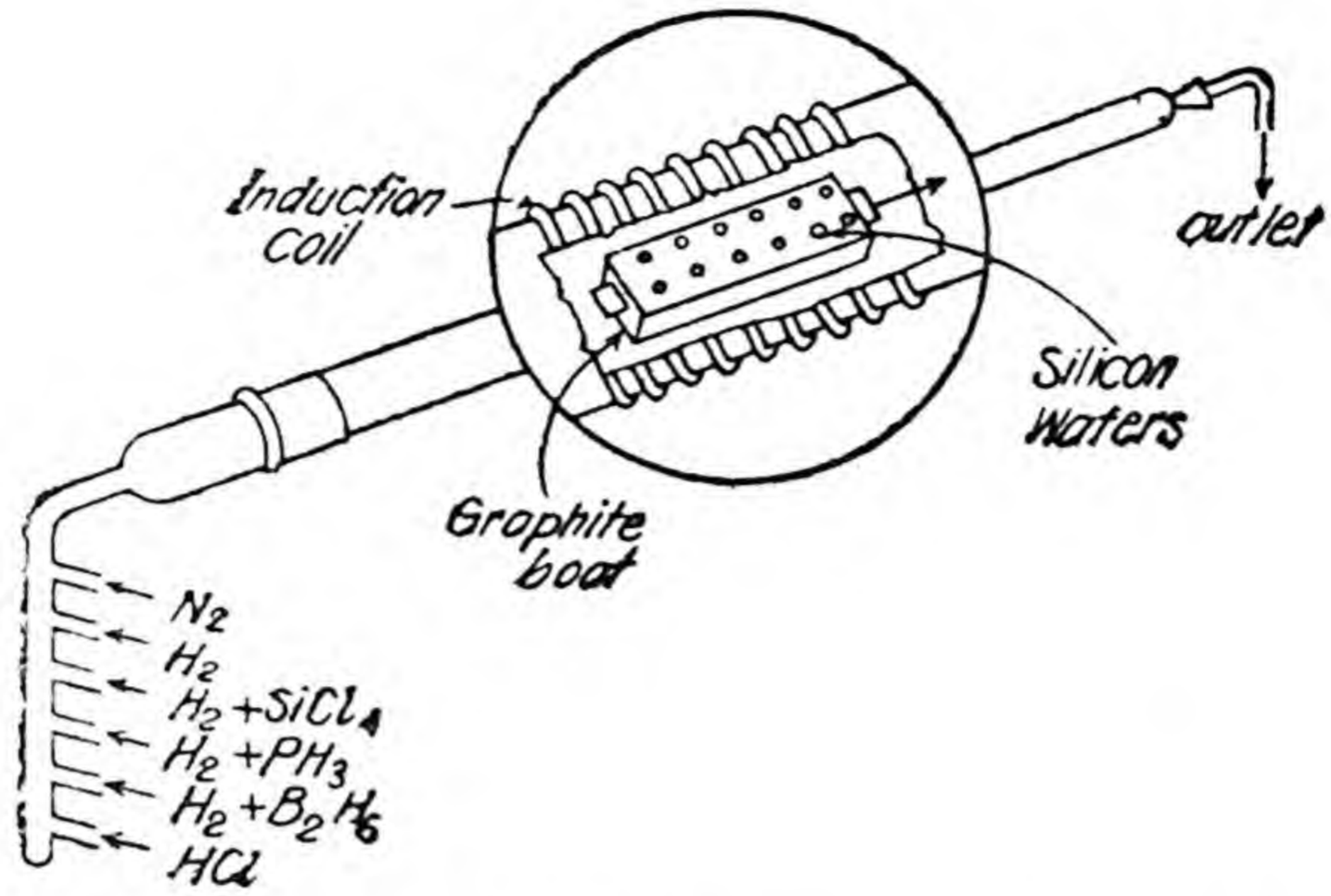


Fig 7. Apparatus used for epitaxial growth.

consists of a long cylindrical quartz tube which is encircled by a radio frequency induction coil. The silicon wafers are placed on a rectangular graphite rod called a boat. This boat is inserted in the reaction chamber (*i e.*, a special type of furnace) where the graphite is heated inductively to about 1200°C. At the input of reaction chamber a control console permits the introduction of various gases required for the growth of appropriate epitaxial layers. Silicon and dopant atoms from the vapour skid about on the surface of the growing epitaxial film until they find a correct position in the lattice and becomes fastened into the growing structure by interatomic forces.

The important features of the epitaxial process are as follows:

(i) The atoms of newly grown layer are arranged in single-crystal fashion on the single crystal substrate. Moreover, the lattice structure of the newly grown layer is an exact extension of the substrate crystal structure as shown in fig. 8).

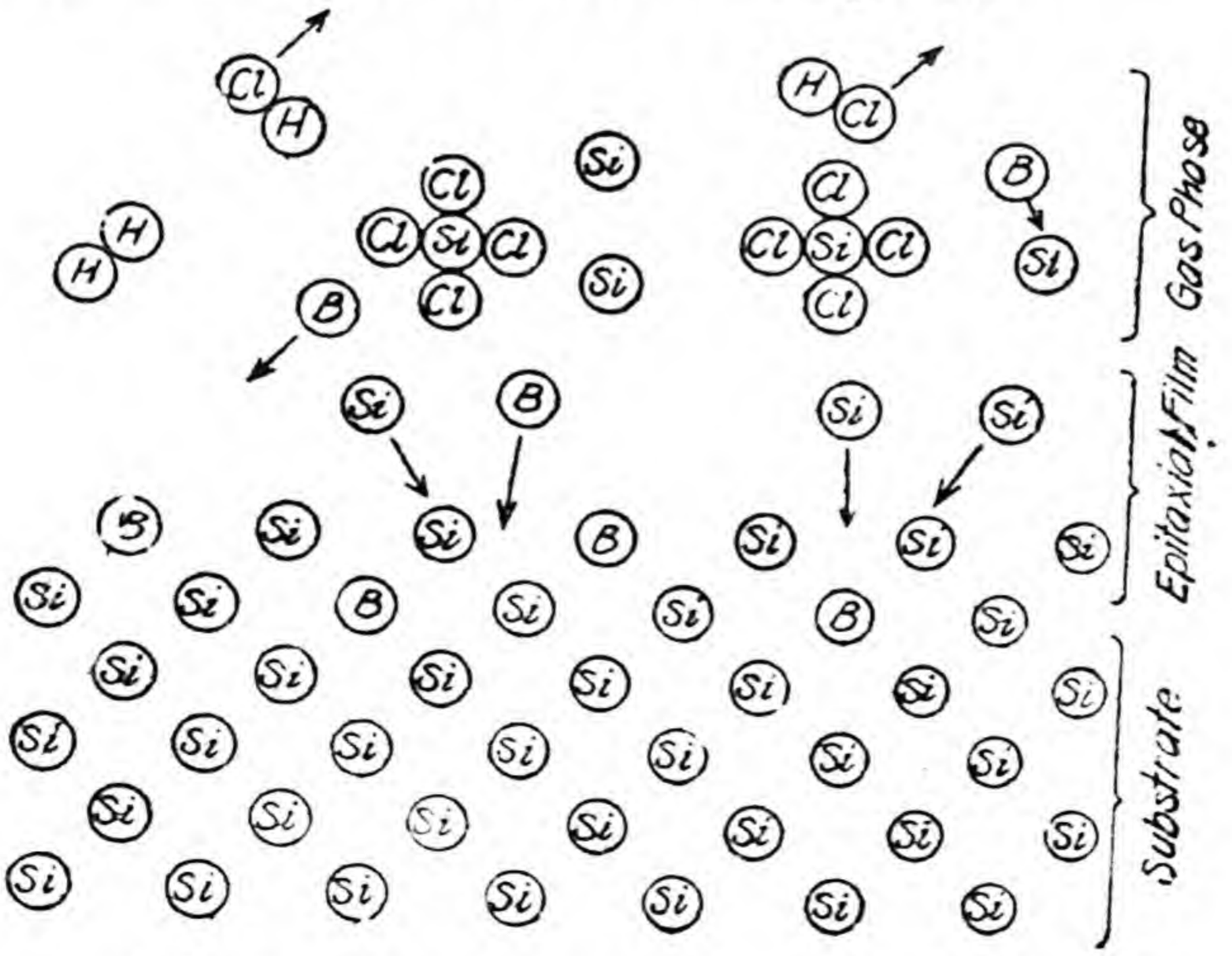
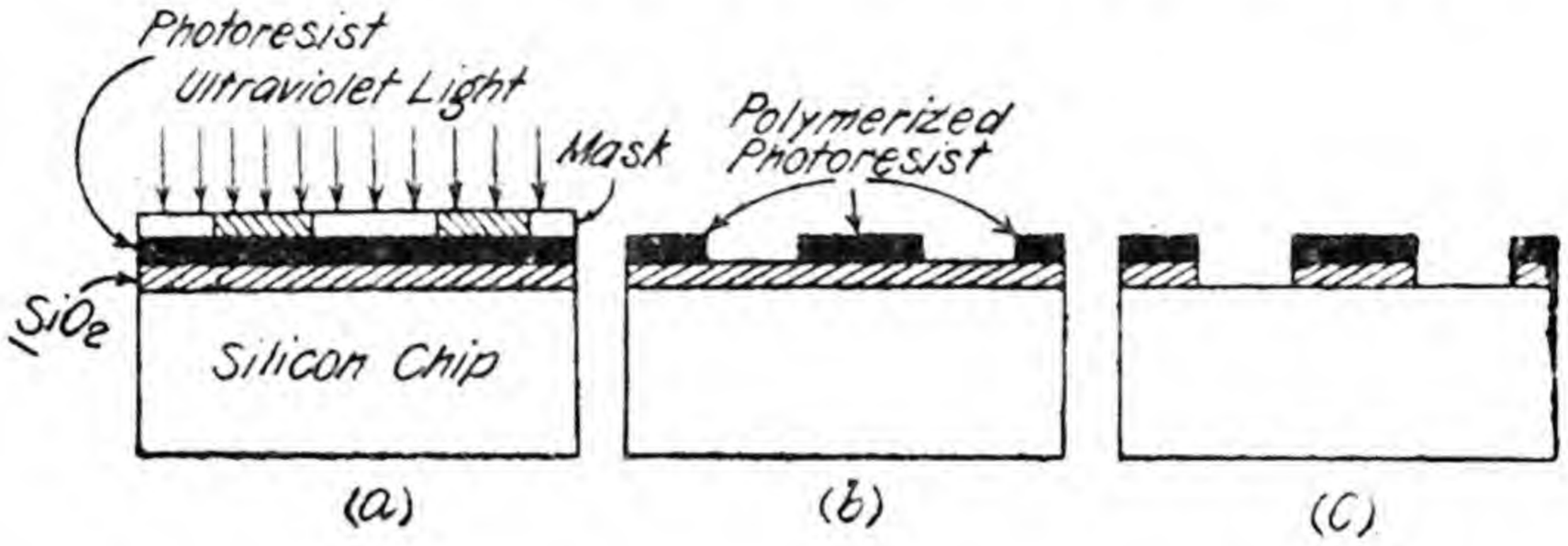


Fig. 8. Single crystal silicon substrate.

(ii) The impurity concentration is epitaxial layer may be controlled within wide limits and complex impurity profiles may be grown.

(2) Masking and Etching : In monolithic technique a selective removal of SiO_2 to form openings through which impurities may be diffused is required. For this purpose *photolithography* is used. This process is shown in fig. (9). The following steps are followed:



(*a*) Masking and exposure to ultraviolet radiation. (*b*) Photo resist after developement. (*c*) After etching

Fig. 9.

(i) The wafer is coated with a film of photosensitive emulsion (such as kodak *photoresist* KPR).

(iii) A large black and white layout of the desired pattern of openings is made and then reduced photographically. This mask (negative or stencil of required dimensions) is placed over the photoresist as shown in fig. (9*a*). Now the photoresist is exposed to ultraviolet light through the mask. Due to ultraviolet light, the photoresist becomes polymerized under transparent regions of the mask.

(iii) The mask is removed and the wafer is developed by using a chemical like trichloroethylene. The chemical dissolves the unexposed or unpolymerized portion of KPR and leaves the surface pattern as shown in fig. (9*b*).

(iv) The unremoved emulsion is now fixed and becomes resistant to the corrosive etches used next.

(v) The chip is immersed in etching solution of hydro-fluoric acid. This removes SiO_2 from the area through which diffusion has to occur. Here it should be noted that the portions of SiO_2 which are protected by photoresist are not affected by the acid

(vi) After diffusion of impurities, the KPR mask is removed with a chemical solvent like hot H_2SO_4 and by means of a chemical abrasion process.

(3) Diffusion of impurities : The diffusion of impurities into the silicon chip is the most important process used in the fabrication

of integrated circuit The process is carried out in the following two steps :—

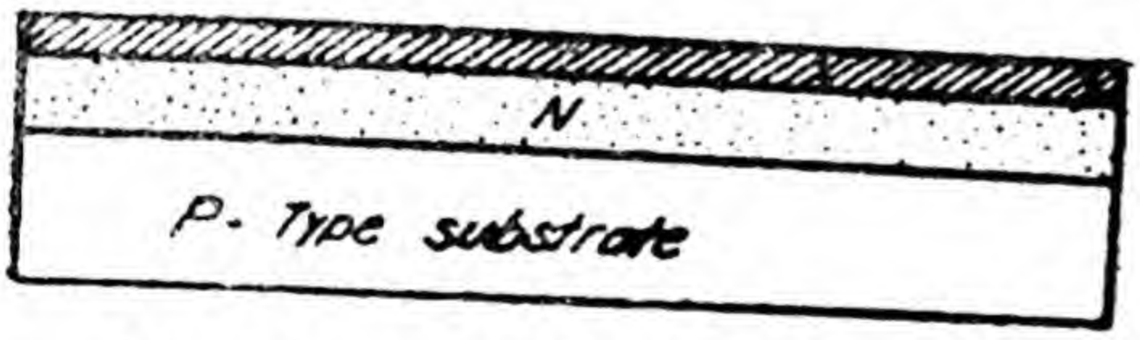
(i) The first step is called *deposition.* In deposition process the silicon slice is heated in the impurity dopant vapour so that a high concentration of dopant atoms is formed on its surface.

(ii) The second step is called *diffusion step.* In the diffusion step the dopant deposited silicon slice is moved to another furnace where it is heated to a higher temperature. Now the dopant atoms diffuse into the siiicon. The depth of penetration or diffusion and the concentration within the slice depend on time, temperature, the kind of dopant and its original concentration.

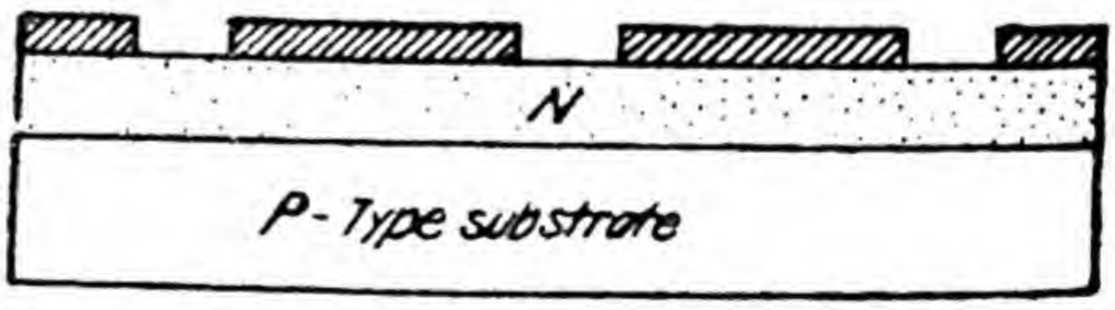
When a P type impurity is diffused into the surface of N-type slice in such a way that the density of P-type atoms exceeds the original density of N-type atoms in slice, the surface will be changed to P-type. A P-N junction is formed a small distance in from the surface where the density of diffused P-type atoms equals the or ginal N-type density. It is also possible to make a second diffusion into a region that has been formed by first diffusion. In the second diffusion the impurity concentrat on should be greater than that produced by first diffusion. Again a second P-N junction will be formed. The combined result will be a NPN structure.

(4) Isolation techniques : In the fabr cation process it is necessary to employ some means by which the co npon nts may be separated electrically This is due to the fact that the resistivity of silicon is low which results in low impedance paths between the components The *diode isolation* and *oxide isolation* are commonly used for this purpose.

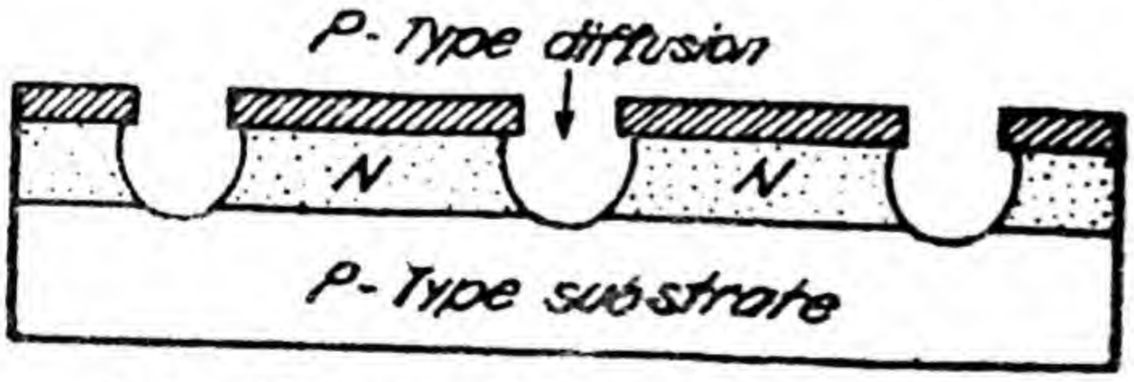
In diode isolation, the components are surrounded by a reverse biased P-N junction. Fig. (10) shows the electrical isolation of



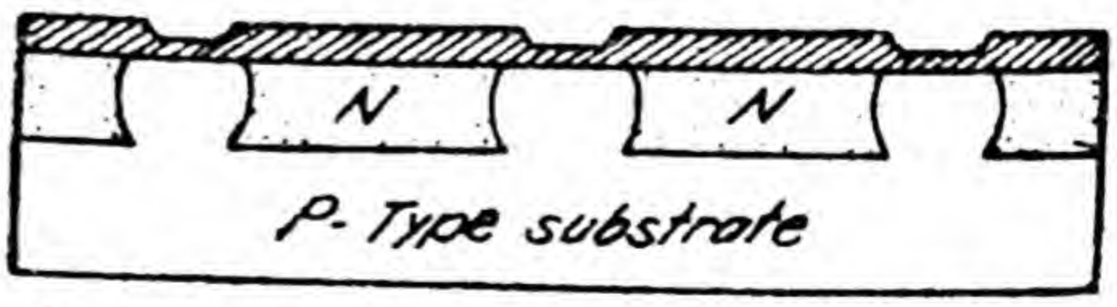
(*a*) N-type epitaxial layer grown over substrate is oxidized.



(*b*) Windows are opened on oxide by photoresist process.



(*c*) P-type diffusion taking place through windows to join up with P-type substrate.

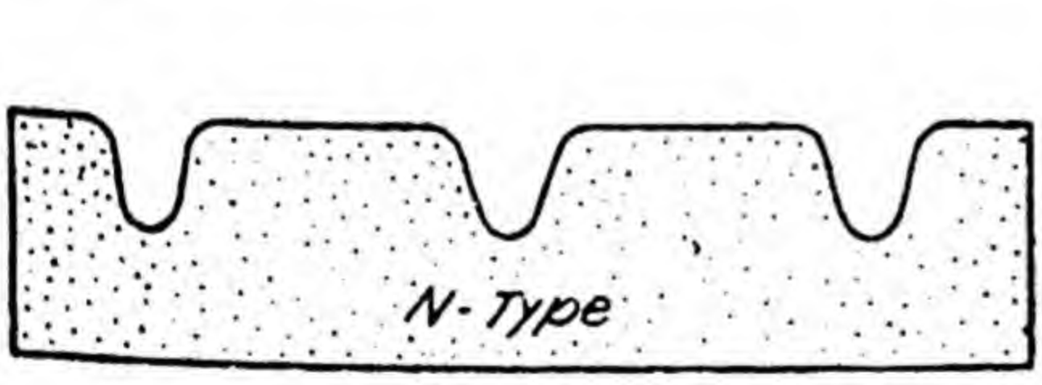


(*d*) The surface is again reoxidised.

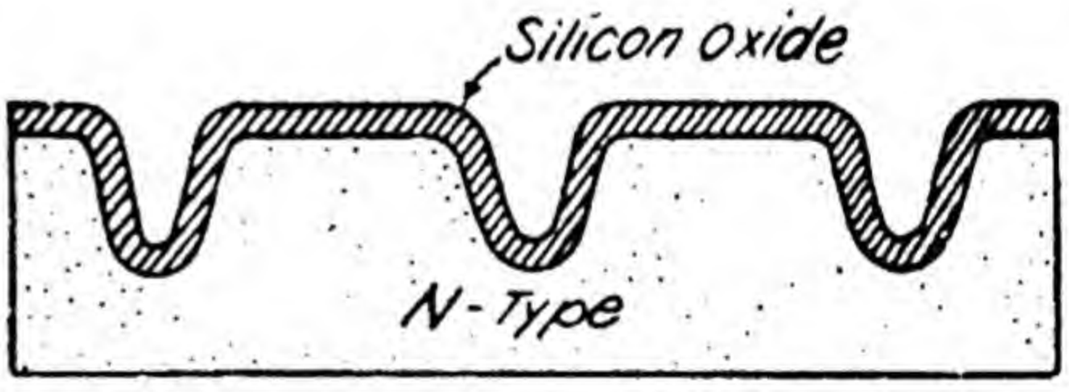
Fig. 10. Electrical isolation of integrated circuit elements by P-N junction.

integrated circuit elements by P-N junction. In this process, first of all, N-type epitaxial layer is grown on a P type substrate slice of silicon and the surface of the epitaxial layer is oxidized as shown in fig. (10*a*). Now the oxide is selectively removed from the regions where the elements are to be formed as shown in fig. (10*b*). After this the P type diffusion is carried in such a way as to join up with P-type substrate as shown in fig. (10 *c*). Finally the surface is reoxidized. In this way all the N-regions are separated from the substrate by a P-N junction When final integrated circuit is operated, each P-N junction is reversed biased by connecting the P-type substrate to a potential more negative than any part of the circuit Each junction offers a very high resistance which isolates the element formed in N-type region.

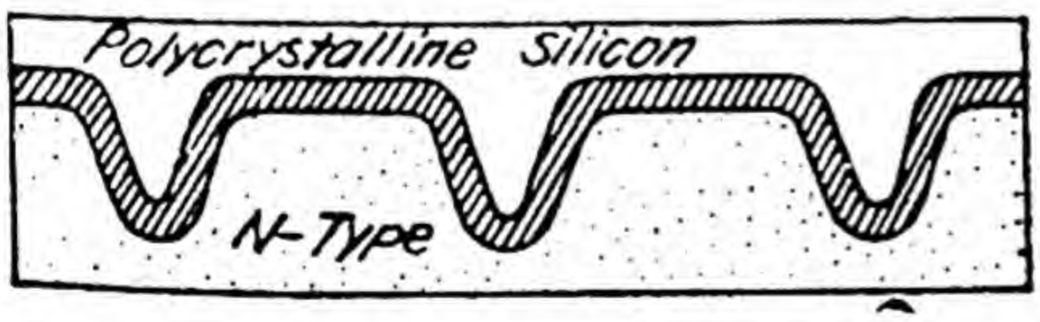
The second isolation technique is oxide isolation. In oxide isolation, a layer of silicon oxide is formed around each element of the integrated circuit. Fig. (11) shows the oxide isolation. In this process, a slice of N-type single crystal silicon is taken and deep grooves are etched in the surface at proper places where elements are to be formed as shown in fig. (11*a*). Now the whole surface is oxidized to form a continuous layer of silicon oxide as shown in fig. (11*b*). After this a uniform layer of polycrystalline silicon is deposited on the top of the oxide in an epitaxial reactor as shown in fig (11*c*). Finally the slice is inverted and original silicon is lapped down so that only the regions between the channels are left. Then the N-type areas are completely surrounded by layer of oxide. The devices can now be diffused into these isolated regions.



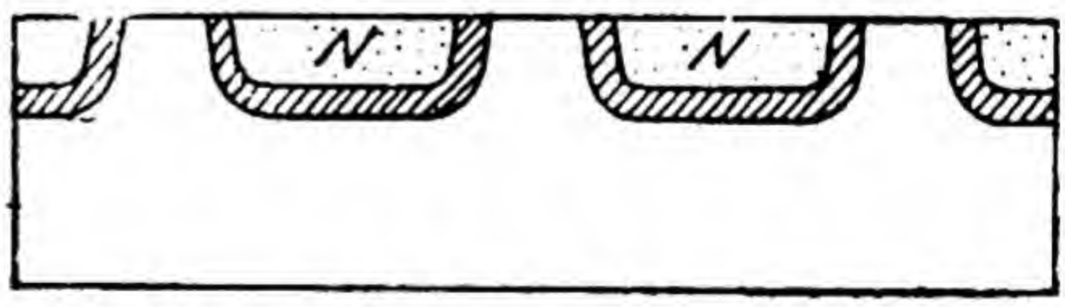
(*a*) Channels are etched at proper places.



(*b*) The surface is oxidized.



(*c*) Polycrystalline epitaxial layer is deposited on the top of the oxide.



(*d*) Slice is inverted and lapped down.

Fig. 11. Showing the electrical isolation by a layer of silicon oxide.

23·5. MONOLITHIC INTEGRATED COMPONENTS :

Integrated resistors : A resistor can be formed in a silicon wafer by diffusing a suitable impurity into a defined region. The value of the resistor so formed depends upon (i) the concentration of impurity, (ii) the dimensions of the region at the surface, and (iii) the depth to which the impurity is diffused in. In the fabrication of integrated resistors, P-type base diffusion is most commonly used, although N type emitter diffusion may also be employed. Most resistors in integrated circuits are formed at the same time as the P-type transistor base region. The diffusion layers are extremely thin so that it is more convenient to express the resistivity of the layer in terms of what is called as sheet resistance R_S of the layer.

Sheet resistance : Consider the case of a sheet of resistive material, resistivity (ρ), of length L and width w equal in length as shown in fig. (12). As shown in fig. (12), the resistance of this conductor in ohm square for the direction of current flow is given by

$$R_S = \frac{\rho L}{Ly} = \frac{\rho}{y} \qquad (1)$$

where y is the thickness.

Equation (1) shows shat R_S is independent of the size of the square and depends only on the resistivity of the material and the thickness of the layer. In order to design a resistor of 100 ohms from a material of 10 ohms/square, we merely put 10 squares in succession and make contact to the ends.

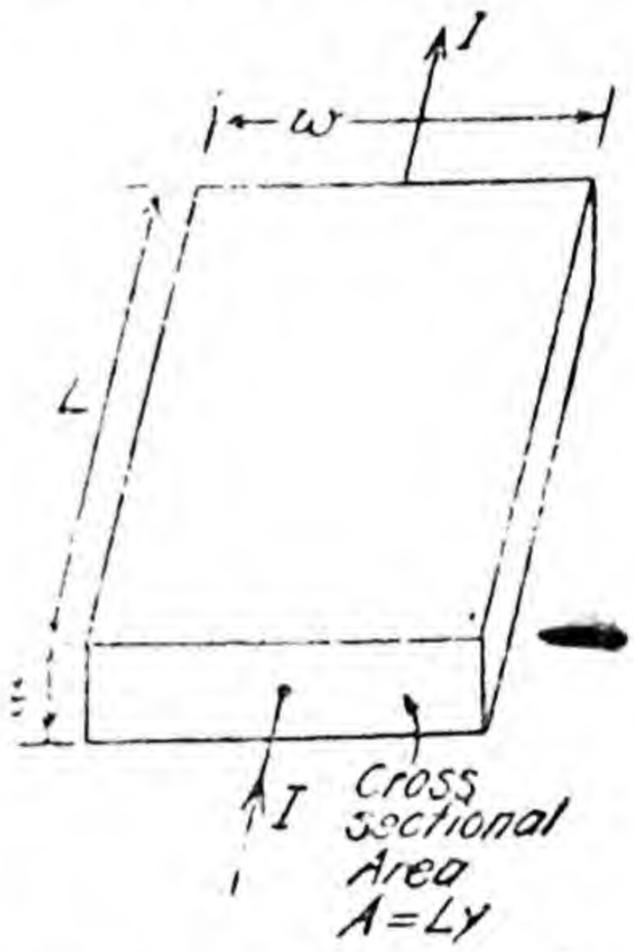


Fig. 12. Sheet of resistive material.

The construction of base diffused resistor is shown in fig. (13*a*). Figure (13*b*) gives the top view of the same. The total resistance is given by

$$R = \frac{\rho L}{yw} = R_S \frac{L}{w} \qquad \ldots(2)$$

where L and w represent length and width of diffused area respectively. Consider a base diffused resistor strip 2 mils wide and 10 mils long. The strip contains 20 (one mil by one mil) square and its resistance is

$$R_S \times 20 = 200 \times 20 = 4000\,\Omega.$$

Corrections for the end contacts are usually included in calculation of R

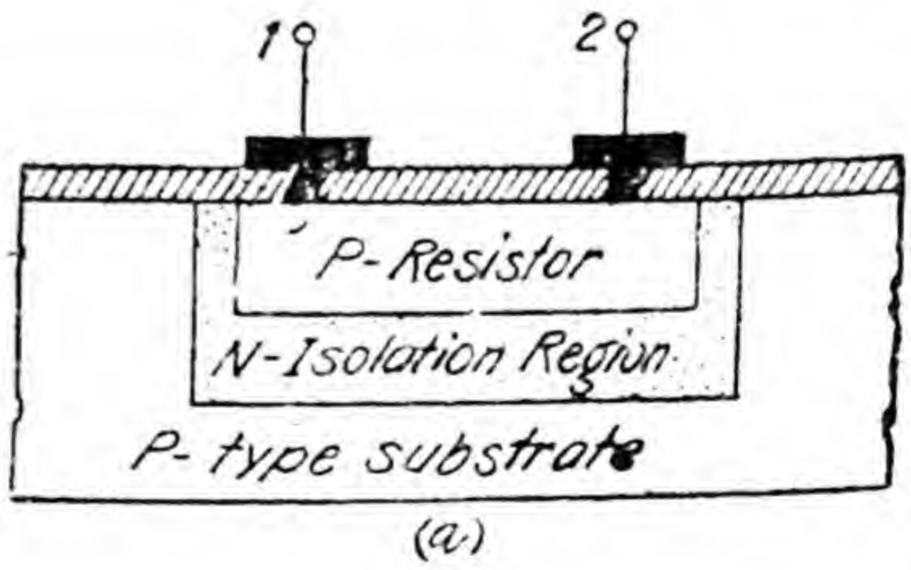


Fig. 13 (*a*) Cross sectional view. (*b*) Top view.

Values of monolithic resistors : Since the sheet resistance of the base emitter diffusions is fixed, hence only variables available in the design of diffused resistors are : (i) the strip length L and (ii) the strip width w. Strip widths of less than one mil is not normally used because with $w = 1$ mil a line width variation of 0·1 mil due to mask drawing error or due to mask alignment or due to photographic resolution error may result in 10% resistance tolerance error. In practice monolithic base diffusion resistors have values from 20Ω to 30 $K\Omega$ when high values of resistance are required, an alternative to increasing the strip length is to reduce the P type thickness and effective concentration by diffusion into it an N-type region at the same time as the transistor emitter diffusion. This method is however more difficult to control than increasing the length of the strip.

Equivalent circuit : Fig. (14) shows the equivalent circuit of a diffused resistor. This circuit includes (i) parasitic capacitance C_1 of base isolation junction, (ii) parasitic capacitance C_2 of isolation substrate junction and (iii) a parasitic P-N-P transistor, with P-type substrate acting as collector, N-type isolation region acting

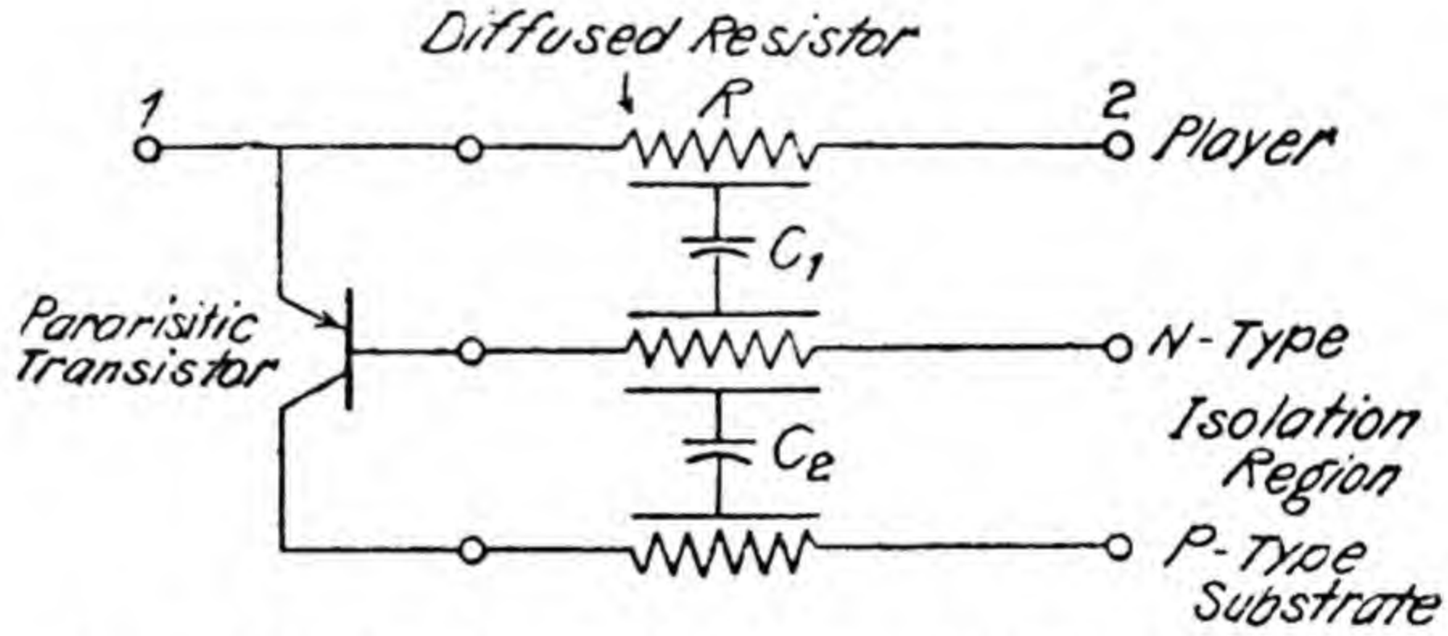
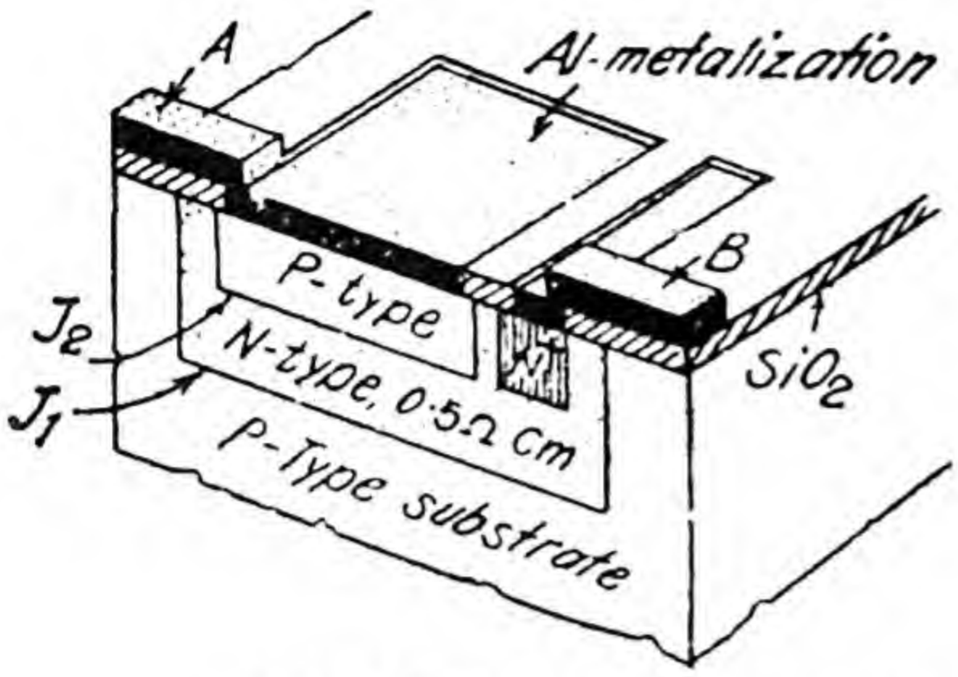


Fig. 14. Equivalent circuit of diffused resistor.

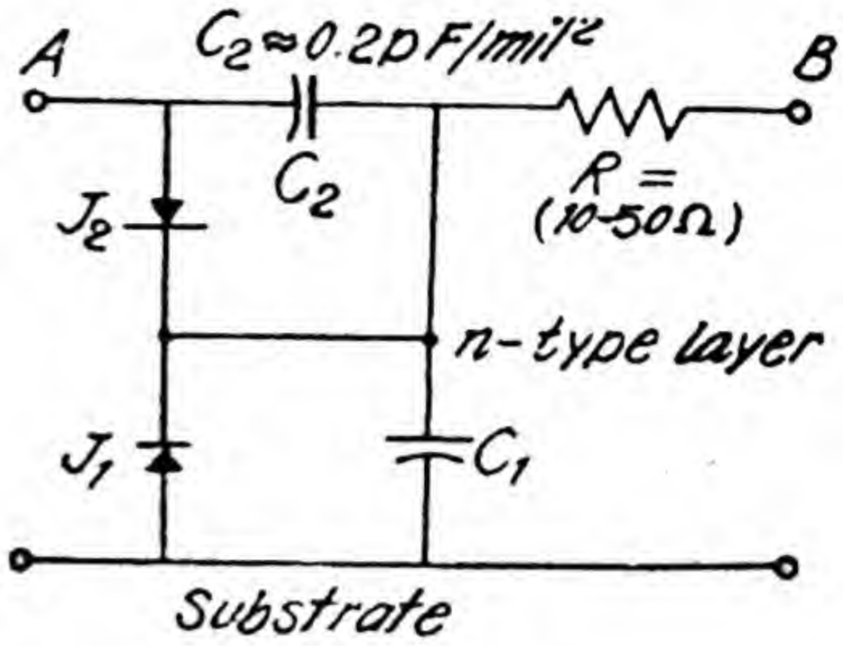
as base and P-type resistor material acting as emitter. In order to maintain the parasitic transistor cut-off, it is essential to keep the emitter reversed biased. This is achieved by placing all the resistors in same N-type isolation region and connecting this isolation region to most positive voltage present in the circuit.

Integrated capacitors : For monolithic integrated circuits, the capacitors are formed either by a junction technique or by a thin film technique. A junction capacitor uses the capacitance of reversed biased P-N junction which can be formed at the same time as the emitter junction or the collector junction of the transistor. The two techniques are discussed below :

Junction capacitors – Fig (15) shows a junction monolithic capacitor. In fig. (15 *a*) a cross sectional view is shown while in fig. (15*b*) its equivalent circuit is shown. The capacitor is formed by a reversed biased junction J_1 between the epitaxial N-type layer and the P-type diffusion area. Here one more junction J_2 exists between N-type epitaxial layer and P-type substrate which is associated with a parasitic capacitance C_2. In the equivalent circuit C_1



(*a*) Cross-sectional view

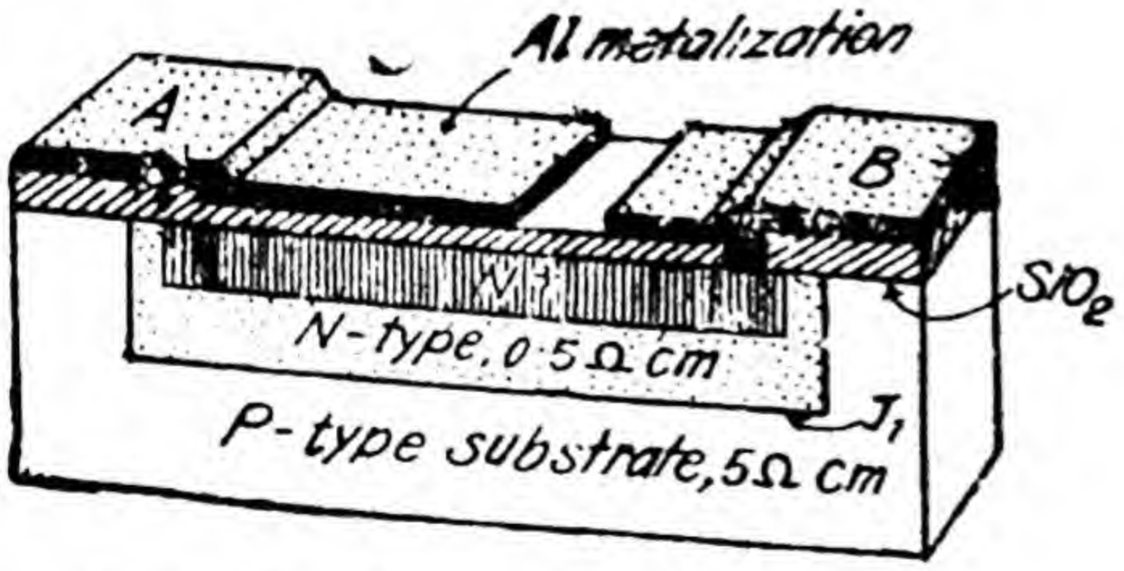


(*b*) Equivalent circuit.

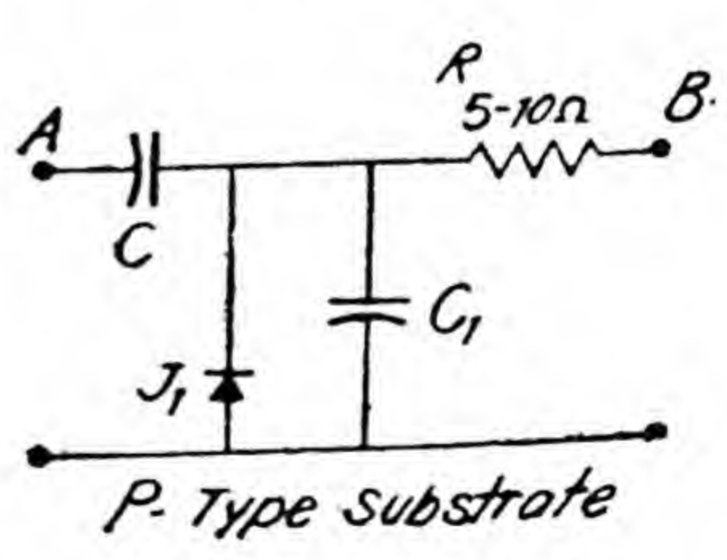
Fig 15. Junction monolithic capacitor.

forms the desired capacitor which is made as large as possible relative to C_2. The value of this capacitor C_1 depends upon (i) area of the junction and (ii) impurity concentration. The junctional capacitance is given by A/W, where A is the area of junction and W is the total space charge width of the junction. In figure, R represents the resistance of P-type diffusion layer Here two things should be remembered *i e.*, (i) to minimise the parasitic capacitance C_2, the P-type substrate should be at most negative potential and (ii) the junction J_2 should be reversed biased so that capacitor C_2 may be isolated from the rest of the circuit.

Thin film MOS (*metal oxide semiconductor*) *capacitors* – Fig. (16) shows the *MOS* capacitor. In figure (16*a*), the cross sectional view is shown while in fig. (16 *b*). the equivalent circuit is shown. Basically a *MOS* is a parallel plate capacitor with SiO_2 as the dielectric. An N^+ region is diffused into the silicon at the same time as the transistor emitter diffusion to form the bottom electrode of the capac tor. The controlled thickness of silicon oxide is formed on this surface of this region to give the dielectric. The top electrode consists of a layer of metal (aluminium) deposited at the same time as the interconnection pattern. For oxide thickness of 500Å. the capacitance is typically 0·3 PF/mil² and varies inversely with oxide thickness.



(a) Cross-sectional view.



(b) Equivalent circuit

Fig. 16. *MOS* capacitor.

23·6 TRANSISTORS OF MONOLITHIC INTEGRATED CIRCUITS :

Transistors of monolithic integrated circuits are formed in the epitaxial layer by successive diffusions. First of all an epitaxial

layer (N-type) is formed on the substrate (P-type) and then it is covered by an oxide layer about 1 micron thick The oxide layer is then coated by a photosensitive material and exposed to light through a mask. The regions where the photosensitive material is exposed – harden on developing while regions not exposed are comparatively soft. When this surface is exposed to hydroflouric acid, the oxide layer not protected by the hardened photoresist. gets dissolved forming windows in the protective oxide layer The crystal is now exposed to a third group impurity at high temperature. The epitaxial layer is now divided into a number of islands. The N-type epitaxial islands are isolated from one another by two reversed biased junctions NP-PN. Transistors are now formed in the islands by repeated diffusion. One island is masked and boron is diffused in to form the P-type base region. The impurity extends to about 10 microns having about 10 microns thick N layer below, which functions as the collector After base diffusion, the crystal is masked again and phosphorous is diffused in to form the high concentration N^+ type emitter region At the same time, another N^+ region is diffused into N–type collector region so that a low resistance contact to the collector region can be made. Monolithic planer epitaxial NPN transistor is shown in fig (17) Since the anode of the isolation diode covers the back of the entire wafer, it is necessary to make the collector contact on the top as shown in fig. (17). The isolation diode of this transistor has three undesirable effects. They are (i) it produces a parasitic shunt capacitance to the collector, (ii) it provides a leakage current path and (iii) the collector contact at the top increases the length of the collector current path and hence increases the collector resistance. All these undesirable effects are absent in discrete planar expitaxial transistor. The discrete planar epitaxial transistor is shown in fig. (18). In this transistor collector contact is made at the bottom of the N^+ substrate. Here it should be remembered that although discrete epitaxial transistor is superior to monolithic epitaxial transistor but there are so many other superiority of monolithic epitaxial transistor over discrete

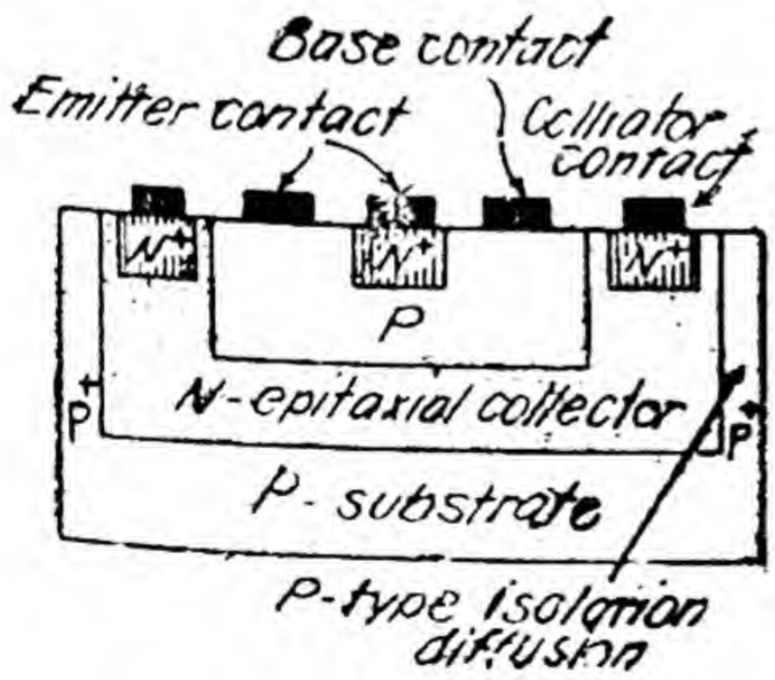


Fig 17. Monolithic integrated circuit NPN transistor.

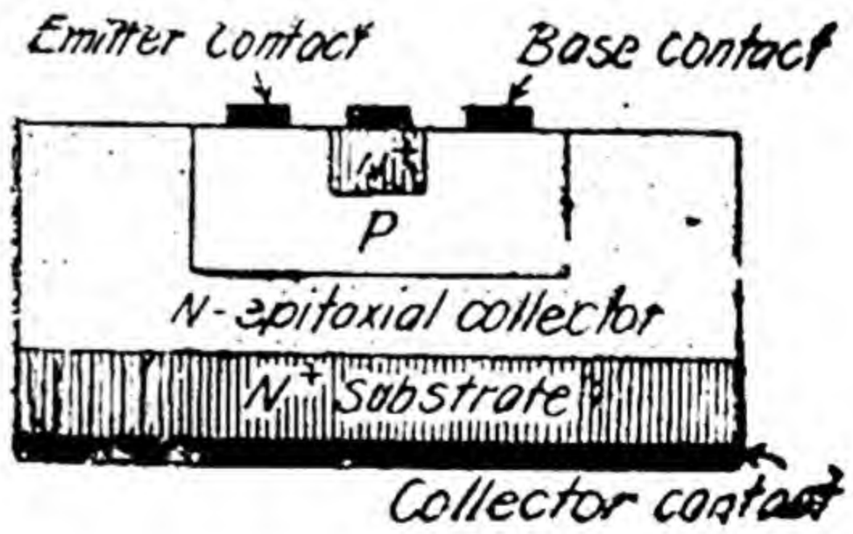


Fig. 18. Discrete epitaxial transistor.

epitaxial transistor. The main superiority are : (i) integrated transistors are located physically together and (ii) their electrical characteristics are closely matched.

The collector series resistance in monolithic epitaxial transistor may be reduced by placing a heavily doped N^+ buried layer sandwiched between the P-type substrate and N-type epitaxial collector as shown in fig. (19). The buried layer may be formed by diffusing the N^+ layer into P-type substrate before the N-type epitaxial collector is grown. This layer may also be obtained by selectively growing the N^+ type layer using masked epitaxial technique.

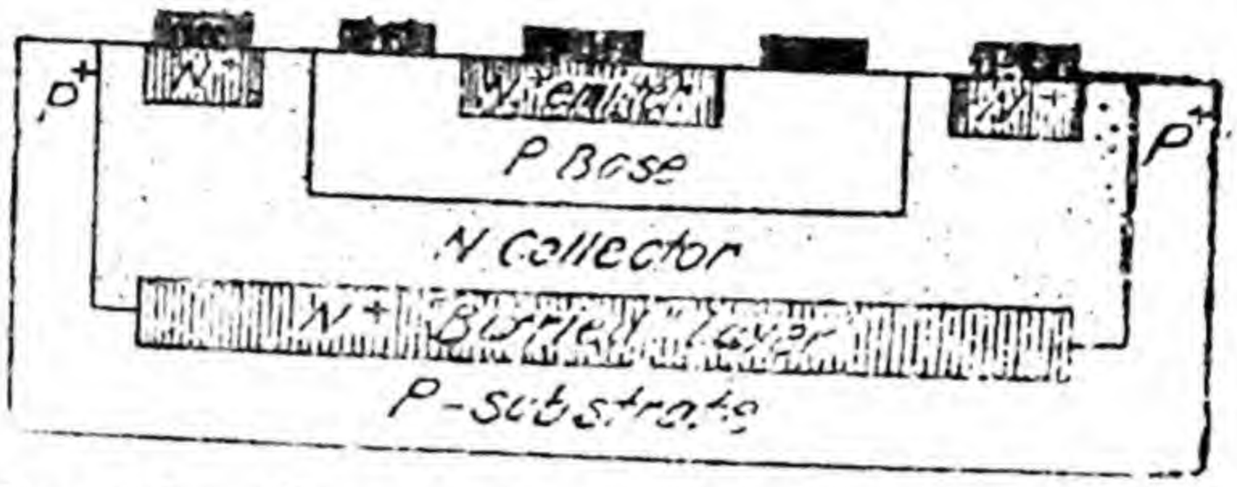


Fig. 19. Transistor with buried layer.

23 7. MONOLITHIC DIODES :

Integrated circuit diodes are formed by making use of the basic structure of integrated N-P-N transistor. Although five different structures are possible but the following three configurations are widely used :

(i) Using the emitter base diode with collector short circuited to base as shown in fig. (20 *a*).

(ii) Using the emitter-base diode with collector open as shown in fig. (20 *b*).

(iii) Using the collector base diode with emitter open (or not fabricated at all) as shown in fig. (20 *c*).

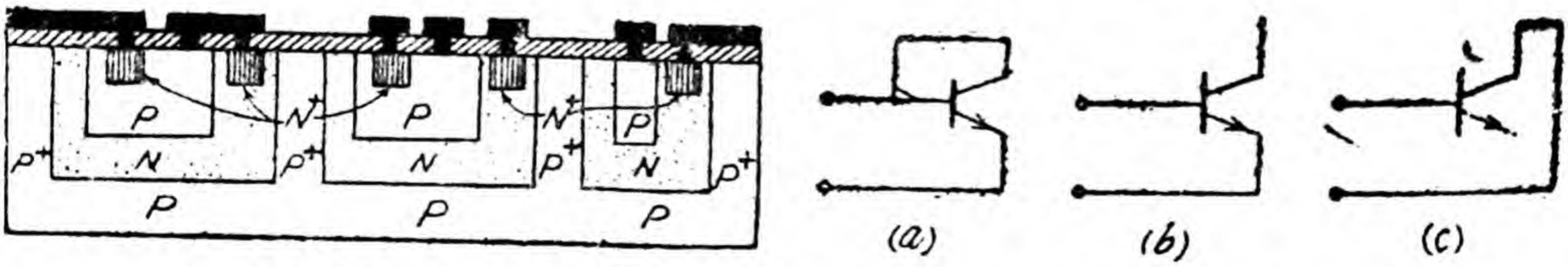


Fig. 20. (*a*) Emitter-base diode with collector short circuited to base.
(*b*) Emitter-base diode with collector open.
(*c*) Collector-base diode with emitter open.

Depending upon the circuit requirement, any configuration out of the three may be fabricated.

In configuration 20 (*a*), the collector and base regions are connected together. Here no charge can be stored in the collector base region. There is a large stray capacitance from anode to substrate. Since base-emitter junction is used, the breakdown voltage is low.

The configuration 20 (*b*) has a low breakdown voltage due to low resistivity required for the emitter of the transistor. The leakage current path R_{B-E} and junction capacitance C_{B-E} are that of the base emitter junction alone. The voltage drop in the forward direction is mostly caused by the series resistance in the base region.

The configuration 20 (*c*) provides high breakdown voltage and low leakage current of the base-collector junction. The high resistivity of the collector region and the additional charge stored in the floating base-emitter junction increase the switching time considerably. It is possible to omit the emitter diffusion, thus saving space.

EXERCISES AND PROBLEMS

1. What are the advantages of integrated circuits over conventional circuits ? What are their drawbacks ?
2. List the different steps involved in fabricating a monolithic integrated circuit.
3. Name and explain different processes involved in the fabrication of monolithic integrated circuits.
4. (*a*) Describe a photoetching process.
   (*b*) How many masks are required to complete an IC ? List the function performed by each mask.
5. Explain how isolation between components is obtained in an IC.
6. Define sheet resistance. What is order of magnitude of this resistance for the base region and for the emitter region ? Sketch the cross section of an IC resistor.
7. Sketch the cross-section of a junction capacitor. Draw equivalent circuit showing all parasitic elements.
8. Sketch the equivalent circuit of a base-diffused resistor, showing all parasitic elements What must be done externally to minimise the effect of parasitic elements ?
9. How will you fabricate an NPN transistor ?
10. Draw different integrated diode configurations. Which configuration is having (i) highest junction capacitance, (ii) highest parasitic effect, (iii) highest leakage current and (iv) highest storage time ?

Select correct answers of the following questions :

11. In a monolithic integrated circuit—
    (*a*) All circuit elements are formed in a single wafer of semiconductor material.
    (*b*) Each circuit element is fabricated as a separate wafer
    (*c*) Active circuit elements are assembled as separate wafers.

    [Ans (*a*)].

12. Photomasking :
    (*a*) Controls the depth of diffusion.
    (*b*) Is used in process to remove selected regions of silicon oxide.
    (*c*) Reduces the size of the circuit elements. [Ans. (*b*)].
13. Transistors of monolithic integrated circuits :
    (*a*) Use the isolation junction as the collector junction.
    (*b*) Are made as separate wafer.
    (*c*) Are similar to discrete planar transistors, but have the collector contact on the top surface. [Ans. (*c*)].
14. Capacitors for integrated circuits :
    (*a*) Can be made using silicon oxide as the dielectric.
    (*b*) Are not possible.
    (*e*) Can not be made using diffusion techniques. [Ans. (*a*)]

24 LINEAR INTEGRATED CIRCUITS : OPERATIONAL AMPLIFIER (OP-AMP)

Basically operational amplifier consists of a very high gain d c. amplifier with feedback, having high input impedance, a low output impedance, and acting as a differential amplifier Operational amplifier, originally used to perform mathematical functions such as addition, integration, differentiation etc. in analogue computers are now put to a variety of other uses, *e.g.*, as comparator, pulse generator, square wave generator, Schmitt trigger etc.

These days operational amplifier uses integrated circuit technology and is referred to as basic linear or analogue integrated circuit. IC OP-AMP are widely used as they possess all the me its of monolithic integrated circuits, *e.g.*, small size low cost, high reliability, low offset voltage and current, and temperature tracking properties.

24·1. BASIC CONCEPTS :

As stated earlier, OP-AMP is basically a difference amplifier whose basic function is to amplify the difference between two input signals. The advantage of using differential amplifier in OP-AMP is its rejection capability of unwanted signals.

Ideal Operational Amplifier : The ideal operational amplifier is shown in fig. (1 *a*) and its low frequency equivalent in fig. (1 *b*). A signal appearing at the negative terminal (1) is inverted at the output and is called *inverting terminal* while a signal at the positive terminal (2) appears at the output without any change in sign and is called *non-inverting terminal*. In general, the output voltage is directly porportional to the input voltage which is difference of V_1 and V_2 *i.e.* $V_t = V_2 - V_1$. $(-A)$ is the voltage gain of the amplifier.

Ideal OP-AMP has the following characteristics :

(i) Infinite input impedance

$$Z_t = \infty$$

(ii) Zero output impedance

$$Z_0 = 0$$

(iii) Infinite voltage gain,

$$A = -\infty$$

(iv) Infinite bandwidth, $BW = \infty$

(v) Perfect balance, $V_0 = 0$ when $V_2 = V_1$

(vi) Zero drift, *i.e.*, characteristics do not drift with temperature.

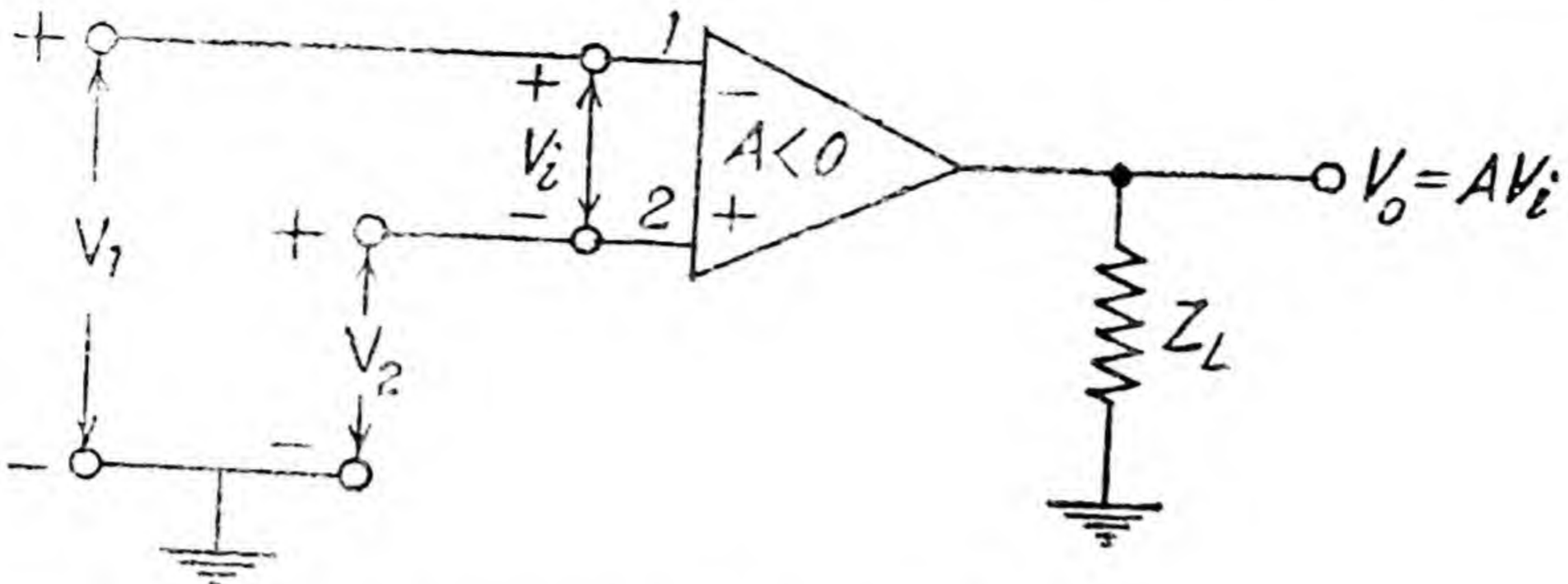


Fig. 1 (*a*) 1—Inverting terminal.
2—Non inverting terminal.
V_i—Input voltage ($=V_2-V_1$)
A—Voltage gain under load conditions.

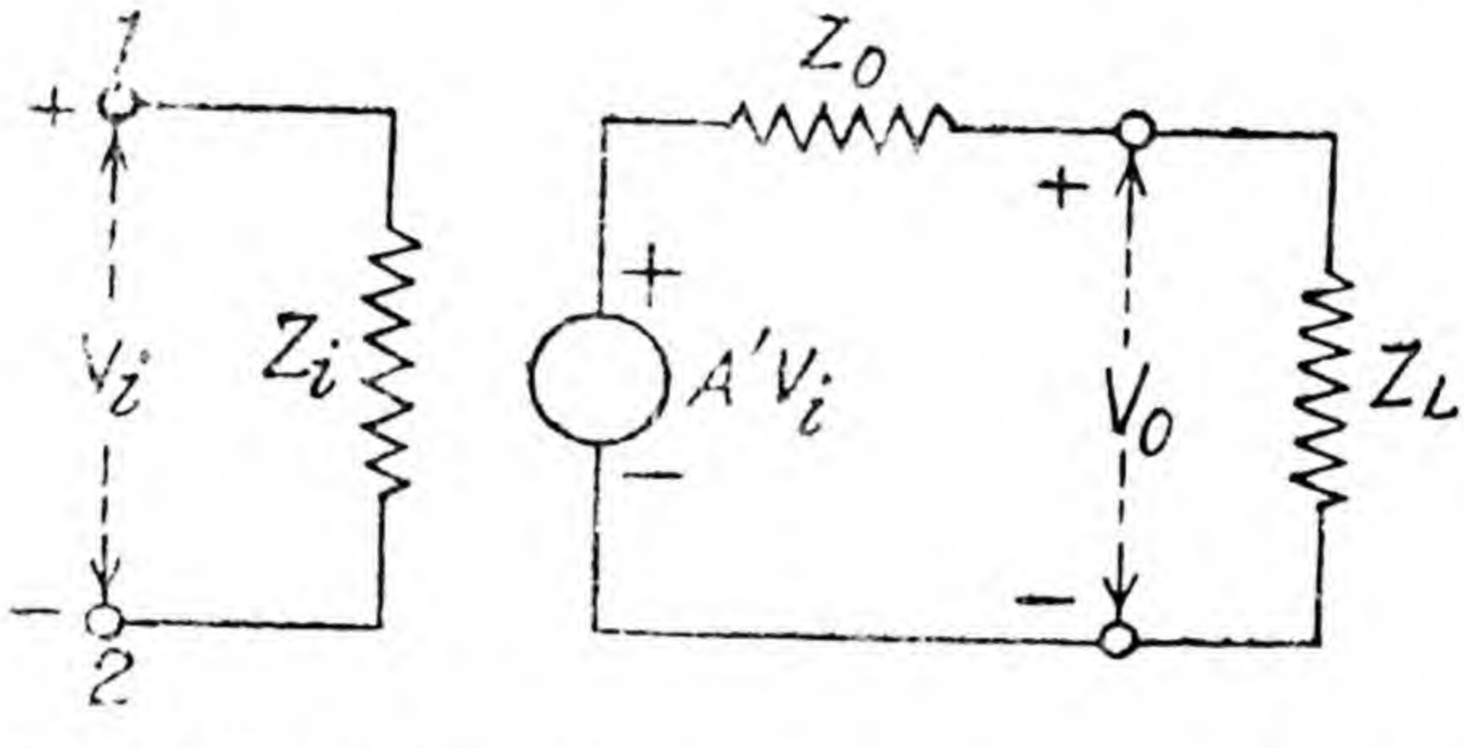


Fig. 1 (*b*) Low frequency equivalent circuit of OP-AMP.
A'—Open circuit voltage gain.

Basic Inverting OP-AMP :

Circuit diagram of basic inverting OP-AMP is shown in fig. 2 (*a*) and its equivalent is shown in fig. 2 (*b*). In this mode of operation, the positive input terminal of the amplifier is grounded and the input signal is applied to the negative input terminal

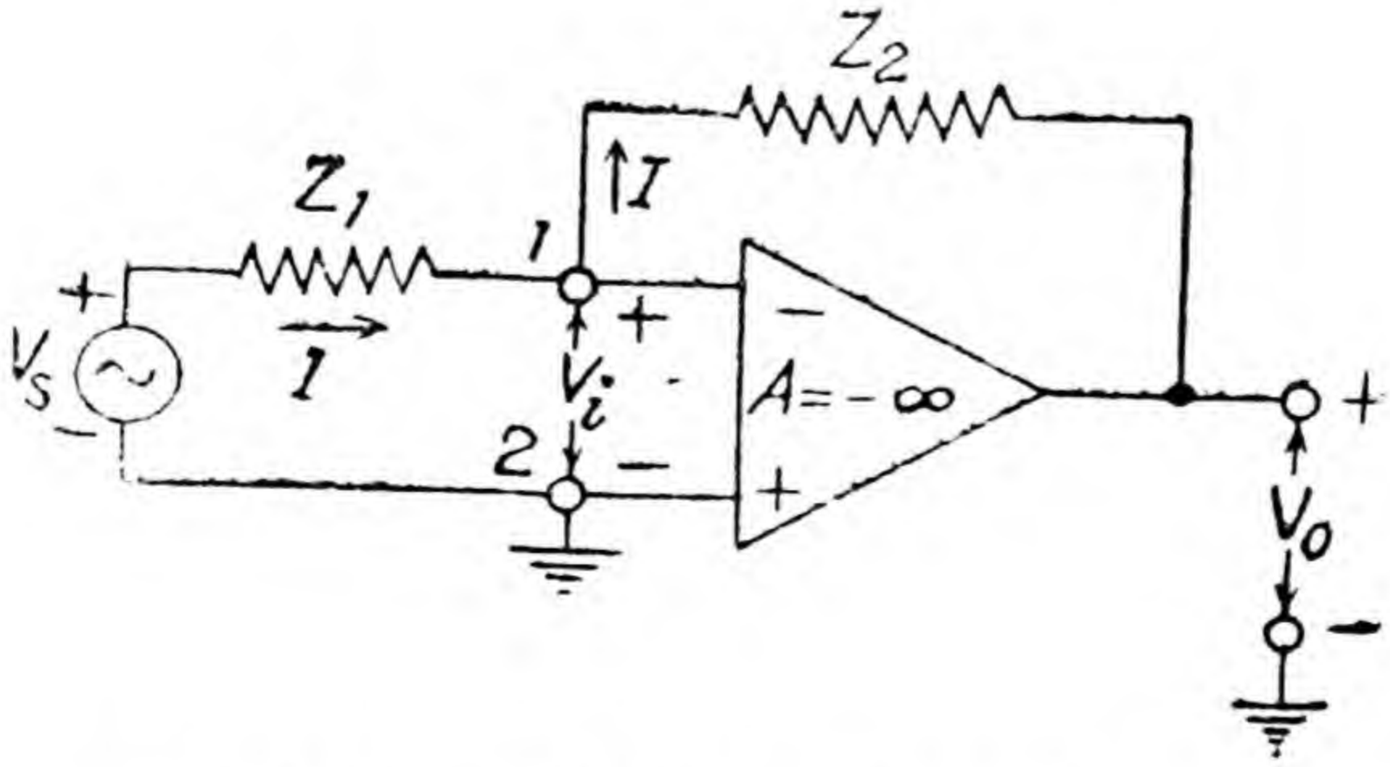


Fig. 2 (*a*) 1—Inverting terminal.
2—Non-inverting terminal.

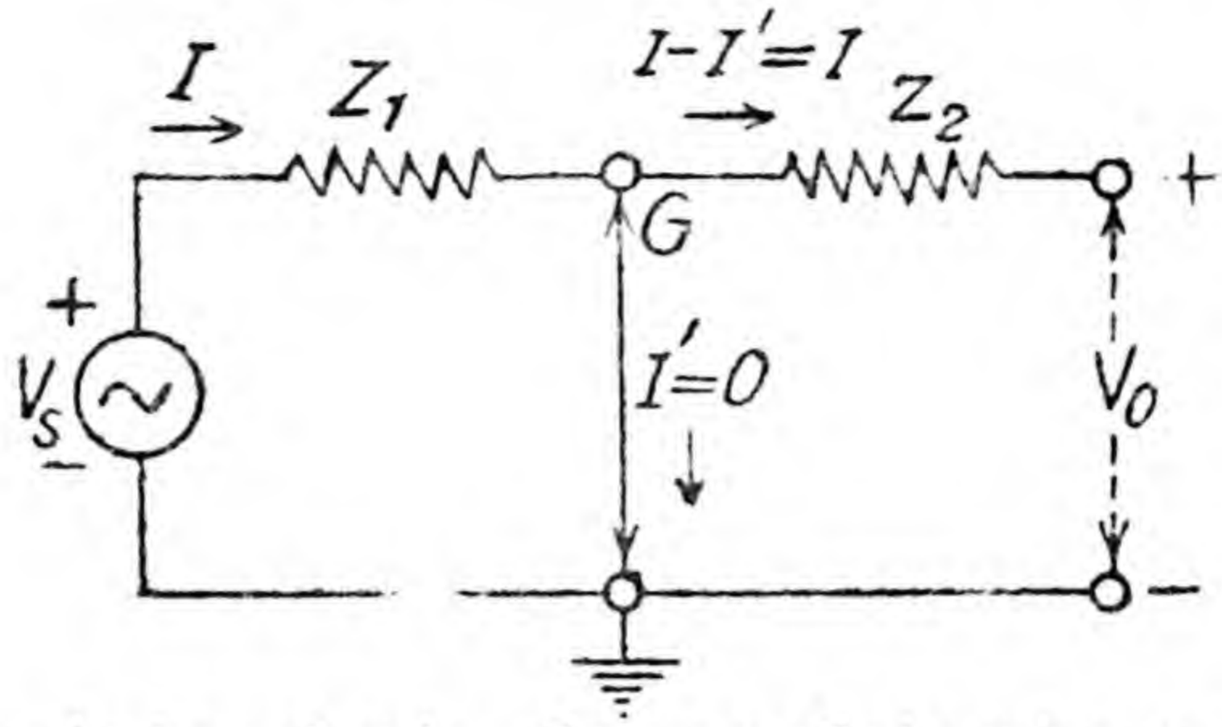


Fig. 2 (*b*) Equivalent circuit : Concept of virtual ground in OP AMP. Virtual ground exists at *G*.

through impedance, Z_1. The feedback applied through the impedance, Z , from the output to input terminal is negative. Feedback impedance, Z_2 and input impedance, Z . determine the inverting operation of the amplifier as we shall infer from eq (2).

As OP-AMP is considered as ideal, it will have infinite voltage gain, A. But

$$|A| = \frac{V_0}{V_i}$$

Therefore with finite value of V_0, the output voltage, gain A can be infinite only if input voltage, V_i, is zero. The negative feedback from output to input through Z_2 serves to keep the voltage V_i, at zero. Thus terminal, 1, will be at the potential of terminal, 2. As terminal 2 is grounded, terminal, 1, is thus *virtually grounded*, and no current ($I'=0$) flows from input terminal, 1 to input terminal 2. Current I flowing through Z_1 will also flow through Z_2 (since input impedance of an ideal OP-AMP is infinite, $Z_i=\infty$). Thus

$$\text{Current through } Z_1 = I = \frac{V_S - V_i}{Z_1}$$

$$\text{Current through } Z_2 = I = \frac{V_i - V_0}{Z_2}$$

so that

$$\frac{V_S - V_i}{Z_1} = \frac{V_i - V_0}{Z_2}$$

or

$$\frac{V_0}{Z_2} = \frac{V_i}{Z_2} + \frac{V_i}{Z_1} - \frac{V_S}{Z_1}$$

$$= V_i\left(\frac{1}{Z_2} + \frac{1}{Z_1}\right) - \frac{V_S}{Z_1}$$

But

$$A = -\frac{V_0}{V_i} \text{ so that } V_i = -\frac{V_0}{A}.$$

Substituting in above equation we get

$$\frac{V_0}{Z_2} = -\frac{V_0}{A}\left(\frac{1}{Z_2} + \frac{1}{Z_1}\right) - \frac{V_S}{Z_1} \qquad \ldots(1)$$

or $$V_0\left[\frac{1}{Z_2}+\frac{1}{A}\left(\frac{1}{Z_2}+\frac{1}{Z_1}\right)\right]=-\frac{V_S}{Z_1}$$

or $$\frac{V_0}{V_S}=-\frac{Z_2}{Z_1}\frac{1}{\left[1+\frac{1}{A}\left(1+\frac{Z_2}{Z_1}\right)\right]}$$

where $\frac{V_0}{V_S}=A_f$ is called closed loop amplification (with feedback) while A is called open loop amplification (without feedback).

Let us approximate the values of the two terms on right hand side of above equation (1). V_0 is typically 10V maximum, A is atleast 10^4, Z_2 and Z_1 are often in megaohm range (never less than 10 $K\,\Omega$). Thus

$$\text{first term}=-\frac{V_0}{A}\left(\frac{1}{Z_2}+\frac{1}{Z_1}\right)=\frac{-10}{10^4}\left(\frac{1}{10^4}+\frac{1}{10^4}\right)=-2\times10^{-7}$$

$$\text{second term}=-\frac{V_S}{Z_1}=\frac{-10}{10^4}=-10\times10^{-3}$$

We can then neglect first term in comparison to second term in eq. (1) to write

$$\frac{V_i}{Z_2}=-\frac{V_S}{Z_1}$$

or $$\frac{V_0}{V_S}=-\frac{Z_2}{Z_1}, \qquad \ldots(2)$$

and is true so far as gain A is high enough.

Here $\frac{V_0}{V_S}$ is referred to as the *closed loop gain of the inverting amplifier*. It is a negative quantity because closed loop amplifier reverses the sign of input voltage, *i.e.* output is out of phase with input. It clearly depends on the ratio of feedback impedance, Z_2 and input impedance, Z_1.

Input impedance : Since $V_i=0$ due to feedback, we have

$$Z_i=\frac{V_S}{I}=Z_1$$

Which predicts that input impedance of the amplifier depends only on the external impedance Z_1.

Output impedance : It is defined as the impedance seen at the output of the amplifier when the input terminal is set equal to zero. For ideal OP AMP it is zero.

Thus we have observed that in inverting OP-AMP circuits :

(*i*) *Current into each input terminal of the amplifier is zero,*

(*ii*) *potential difference between input terminals is zero.*

(*iii*) *a virtual short circuit exists at input terminals.*

Practical inverting OP-AMP : Let us consider an amplifier which does not satisfy the above conditions of $A=\infty$, $Z_i=\infty$ and $Z_0=0$ of ideal OP-AMP. Let us take that

$$A\neq\infty,\ Z_i\neq\infty \text{ and } Z_0\neq 0.$$

Small signal model of practical inverting OP-AMP is shown in fig. 3. Using Miller's theorem, effect of feedback impedance, Z_2, on the input and output of the amplifier is accounted by replacing it by two impedances, *viz.* $\frac{Z_2}{(1-A)}$ across the input and $\frac{Z_2A}{(A-1)}$ across the output.

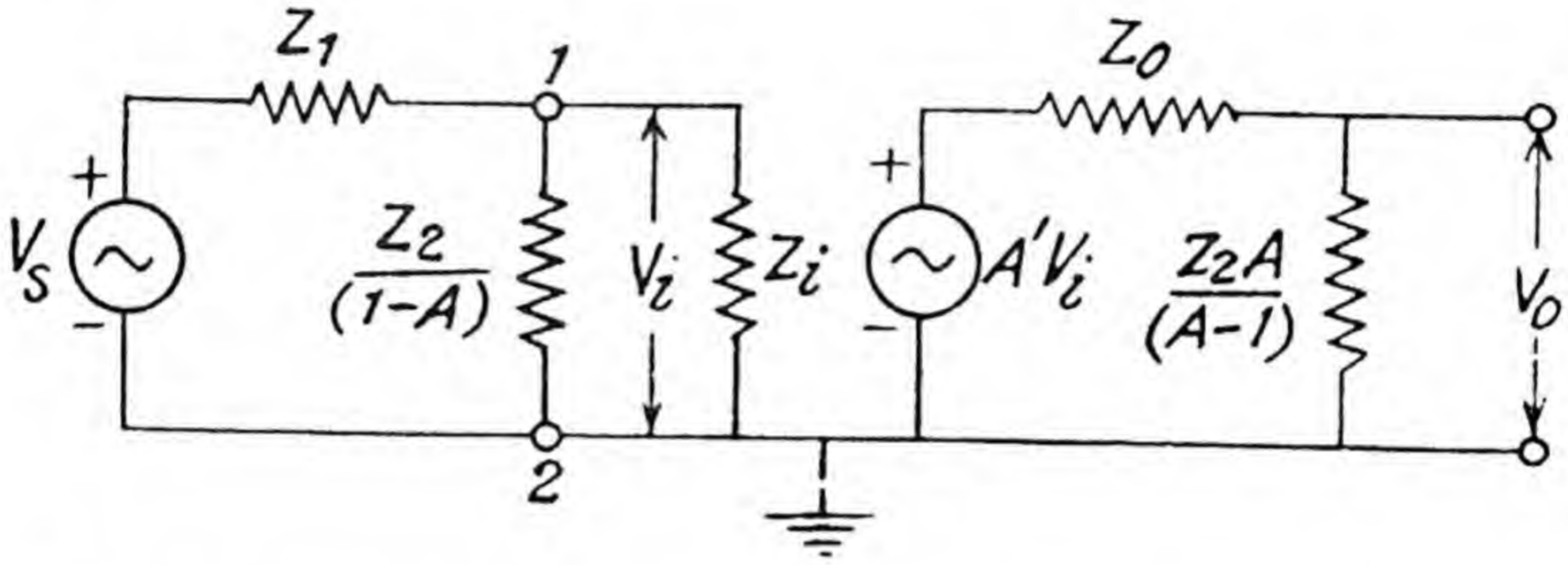


Fig. 3. A' is open circuit (or unloaded) voltage gain.
Z_2 is not shown as it has been accounted.

We can show that for such a circuit

$$\frac{V_0}{V_S}=\frac{-Z_2/Z_1}{\frac{1}{A}\left(1+\frac{Z_2}{Z_1}+\frac{Z_2}{Z_i}\right)-1} \qquad ..(3)$$

and

$$A=\frac{V_0}{V_i}=\frac{A'+Z_0/Z_2}{1+Z_0/Z_2}$$

If $Z_0=0$, *i.e.* loading is effectively removed, then above relation gives that $A'=A$. If $|A'|\to\infty$ then $|A|\to\infty$ so that from eq. (2), we have

$$\frac{V_0}{V_S}=-\frac{Z_2}{Z_1}$$

which is same as eq. (2). Thus for high gain OP-AMP output voltage is

$$V_0=-\left(\frac{Z_2}{Z_1}\right)V_S$$

which clearly shows that output is dependent on ratio Z_2/Z_1. Thus feedback impedance Z_2 and input impedance Z_1 determine the precision of the operation to which this amplifier is put.

Practical Noninverting OP-AMP : In noninverting OP-AMP output is equal to and in phase with the input voltage. Like an emitter follower, here the source and the load are *effectively isolated, i.e.*, $Z_i=\infty$ and $Z_0=0$. As input impedance $Z_i=\infty$, no current

flows into either input terminals of the OP-AMP. That is $I_1=0$ and $I_2=0$. The circuit diagram is shown in fig. 4. The input signal, V_S, is applied directly to the non-inverting terminal 2, so no phase inversion results at the output. Some of the output is fedback to the inverting input providing effectively some voltage, V_1, at the inverting terminal of the same polarity as at the noninverting input.

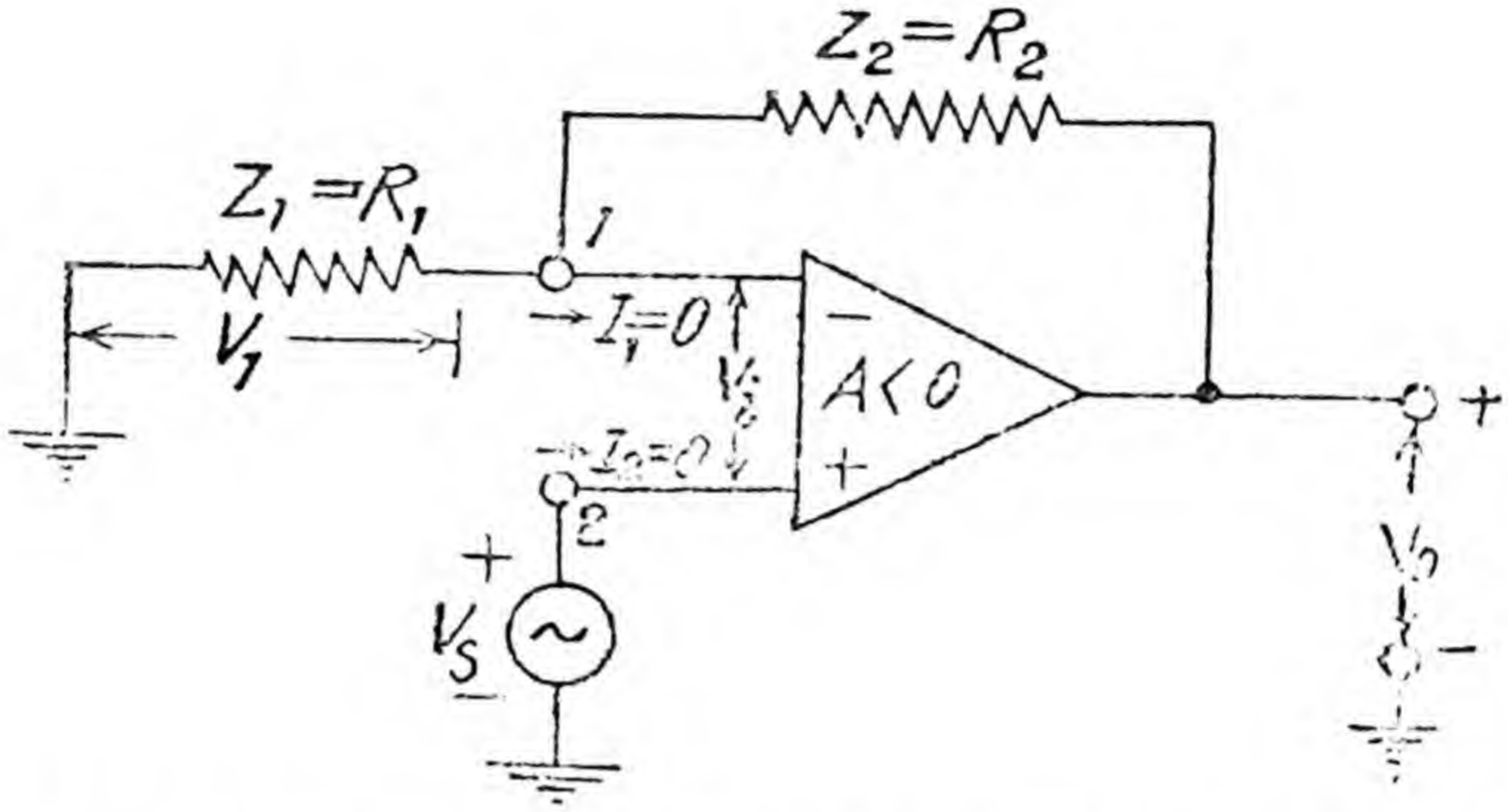


Fig 4. Non inverting OP-AMP Feedback impedance $Z_2=R_2$ and input impedance $Z_1=R_1$. Output voltage follows input voltage $V_0\cong V_S$ and $V_i=0$.

With $Z_i=\infty$, $A=-\infty$ and $V_0=A(V_1-V_S)$ we conclude that for finite V_0 (while A is infinity),

$$V_1-V_S=0 \qquad \text{or} \qquad V_1=V_S$$

Thus like to the case of inverting amplifier, here again input voltage V_i $(=V_1-V_S)$ is zero but since $V_1\neq 0$, noninverting OP-AMP has *no* virtual ground at either one of its input terminals. As the current into the input terminal, $I_1=0$, we conclude that current, I, through R_1 will also flow through R_2, we have from fig. 4 that

$$V_1=R_1\left(\frac{V_0}{R_1+R_2}\right)$$

or

$$V_S=R_1\left(\frac{V_0}{R_1+R_2}\right)$$

or

$$\frac{V_0}{V_S}=\frac{R_2+R_1}{R_1}=1+\frac{R_2}{R_1} \qquad \ldots(4)$$

Thus we find that output is dependent on the ratio (R_2/R_1). If $R_1=\infty$, then $V_0=V_S$ *i.e.* output voltage follows the input voltage, *i.e.*, OP-AMP circuit acts as a *voltage follower*. We have thus observed that in noninverting OP-AMP circuits :

(*i*) *No current flows into either input terminals.*

(*ii*) *The voltages at the two input terminals to the amplifier are equal* ($V_1=V_S$), and

(iii) Since $V_1 = V_S$, the amplifier is said to be operating with some common mode voltage at its input terminals because some voltage is common to both terminals. Thus there is effectively no voltage drop at the input terminals as V_1 follows V_S. It implies that effectively an open circuit results between input terminals.

24·2. DIFFERENTIAL AMPLIFIER :

In fig. (5 *a*) circuit of a differential amplifier is shown. This

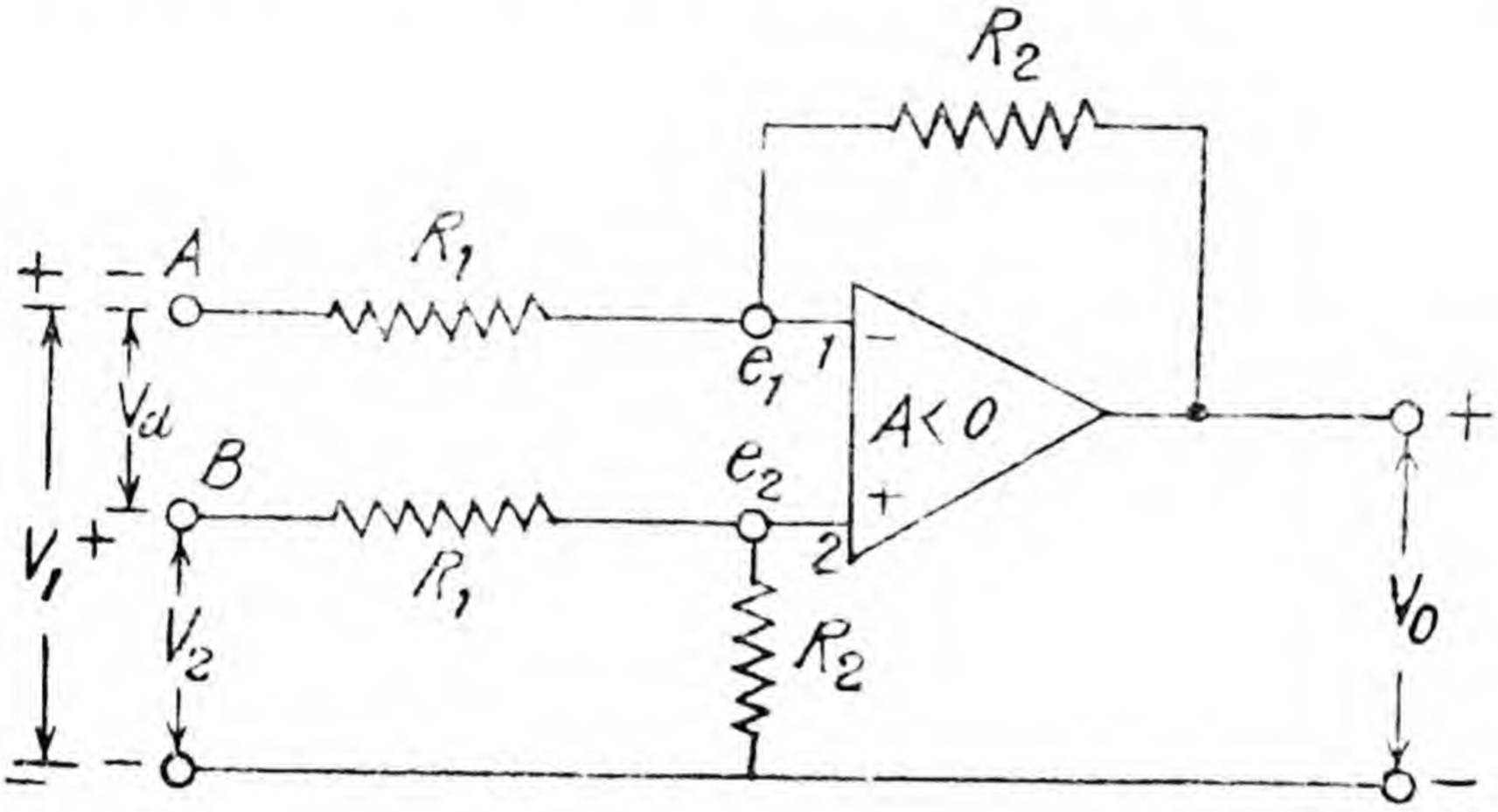


Fig. 5 (*a*)

amplifier provides the gain for differential input $(V_2 - V_1)$ and rejects the input voltage common to both. The voltage e_2, at the noninverting terminal, 2, remembering that input current to the ideal amplifier is zero, is given by

$$e_2 = \left(\frac{R_2}{R_1 + R_2}\right) V_2$$

Here $R_2/ R_1 + R_2)$ is termed as the transfer function $T(S)$ of the network involving R_1 and R_2 at the terminal 2. Similarly by the principle of superposition, the voltage at the inverting input terminal, 1,

$$e_1 = \left(\frac{R_2}{R_1 + R_2}\right) V_1 + \left(\frac{R_1}{R_1 + R_2}\right) V_0,$$

since the input current to the ideal amplifier is zero.

As the potential difference between two inputs terminals is forced to zero by the feedback through R_2, we haze $e_1 = e_2$; that is,

$$\left(\frac{R_2}{R_1 + R_2}\right) V_1 + \left(\frac{R_1}{R_1 + R_2}\right) V_0 = \left(\frac{R_2}{R_1 + R_2}\right) V_2$$

or

$$R_2 V + R_1 V_0 = R_2 V_2$$

$$V_0 = \frac{R_2}{R_1} (V_2 - V_1) \qquad \ldots (1)$$

$$= \frac{R_2}{R_1} V_d,$$

where V_d is the differential input. Thus output is directly proportional to the difference, V_d, between two input voltages V_1 and V_2. In this case we note that :

(i) there is no virtual ground at the input to the amplifier in this circuit, and

(ii) the feedback in the circuit forces e_1 to equal e_2, *i.e.*, amplifier operates in such a manner as to maintain near zero volts between the input terminals. Thus if $R_2 = 100K$ ohm, $R_1 = 10K$ ohm, $V_2 = +3{\cdot}1$ volts and $V_1 = +3{\cdot}0$ volts, then

$$V_0 = \frac{100 \times 10^3}{10 \times 10^3} \times (3{\cdot}1 - 3) = 1{\cdot}0 \text{ volt.}$$

There are three modes of operation of differential amplifier :

(1) *Single ended mode* : The operation in which either V_1 or V_2 is zero, is called single ended mode of operation. If $V_1 = 0$, the differential amplifier operates in noninverting mode and if $V_2 = 0$ then it operates in inverting mode.

(2) *Differential mode* : In this mode, the two input signals are equal but of opposite polarity at every instant of time.

(3) *Common mode* : In this mode, the input signals are identical both in amplitude and phase at every instant of time, *i.e.*, $V_1 = V_2$ so that from eq. (1), $V_d = 0$ giving $V_0 = 0$. Thus common mode input signals produce no output voltage.

Emitter Coupled Differential Amplifier : Fig. (5*b*) shows the basic form of differential amplifier. It consists of a pair of identical transistors Q_1 and Q_2 connected to a common emitter resistor (the tail). The current through this common resistor is called as the tail current, I_T. There are two inputs V_1 and V_2 and output, V_0 is the voltage between the two collectors.

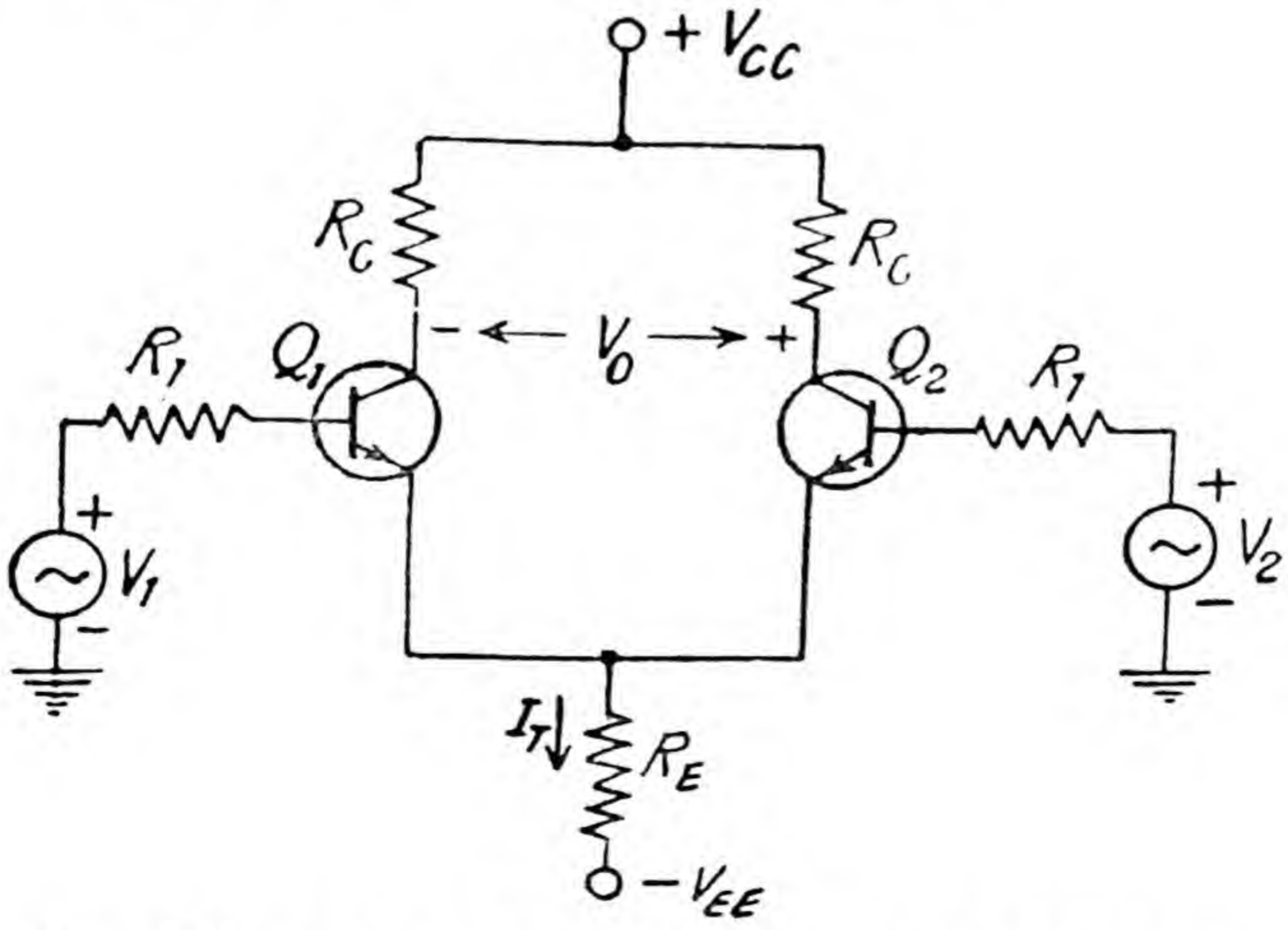


Fig. 5 (*b*). Basic form of a differential amplifier.

d.c. analysis of a differenttial amplifier : The d.c. equivalen of fig. 5 (*b*) is shown in fig. 5 (*c*). The differential amplifier uses emitter bias The top of the emitter resistor is an approximate ground point so that full supply voltage V_{EE} appears across R_E. Therefore d.c. tail current is

$$I_T = \frac{V_{EE}}{R_E}$$

Total current divides equally between two transistors as they are identical. That is d.c. emitter current in each transistor is half the total current,

$$I_E = \frac{I_{dc}}{2}$$

Further to a close approximation $I_C = I_E$. In fig. 5 (*c*) d c. voltage from the collector of Q_1 to ground is equal to

$$V_{C_1} = V_{CC} - I_C R_C$$

where $I_C R_C$ is the drop across collector resistor. Similarly

$$V_{C_2} = V_{CC} - I_C R_C$$

so that

$$V_0 = V_{C_1} - V_{C_2} = 0$$

That is d c. output voltage is zero.

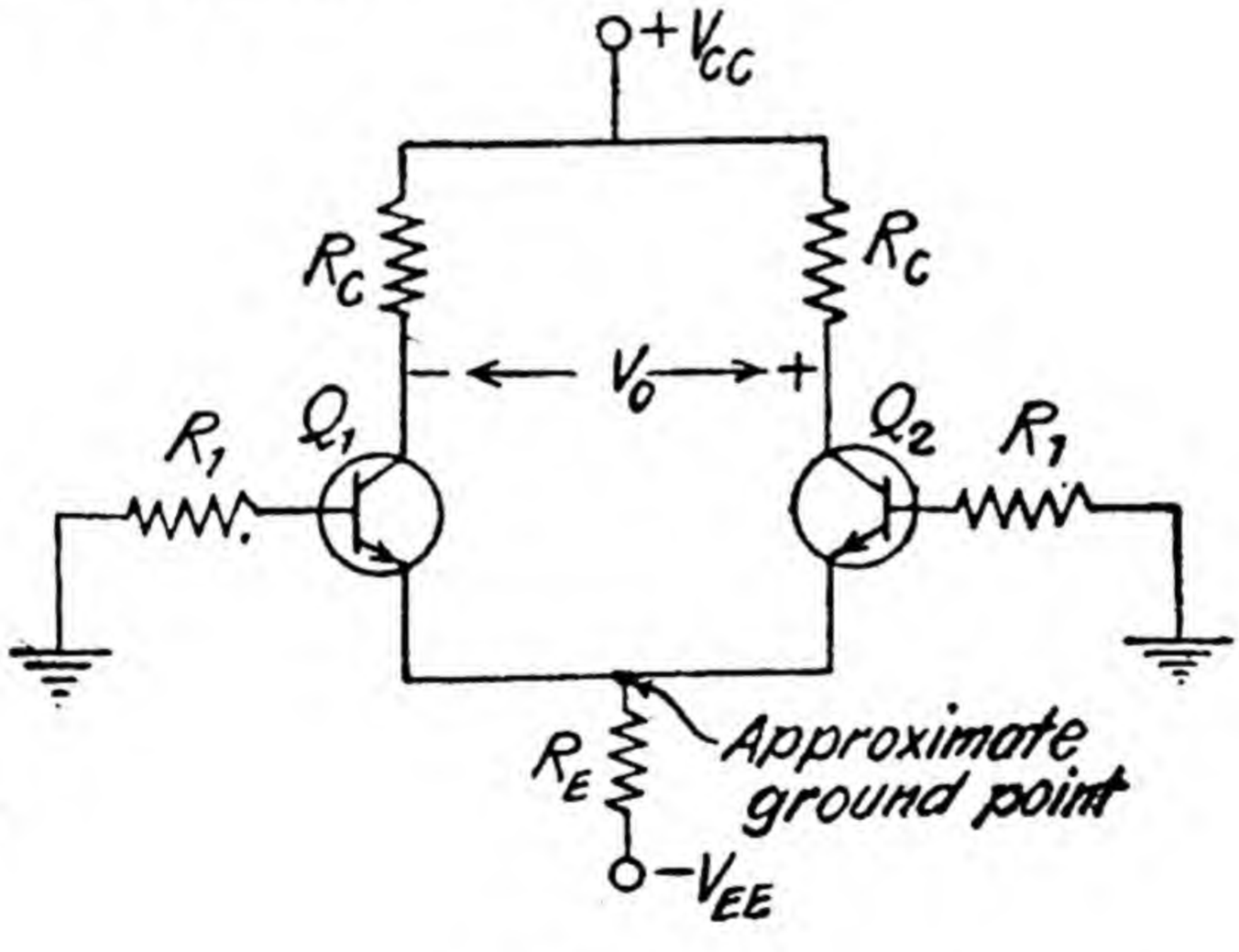


Fig. 5 (*c*).

a.c. analysis of differential amplifier : Refer to fig. 6 (*a*). When the two halves of differential amplifier are identical, output voltage is zero. If $V_1 > V_2$ more collector current will flow through the transistor Q_1 so that polarity of V_0 will be negative at Q_1. Fig. 6 (*b*) is the equivalent circuit for differential amplifier.

To get a.c. equivalent circuit, visualize all d c. sources to fig. 6 (*b*) reduced to zero. This is equivalent to a.c. grounding the V_{CC} supply point and opening the I_T current source. Since two a.c. sources drive the differential amplifier, we have to use superposition theorem to obtain output voltage. That is, first we shall consider that V_1 is active and V_2 is off, and then V_1 is off and V_2

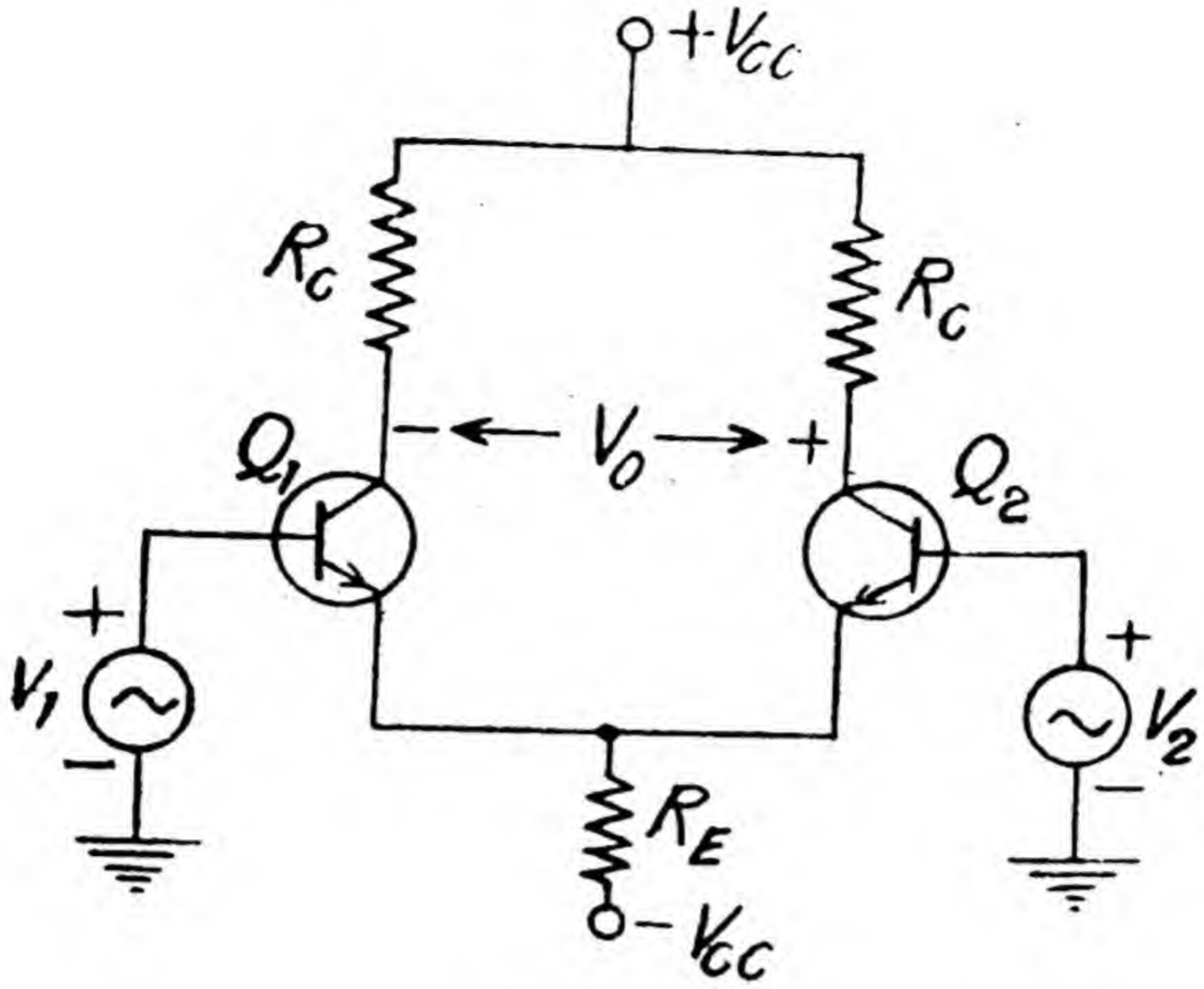


Fig. 6 (*a*)

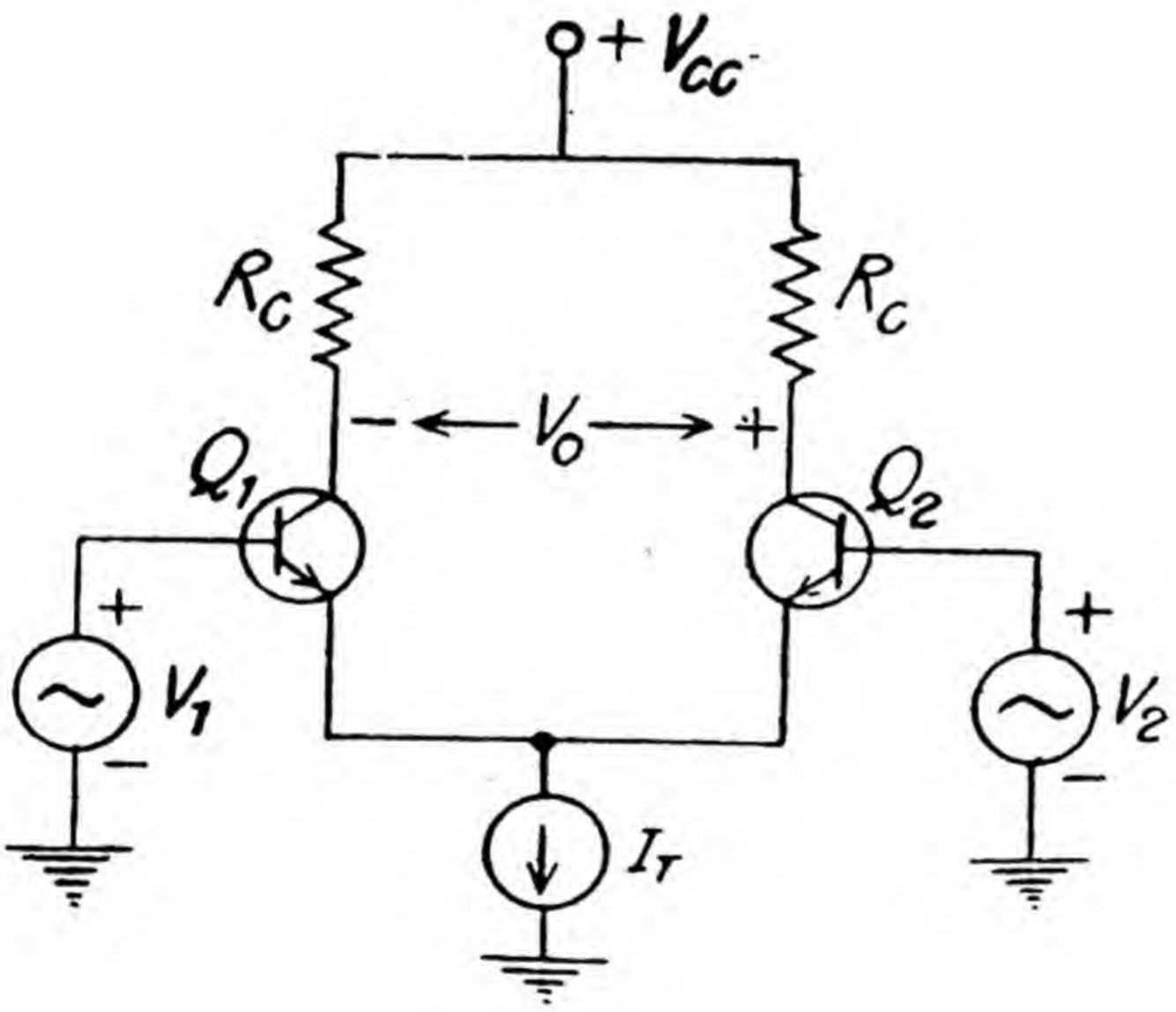


Fig. 6. (*b*)

is active. Refer to fig. 7 (*a*). If V_1 is active and V_2 is off then Q_1 acts like a common emitter (CE) amplifier so that an *inverted* signal appears at its collector. This a.c. voltage is between collector of Q_1 and ground. Since I_T current source is open to a.c., the emitter current of Q_1 will flow into the emitter of Q_2. Q_2 therefore acts like common base (CB) amplifier and an *inphase* signal appears at its collector. Thus output voltage taken between collectors of Q_1 and Q_2 will be the algebraic sum of two equal and opposite sine waves, the output has twice the amplitude. The contribution of first source acting alone is

$$V_0(1) = AV_1$$

Now refer to fig. 7 (*b*) V_2 is made active while V_1 is off. Now Q_2 acts like a *CE* amplifier and Q_1 like a *CB* amplifier so that inverted signal appears at the collector of Q_2 while inphase signal

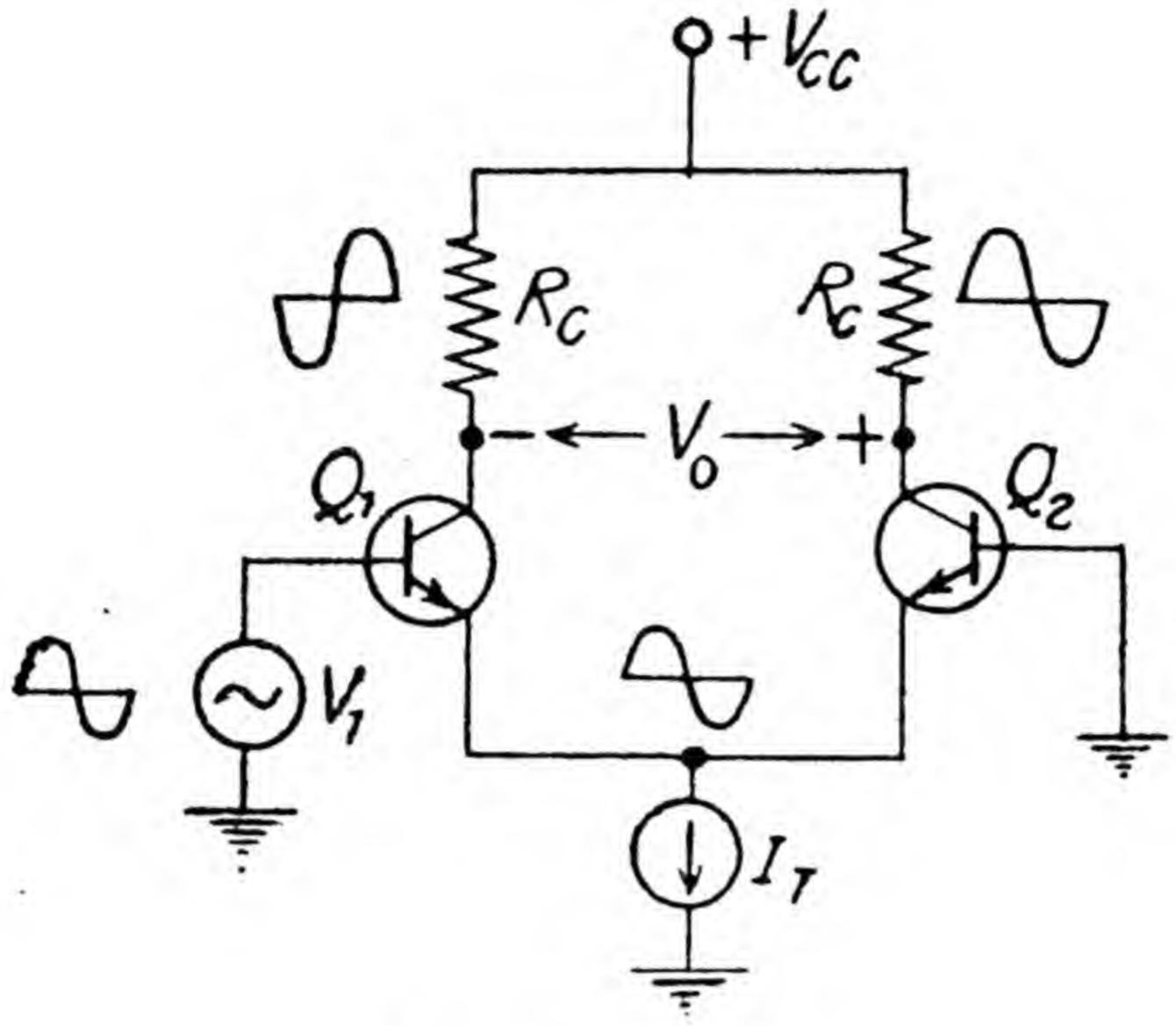


Fig. 7 (*a*)

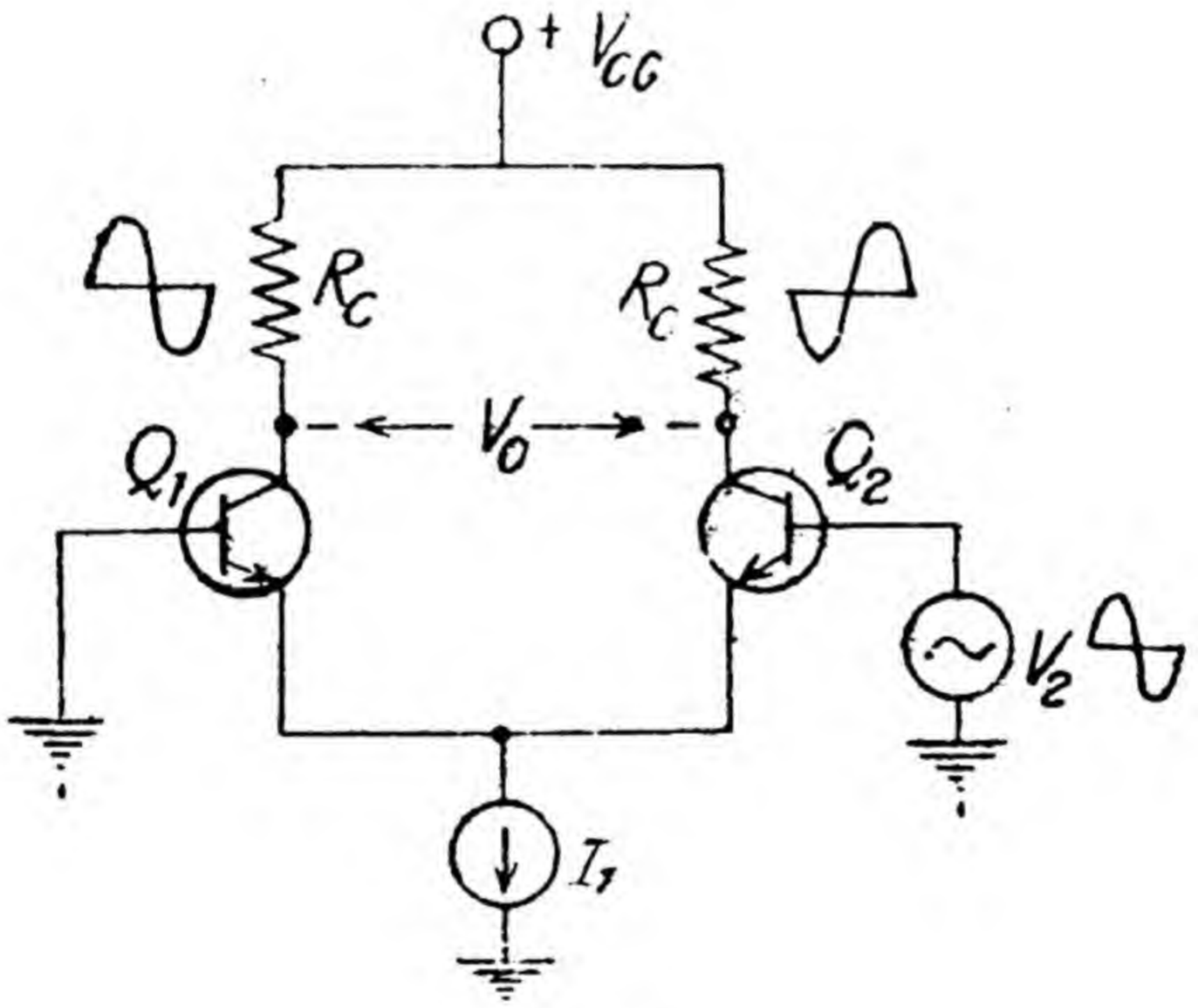


Fig. 7 (*b*)

appears at the collector of Q_1. The contribution of second source acting alone is

$$V_0\,(2) = -AV_2$$

Applying superposition theorem, when both sources act simultaneously

$$V_0 = V_0\,(1) + V_0\,(2)$$
$$= AV_1 - AV_2 \text{ (two inputs)}$$

Let us draw the circuits showing different modes of operation.

(1) *Differential input* : We can drive the differential amplifier with a signal between the bases as shown in fig. 8. Here $V_i = V_1 - V_2$ so that $V_0 = AV_i$.

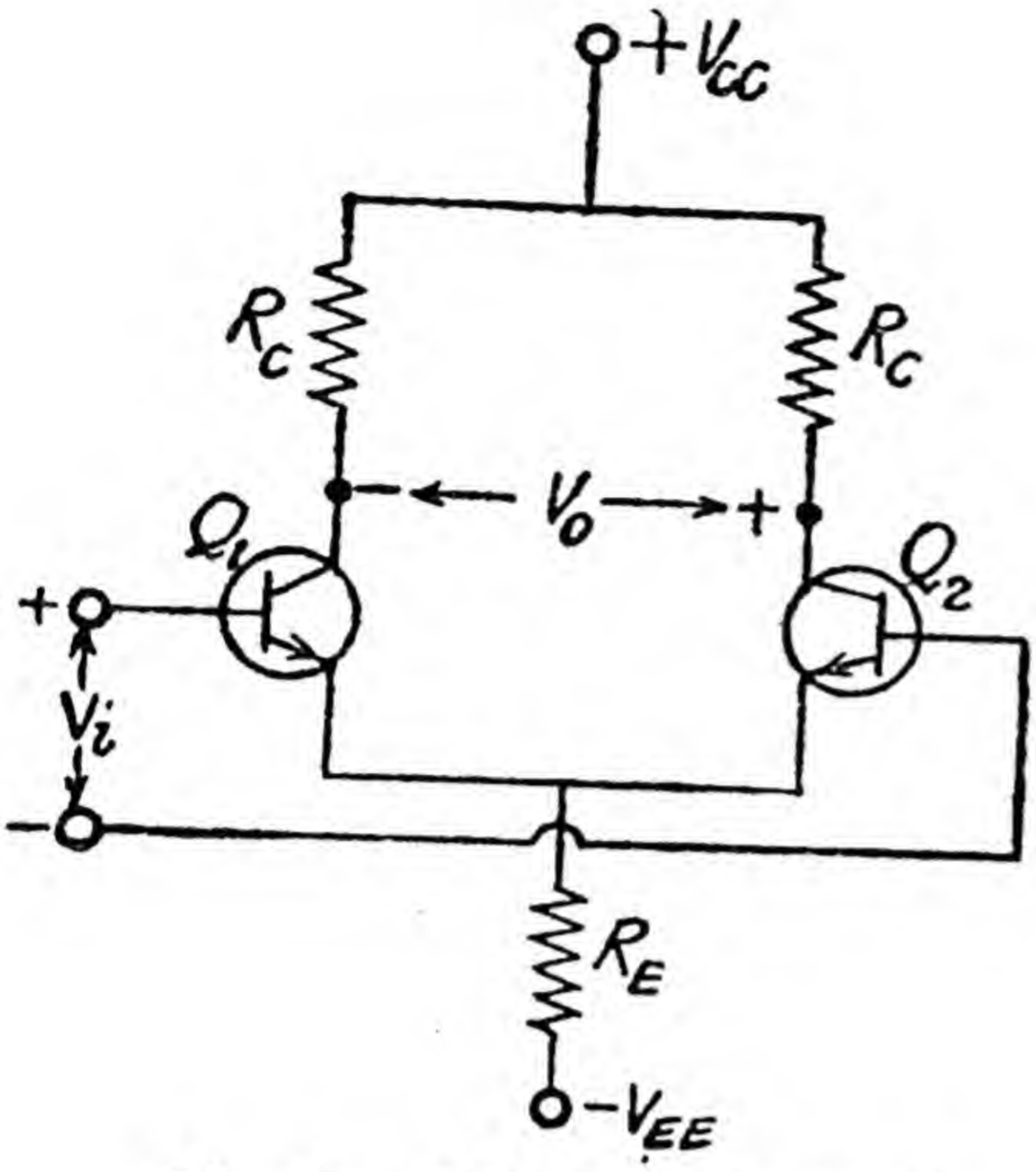


Fig. 8. Differential input.

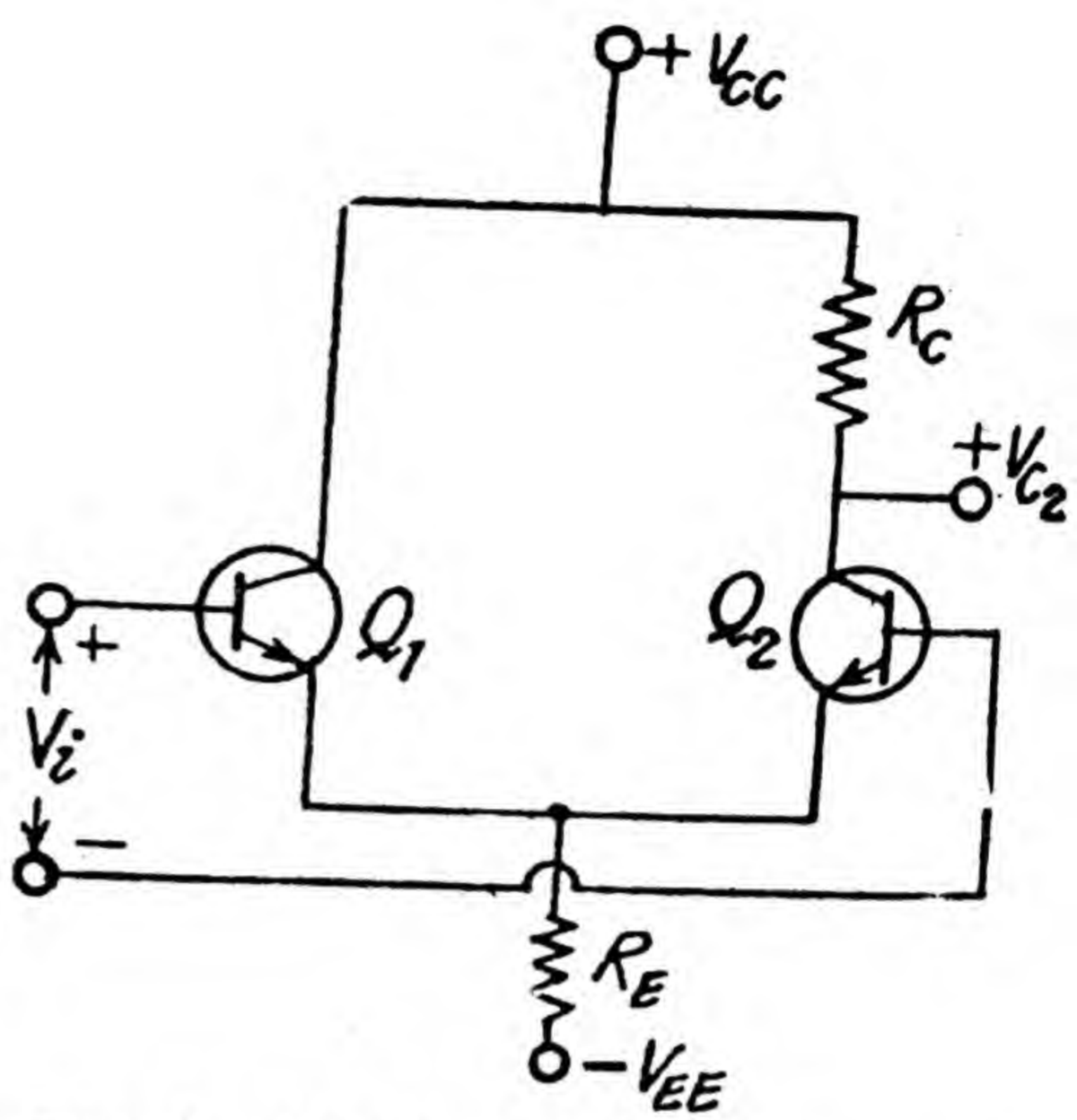


Fig. 9. Single ended input.

(2) *Single ended input* : Circuit is shown in fig. 9. Only half the available output voltage is used, the voltage gain drops in half

$$V_{C_2}=\frac{A}{2}\ V_i.$$

(3) *Common mode input* : The same signal is applied to both inputs as shown in fig. 10. If each half of the differential amplifier is identical, a.c. output voltage will be zero. Common mode input signals are used to test how identical the two halves of the amplifier are.

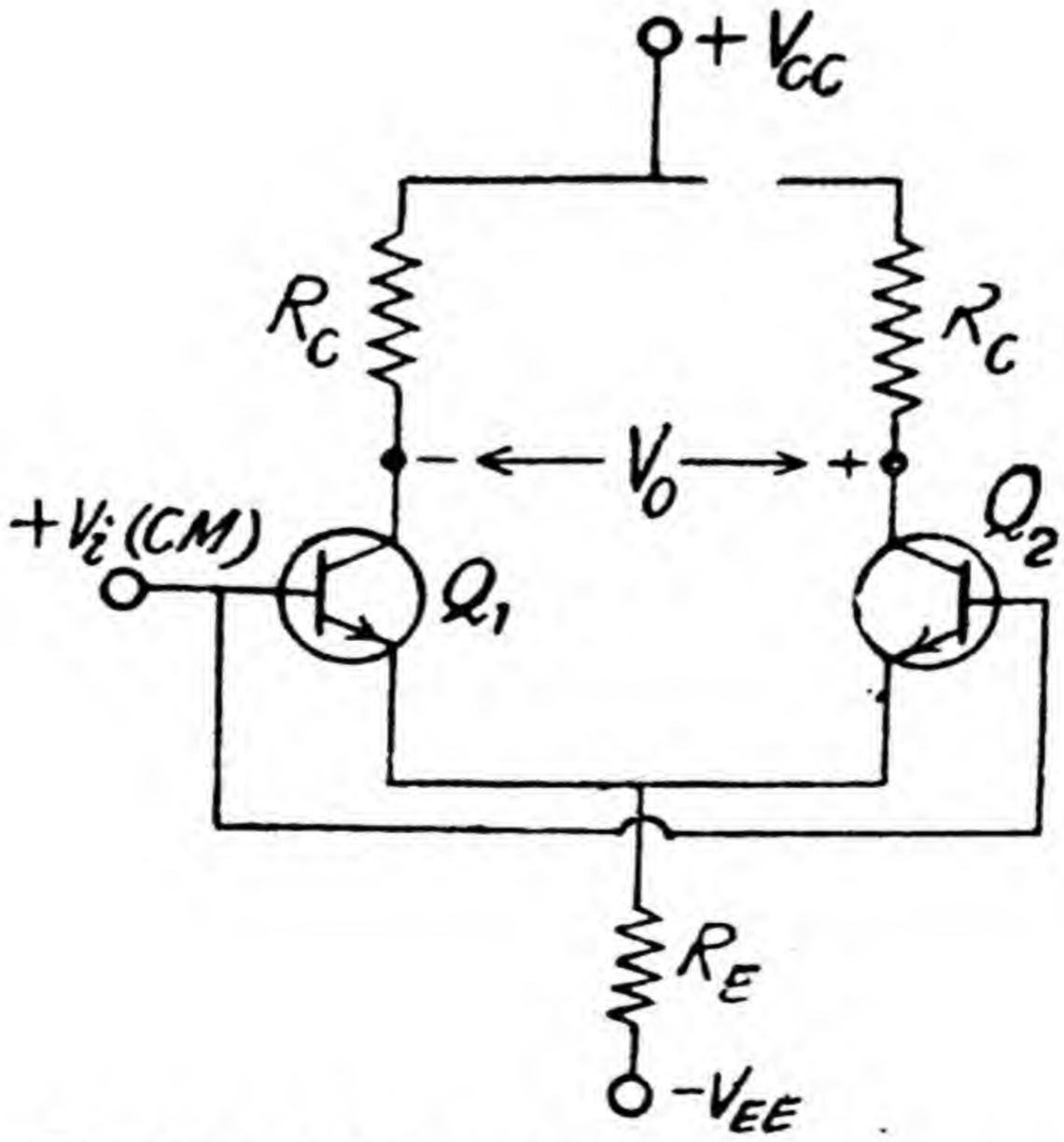


Fig. 10. Common mode input.

Common mode rejection ratio : To express how successful a differential amplifier is in providing gain for the differential input (the difference between two input voltages) and rejecting the com-

mon mode signal (the voltage common to both), a factor called common mode rejection ratio is defined as

$$CMRR=\frac{A_d}{A_c}=\frac{\text{Differential gain}}{\text{Common mode gain}}$$

Alternatively the common mode rejection may be expressed in decibels as

$$CMR=20 \log CMRR$$
$$=20 \log (A_d/A_c)=20 \log A_d-20 \log A_c \quad \text{.. (1)}$$

To explain more about $CMRR$, let us show a linear active device with two inputs V_1 and V_2 and one output V_0, each measured with respect to ground (fig. 11). In ideal differential amplifier,

$$V_0=A_d (V_1-V_2)$$
$$=A_d V_d$$

where A_d is the differential gain of the amplifier (*i.e.* gain for



Fig. 11.

differential input V_d). Ideally a signal common to both inputs produces zero output but in practical differential amplifier, output depends not only on difference signal, V_d but also on the average signal (called common mode signal V_c), where

$$V_c=\left(\frac{V_1+V_2}{2}\right) \text{ and } V_d=V_1-V_2$$

giving

$$\left.\begin{aligned} V_1&=V_c+\tfrac{1}{2}V_d \\ \text{and} \quad V_2&=V_c-\tfrac{1}{2}V_d \end{aligned}\right\} \quad \text{...(2)}$$

Suppose A_1 is the voltage gain for input V_1 with V_2 grounded, and A_2 is the voltage gain for input V_2 with V_1 grounded then we can express output as a linear combination of two input voltages, *i.e.*,

$$V_0=A_1V_1+A_2V_2$$
$$=A_1 (V_c+\tfrac{1}{2}V_d)+A_2 (V_c-\tfrac{1}{2}V_d)$$
$$=\tfrac{1}{2} (A_1-A_2)\, V_d+(A_1+A_2)\, V_c$$
$$=A_dV_d+A_cV_c$$

where $A_d=\frac{1}{2} (A_1-A_2)$ and $A_c=(A_1+A_2)$.

If we put $V_1=0{\cdot}5$ volt, $V_2=-0{\cdot}5$ volt, then

$$V_d=V_1-V_2=1{\cdot}0 \text{ volt and } V_c=\frac{V_1+V_2}{2}=0$$

so that $V_0=A_dV_d+A_cV_c=A_d$

That is, output voltage will directly give A_d. Similarly if we put

$V_1=1$ volt, $V_2=1$ volt then

$$V_d=V_1-V_2=0 \quad \text{and} \quad V_c=\frac{V_1+V_2}{2}=1\cdot0 \text{ volt.}$$

so that $$V_0=A_dV_d+A_cV_c=A_c$$

That is, in this case output voltage will provide directly the value of A_c. Thus we can determine both A_d and A_c and find $CMRR$. For better performance of the differential amplifier A_d should be large and A_c should be zero, *i.e.*, $CMRR$ should approach infinity. In other words larger the value of $CMRR$, better is the differential amplifier.

Let us find total output voltage and $CMRR$ in fig. 5, if voltage $e_1=1\cdot3\ mV$, $e_2=1\cdot5\ mV$, $A_d=1200$ and $A_c=2\cdot14$. It is given by

$$V_0=A_dV_d+A_cV_c$$

$$=A_d\ (e_2-e_1)+A_c\left(\frac{e_2+e_1}{2}\right)$$

$$=1200\times(1\cdot5-1\cdot3)\times10^{-3}+2\cdot14\left(\frac{1\cdot3+1\cdot5}{2}\right)\times10^{-3}$$

$$=0\cdot24+0\cdot003=0\cdot243 \text{ volts.}$$

and $$CMRR=\frac{A_d}{A_c}=\frac{12\cdot00}{2\cdot14}=560.$$

Also if in fig. 10, $V_i\ (CM)=1\cdot0$ volt, $A=100$, $V_0(CM)=0\cdot01$ volt, then for such an amplifier

$$CMRR=\frac{A_d}{A_c}=\frac{AV_i\ (CM)}{V_0\ (CM)}=\frac{100\times1}{0\cdot01}=10^4$$

so that $$CMR=20 \log 10^4=80\ db.$$

24·3. A GENERAL PURPOSE IC OPERATIONAL AMPLIFIER :

A popular monolithic circuit of differential operational amplifier is shown in fig. 12. It consists of

(i) Darlington—connected NPN transistors (Q_1 and Q_2) for high input impedance; such stage can provide input impedance of 10 to 20 $M\Omega$. JFET device offers another approach to obtain high input impedance, and low bias current for input stage.

(ii) Two differential gain stages with a single ended emitter follower output (Q_3, Q_4).

Effect of grounding A and giving B a positive voltage : Then point Y will become less positive so that Q_4 will conduct more and in effect will make the point Z *less negative*. Consequently Q_5 will conduct less so that output will *rise* to some positive value. This will provide some gain A_1.

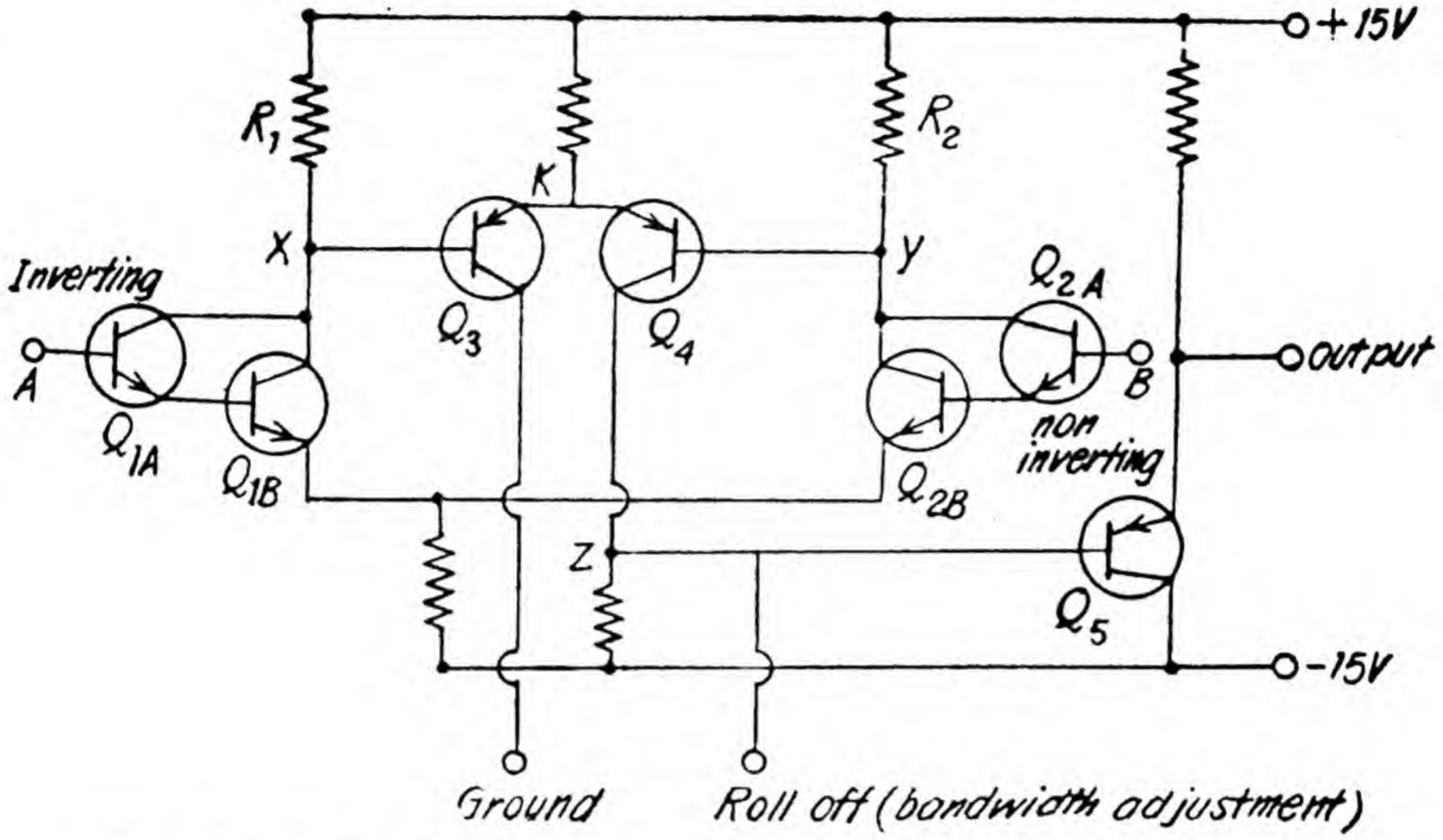


Fig. 12. Circuit of a mono.ithic EC differential cperational amplifier.

Effect of grounding B and giving A a positive voltage : Then point X will become less positive so that Q_3 will conduct more and in effect will make the point K less positive. Q_4 will then conduct less making Z *more negative*. Therefore Q_5 will conduct more, quite opposite to the previous case when A was grounded. Thus the change in voltage at the collector of Q_4 is transferred to the output as a similar negative voltage change which provides some gain A_2. Thus effect of grounding B is opposite to the effect of grounding A and if the two voltages are applied simultaneously at A and B and if $A_1 = A_2$ then, obviously, the output will be zero. But due to slight difference in components, A_1 will not always equal A_2 and so an output will result even though the two inputs are identical. The common mode input will be

$$V_c = \frac{V_1 + V_2}{2} \qquad \text{(if the two inputs } V_1 \text{ and } V_2 \text{ are not equal)}$$

$$= V \qquad \text{(if inputs are equal)}$$

The ratio of the output to the common mode input is called the common mode gain as deroted by A_c earlier. However, what desired is the differential gain, A_d, of the ampiifier (called open loop gain) and is the ratio of the output to the differential input $(V_1 \sim V_2)$.

24·4. SOME OPERATIONAL AMPLIFIER PARAMETERS :

These days operational amplifier has become a universal building block for circuit and system design. Thus various OP-AMP circuits are in use. To compare their merits, as to the performance and design, a number of parameters have been defined. Some of them are briefed as follows :

(1) Input Bias Current : In an ideal OP-AMP output is zero when the two inputs are identical ($V_0=0$ if $V_1=V_2$). In practice, due to mismatch of input transistors unequal bias current flows through the input terminals. Thus input bias current is the current flowing into each of the two input terminals when they are biased to the same voltage levels or the average of the currents into the two input terminals (fig. 13) with output at zero volts, is input bias current. That is

$$I_{bias}=\frac{I_{D_1}+I_{D_2}}{2} \qquad ...(1)$$

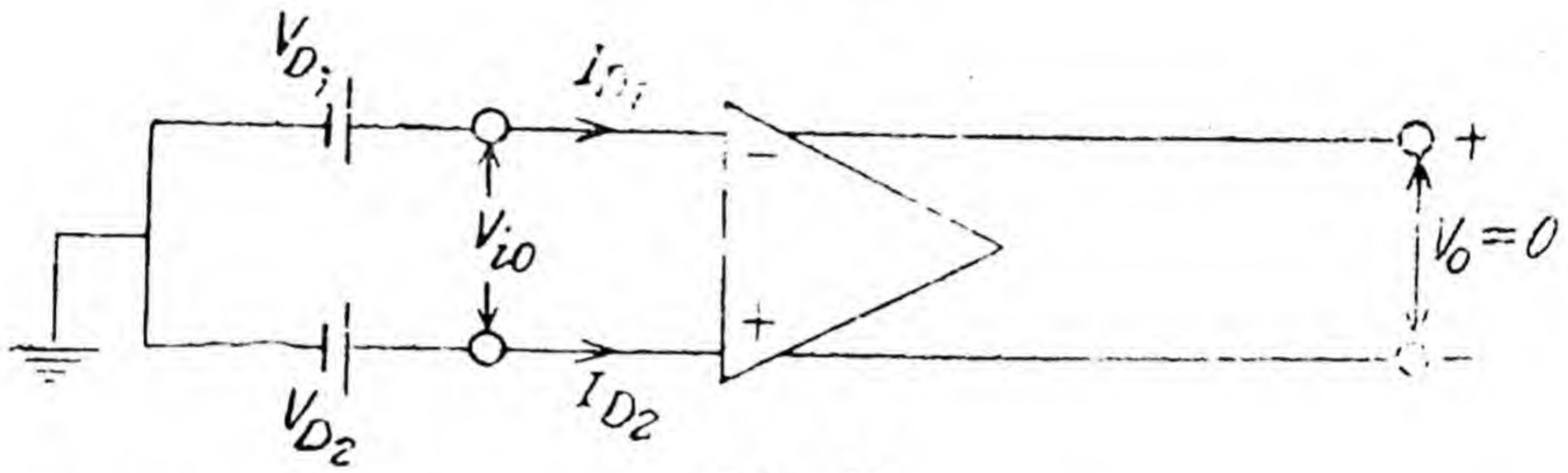


Fig. 13.

(2) Input Offset Current : The input offset current, I_{i0}, is the difference of the currents into the two input terminals with the output at zero volts ($V_0=0$). Thus with $V_0=0$, we have

$$I_{i0}=I_{D_1}-I_{D_2} \qquad ...(2)$$

Input offset current typically lies in the range 20 to 60 *mA.* This current drifts with change in temperature. If $\triangle I_{i0}$ is the change in input offset current due to the change $\triangle T$ in the temperature, then

$$\text{Input offset current drift}=\frac{\triangle I_{i0}}{\triangle T} \qquad ...(3)$$

(3) Input offset voltage : It is the input voltage which must be applied across the input terminals to obtain zero output voltage.

In practical OP-AMP, it is found that even if equal voltage are applied to the input terminals, output voltage is not zero (due to inherent imbalance in the circuits). To set output voltage to zero, we require an input offset voltage, V_{i0}, which typically lies in the range 1 to 4 *mV.*

Input offset voltage drifts with change in temperature. If $\triangle V_{i0}$ is the drift in input offset voltage for a change $\triangle T$ in temperature, then

$$\text{input offset voltage drift}=\frac{\triangle V_{i0}}{\triangle T} \qquad ...(4)$$

(4) Input Common Mode Range : It is the maximum differential signal that can be applied safely to OP-AMP inputs.

(5) Power Supply Voltage Rejection Ratio : It is defined as input offset voltage change per volt of supply voltage change. *i.e.*

$$PSRR = \frac{\Delta V_0}{V_{cc}} \qquad ...(5)$$

(6) Output Voltage Swing (range) : It is the maximum peak to peak output voltage which can be obtained without waveform distortion. Output voltage swing is a function of the supply voltage and may range 50 to 80 percent of the supply voltage.

(7) **Full Power Bandwith :** It is the maximum frequency over which the full output voltage swing can be obtained.

(8) **Slew Rate :** It is the maximum rate of change of output voltage for a *step* input.

The response of an OP-AMP to *large change in input signal is not as fast as might be expected.* Suppose that a step function of large amplitude is applied as V_1 (fig. 14). The internal capaci-

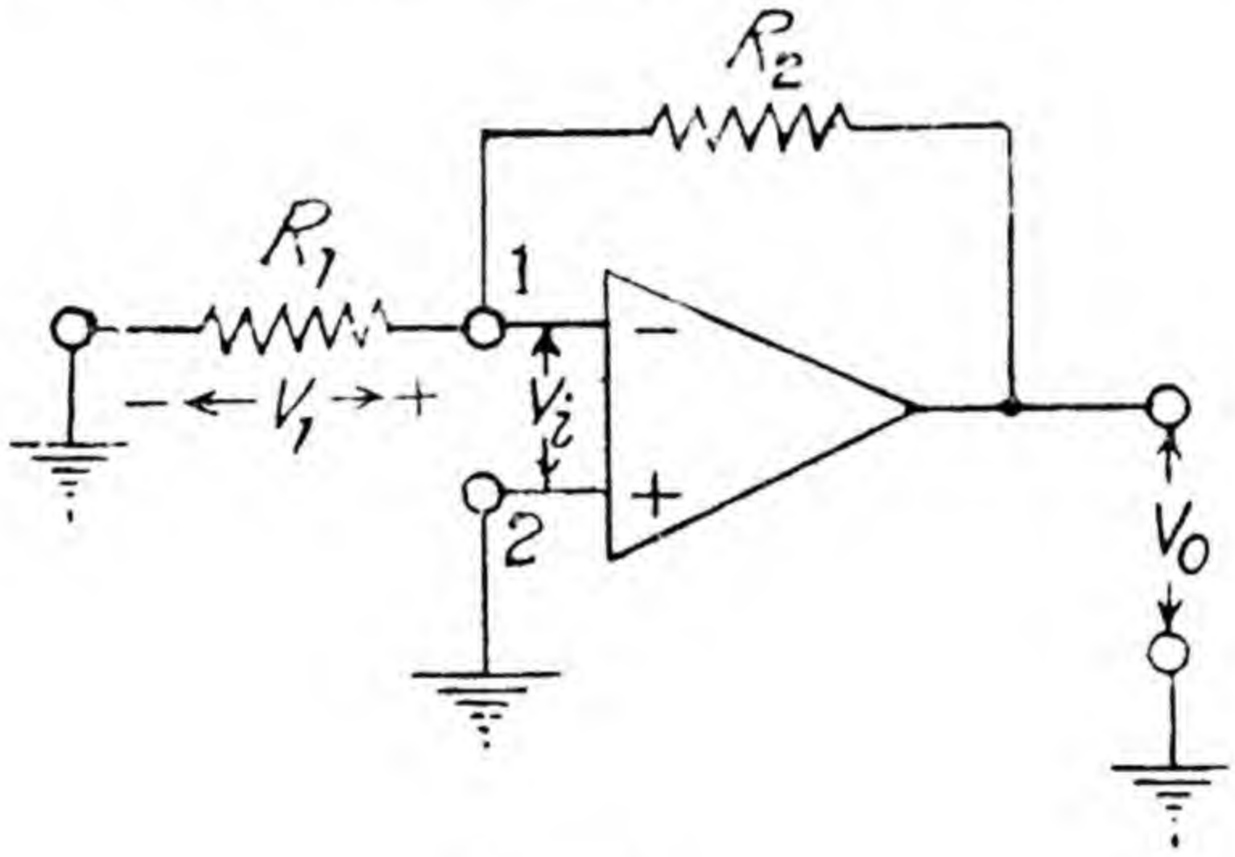


Fig. 14.

tances of the amplifier and feedback loop can not change voltage rapidly so that a *finite time* is required for V_0 to respond to the step input. Further as the feedback voltage, applied through R_2 to input, is proportional to V_0, the feedback is delayed in its return to the input, and the amplifier input voltage V_i, is momentarily large. As a result, until the feedback voltage responds, input voltage V_i drives the amplifier into saturation and a distorted output waveform appears as a transient.

Output voltage, V_0, can rise only as fast as the capacitance can change or slew in voltage. The maximum possible rate of change of output voltage is defined as the slew rate, S, where

$$S = \left(\frac{dV_0}{dt}\right)_{max}$$

For a sinusoidal voltage

$$V_0 = |V_0| \sin \omega t$$

so that

$$\frac{dV_0}{dt} = \omega |V_0| \cos \omega t$$

or $$\left(\frac{dV_0}{dt}\right)_{max} \geqslant \omega \mid V_0 \mid$$

so that maximum possible amplitude at a frequency, f, is

$$\mid V_0 \mid = \frac{(dV_0/dt)_{max}}{\omega} \leq \frac{S}{\omega} \qquad \ldots(6)$$

This slew rate is determined by the capacitances in the amplifier. Further

(i) as the input is a very fast step of amplitude which is much larger than the amplifier can handle linearly, in slew rate determination OP-AMP operates non-linearly. Thus slew rate permits us to determine the extent to which the linear range of operation of an OP-AMP is affected by the speed of the input signal.

(ii) as the large step input drives the output *abruptly* from one limit to other, the feedback cannot be effective (slew rate may range 50 $mV/\mu s$ to 50 $V/\mu s$) and therefore slew rate is rather independent of the amount of feedback.

Problem. An OP-AMP has a slew rate of 4 $V/\mu s$ and peak output swing of 10 volts. ind out the full power bandwidth.

Full power bandwidth

$$\omega_p = \frac{S}{\mid V_0 \mid} \quad \text{(See eq. 6)}$$

or $$f_p = \frac{S}{2\pi \mid V_0 \mid} = \frac{4 \times 10^6 \ V/s}{2 \times 3{\cdot}14 \times 10} = 83{\cdot}7 \ kc/s.$$

24·5. EFFECTS OF OFFSET :

In art 24·4, we have defined two offset terms *viz*, offset voltage and offset current. An ideal OP-AMP has zero offset voltage and zero offset current but a practical OP-AMP has both. The direct current amplifiers and integrators are affected by the input offset and offset changes with time, temperature, and supply voltage.

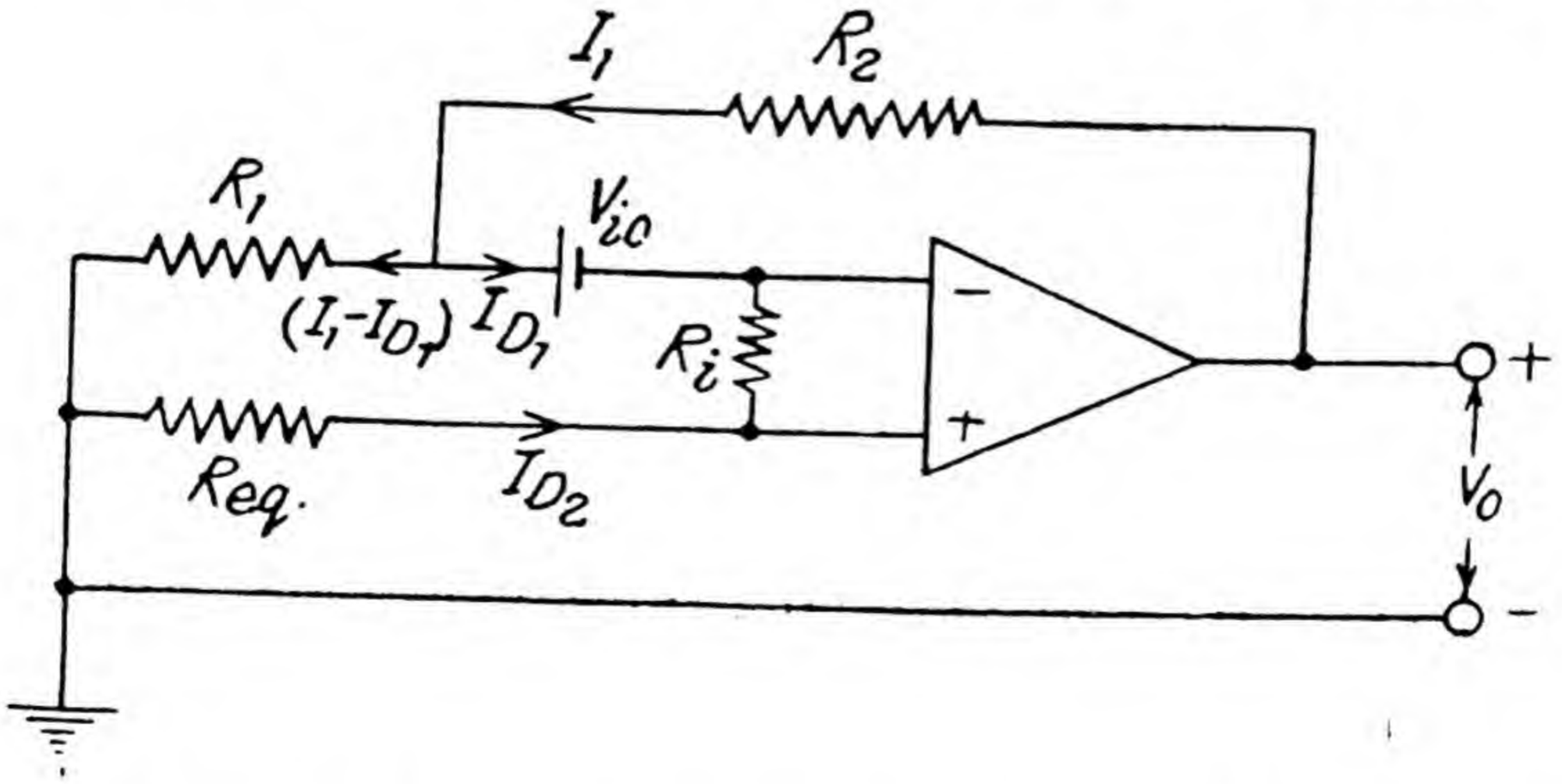


Fig. 15. The equivalent circuit of an OP-AMP with offset.

The equivalent circuit of an OP-AMP with offset is shown in fig. 15. The offset voltage is represented by a battery, V_{i0}, in the input (note that offset is referred to the input of the amplifier and because it is independent of the gain of the amplifier, it can be compared with input signal) and has the same magnitude as the input voltage which must be applied to obtain zero output voltage.

We select few loops and write

$$V_D = I_1 R_2 + (I_1 - I_{D_1})\, R_1 = I_1(R_1 + R_2) - I_{D_1} R_1 \quad \ldots (1)$$

and $$V_0 = [V_{i0} + (I_{D_1} - I_{D_2}) R_1 - I_{D_2} R_{eq}] + I_1 R_2 \quad \ldots (2)$$

But $$V_i = (I_{D_1} - I_{D_2})\, R_i = \frac{-V_0}{A_{(ideal)}} = -\frac{V_0}{\infty} \to 0$$

so that eq. (2) becomes

$$V_0 = V_{i0} - I_{D_2}\, R_{eq.} + I_1 R_2$$

or $$I_1 = \frac{V_0 - V_{i0} + I_{D_2} R_{eq.}}{R_2}$$

Putting I_1 in eq. (1), we get

$$V_0 = \left(\frac{R_1 + R_2}{R_2}\right)[V_0 - V_{i0} + I_{D_2} R_{eq.}] - I_{D_1} R_1 \quad \ldots (3)$$

or $$V_0 = \left[1 - \left(\frac{R_1}{R_2} + 1\right)\right] = (I_{D_2} R_{eq} - V_{i0})\left(1 + \frac{R_1}{R_2}\right) - I_{D_1} R_1$$

or $$V_0 = -(I_{D_2} R_{eq.} - V_{i0})\left(\frac{R_2}{R_1} + 1\right) + I_{D_1} R_2$$

$$= V_{i0}\left(\frac{R_2}{R_1} + 1\right) + I_{D_1} R_2 - I_{D_2} R_{eq.}\left(\frac{R_2}{R_1} + 1\right)$$

$$= V_{i0}\left(\frac{R_2}{R_1} + 1\right) + R_2\left[I_{D_1} - I_{D_2} R_{eq.}\left(\frac{1}{R_1} + \frac{1}{R_2}\right)\right]$$

We can minimise the effect of input bias current $(I_{D1} - I_{D_2})$ by putting,

$$R_{eq.} = 1\Big/\left(\frac{1}{R_1} + \frac{1}{R_2}\right) = \frac{R_1 R_2}{R_1 + R_2}, \text{ we get}$$

$$V_0 = V_{i0}\left(\frac{R_2}{R_1} + 1\right) + (I_{D_1} - I_{D_2})\, R_2$$

$$= V_{i0}\left(\frac{R_1}{R_2} + 1\right) + I_{i0}\, R_2, \quad (4)$$

(see eq. 2 art 24·4)

Where I_{i0} is input offset current. Eq. (4) gives the contribution of offset voltage and offset current to the output voltage, V_0, of thea mplifier,. To counteract the effects of offset voltage and offset current, nulling circuits which effectively introduce an input d.c. voltage can be applied. Refer to art 24 1, practical noninverting OP-AMP acts as voltage follower if $R_1 = \infty$ and we remove the

resistance R_2 then $V_0 = V_{10}$, that is we get minimum output offset voltage.

Problem. An inverting amplifier has $R_2 = 100\ K\Omega$, $R_1 = 10 K\Omega$. Find out the value of resistor which has to be connected between the positive input terminal and ground in order to minimise the effect of input bias current.

In fig. 15, this resistor is R_{eq}. To minimise the effect of input bias current, its value is

$$R_{eq.} = \frac{R_1 R_2}{R_1 + R_2} = \frac{100 \times 10^3 \times 10 \times 10^3}{100 \times 10^3 + 10 \times 10^3} = 9{\cdot}1 \text{ K ohm.}$$

24·6. FREQUENCY RESPONSE AND STABILITY :

An OP-AMP is required to be unconditionally stable so that it will not break into oscillations for any amount of negative feedback. The excess phase shift associated with the gain stage leads to oscillations. Hence it is necessary to compensate frequency

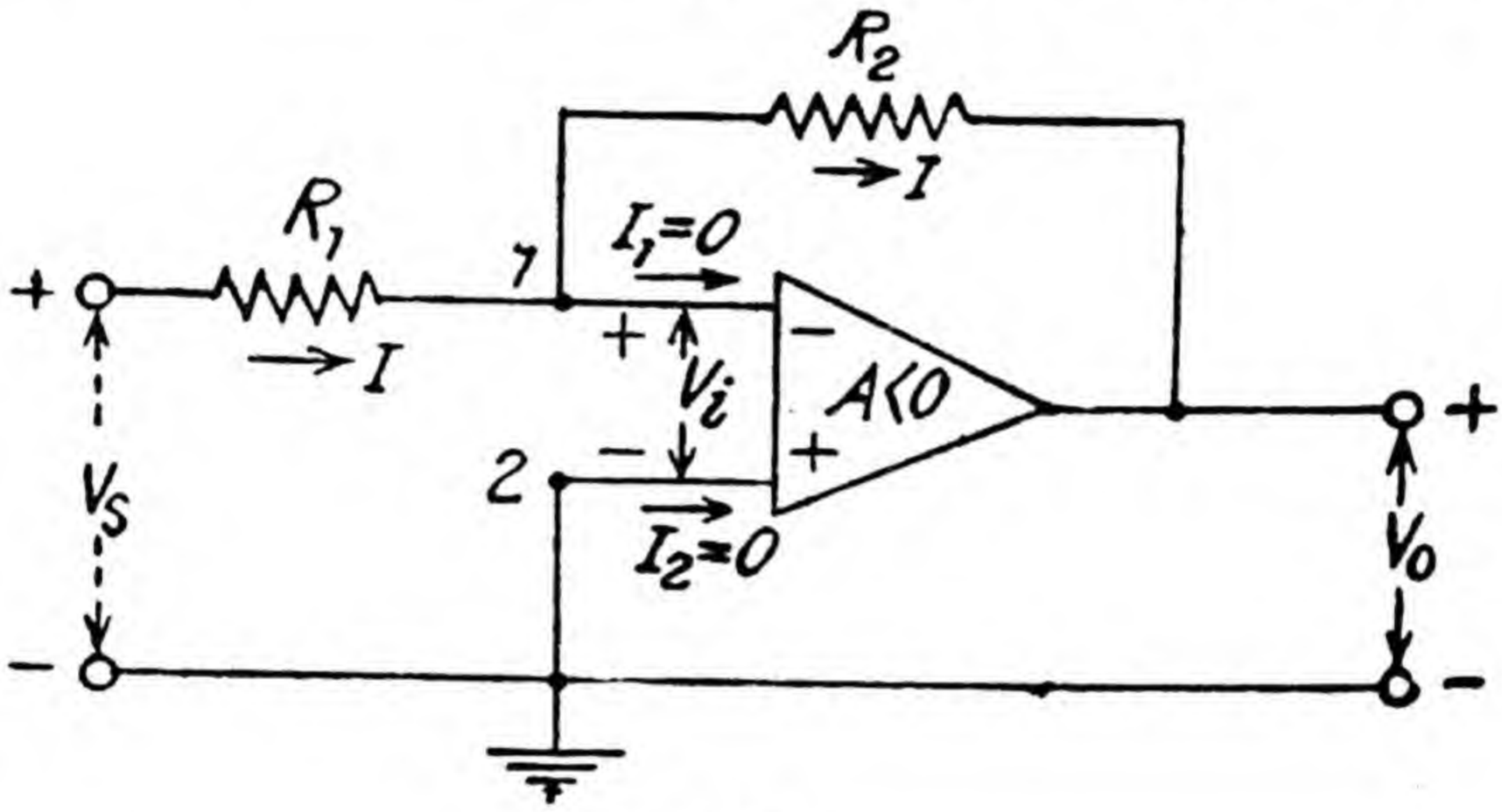


Fig. 16.

performance of the OP-AMP circuit in closed loop response. Let us sketch a closed loop inverting amplifier (fig. 16). Since $R_i \to \infty$, no current flows into the input terminals of the amplifier and the current I which flows through R_1 also flows through R_2 so that

$$I = \frac{V_S - V_i}{R_1} = \frac{V_i - V_0}{R_2}$$

or

$$\frac{V_S}{R_1} = \frac{V_i}{R_1} + \frac{V_i - V_0}{R_2}$$

$$= V_0 \left[\frac{V_i}{V_0 R_1} + \frac{(V_i/V_0) - 1}{R_2} \right]$$

Defining

$+A(\omega) = +\frac{V_0}{V_i} = -A(\omega)$, we have

$$\frac{V_S}{R_1} = -\frac{V_0}{A(\omega)} \left[\frac{1}{R_1} + \frac{1 + A(\omega)}{R_2} \right]$$

or

$$\frac{V_0}{V_S} = \left(\frac{R_2}{R_1 + R_2} \right) \left(\frac{-A(\omega)}{1 + \frac{R_1 A(\omega)}{R_1 + R_2}} \right)$$

The denominator term

$$\left(1+\frac{R_1 A(\omega)}{R_1+R_2}\right)$$

is the term of interst regarding stability and is defined as the loop gain. The gain $A(\omega)$ is complex, having magnitude as well as phase.

If $$\frac{R_1 \mid A(\omega) \mid}{(R_1+R_2)}=1\underline{/180°}$$

that is, it is unity (or greater) in magnitude and becomes negative (or the phase angle of $A(\omega)$ becomes 180°) then $V_0/V_S \to \infty$. It means we shall have an output signal with no input a condition for oscillations and the circuit no more acts as amplifier. This oscillating condition is to be avoided. By sketching the Bode plot of the amplitude and phase of the loop gain function $A(\omega)|/(R_1+R_2)$ *versus* frequency, it can be determined whether the magnitude is greater than zero decibels when the phase angle passes through 180° so that amplifier is unstable. Readers will find a text on operational amplifier useful for study on frequency compensation.

24·7. APPLICATIONS OF OPERATIONAL AMPLIFIER :

Some of the applications of operational amplifier, including linear analog circuits, are given below :

(1) Inverting Amplifier acting as a Scale changing, Phase shifting; and Summing amplifier : Refer to fig. 2(*a*) art 24·1 of basic inverting OP-AMP. Its noninverting terminal is grounded, and virtual short circuit exists at input terminals. From eq. (2), that we have for such amplifier (art. 24·1), we find that if $Z_1=Z_2$ then output voltage V_0 is $(-V_S)$, that is, simply the sign of the input has been changed. Hence such a circuit acts as a phase inverting amplifier, and is called a *sign changer* or *inverter*. If two such amplifiers are connected in cascade, the output of the second stage will be exactly of same magnitude and sign as the input voltage of the first stage, though the output of the two stages will be equal in magnitude but opposite in sign. The system thus acts as a *paraphase amplifier*.

Scale changing amplifier : If $\frac{Z_2}{Z_1}=n$, where n is a real constant then we find from eq. (2) (art 24·1) that $V_0=-nV_S$. Thus the scale has been multiplied by a factor $-n$. For it Z_1 and Z_2 are selected as precision resistors, *e.g.* $Z_1=R$ then $Z_2=nR$.

Phase shifting amplifier : If $|Z_1|=|Z_2|=|Z|$ and if the phase angles of Z_1 and Z_2 are different, then the operational amplifier shifts the phase of a sinusoidal input voltage without affecting its amplitude. Let us write

$$|Z_1| = |Z| e^{j\theta}$$

and

$$|Z_2| = |Z| e^{j\phi}$$

then from eq. (2), art 24·1, we have

$$\frac{V_0}{V_S}=-\frac{Z_2}{Z_1}=e^{j(\phi-\theta)}=1 \ \underline{|(\phi-\theta)}$$

That is, only phase of the input has been shifted through $(\phi - \theta)$ to appear as output.

Summing amplifier : It is the same as inverting amplifier except that it has several input terminals Virtual ground exists at the inverting terminal due to feedback and the input current to the ideal amplifier (as $Z_i = \infty$) is zero. Thus current equation for the node at the inverting terminal aives

$$\frac{V_1}{r_1} + \frac{V_2}{r_2} + \quad + \frac{V_n}{r_n} + \frac{V_0}{R_2} = 0$$

or
$$V_0 = -\left[\frac{R_2}{r_1} V_1 + \frac{R_2}{r_2} V_2 + \ldots + \frac{R_2}{r_n} V_n\right]$$

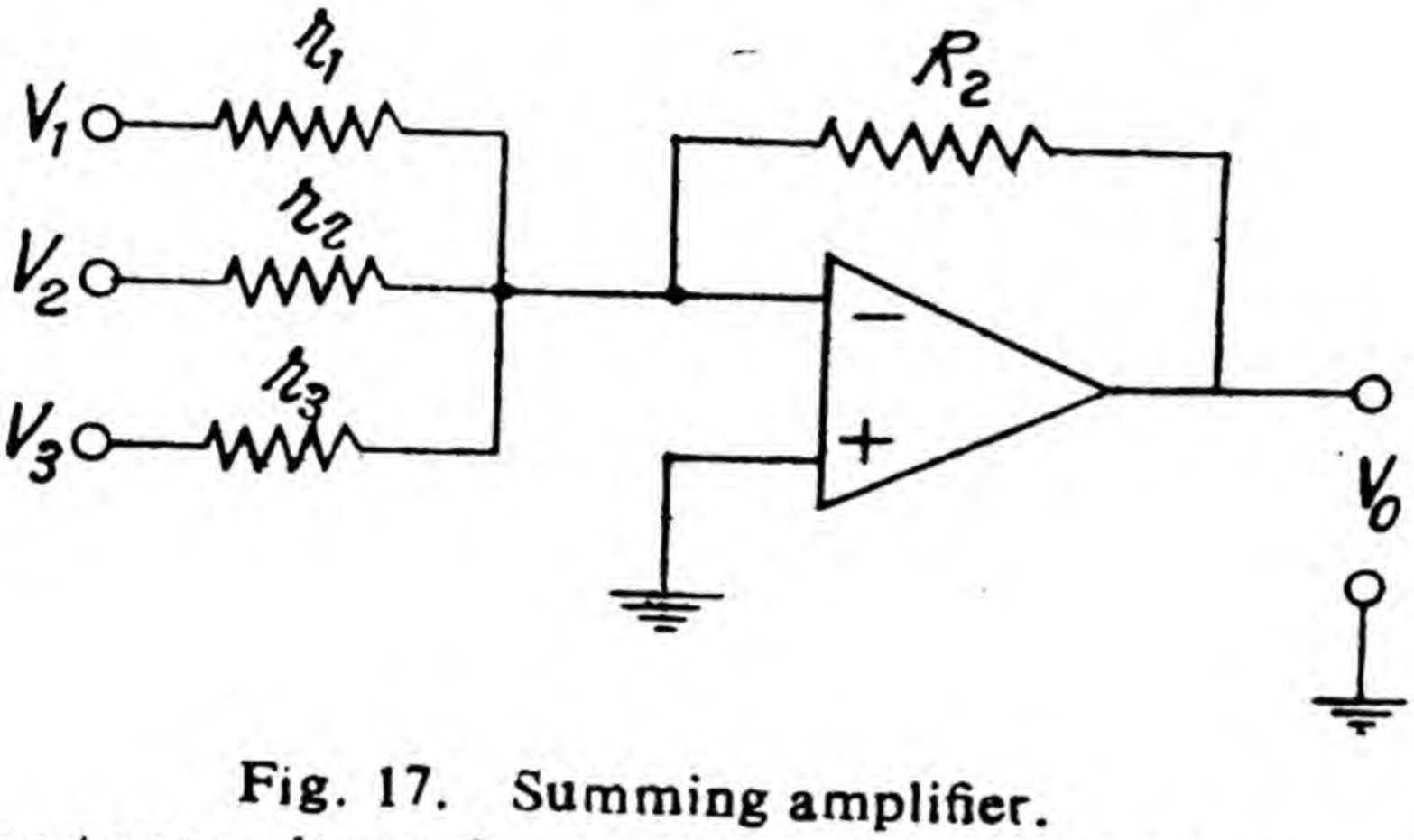


Fig. 17. Summing amplifier.

Thus output voltage is equal to the negative weighed sum of the input voltages. If $r_1 = r_2 = \ldots = r$, then output voltage is

$$V_0 = -\frac{R_2}{r}(V_1 + V_2 + \ldots + V_n),$$

i.e. the output voltage is proportional to the sum of input voltages. Hence the circuit behaves as a summing amplifier. The circuit is used widely to form linear combinations of various signals in analog computers. An important advantage of this circuit is that there is minimum interaction between input sources due to virtual ground existing at the input terminals.

(2) Voltage Follower : Refer to fig. 4, art 24 ! of noninverting OP-AMP. If we remove R_2 and $R_1 = \infty$ then from equation (4), we have

$$\frac{V_0}{V_S} = 1 + \frac{R_2}{R_1} = 1$$

i.e. $$V_0 = V_S,$$

Thus output voltage follows the input voltage. Circuits acts as a voltage follower.

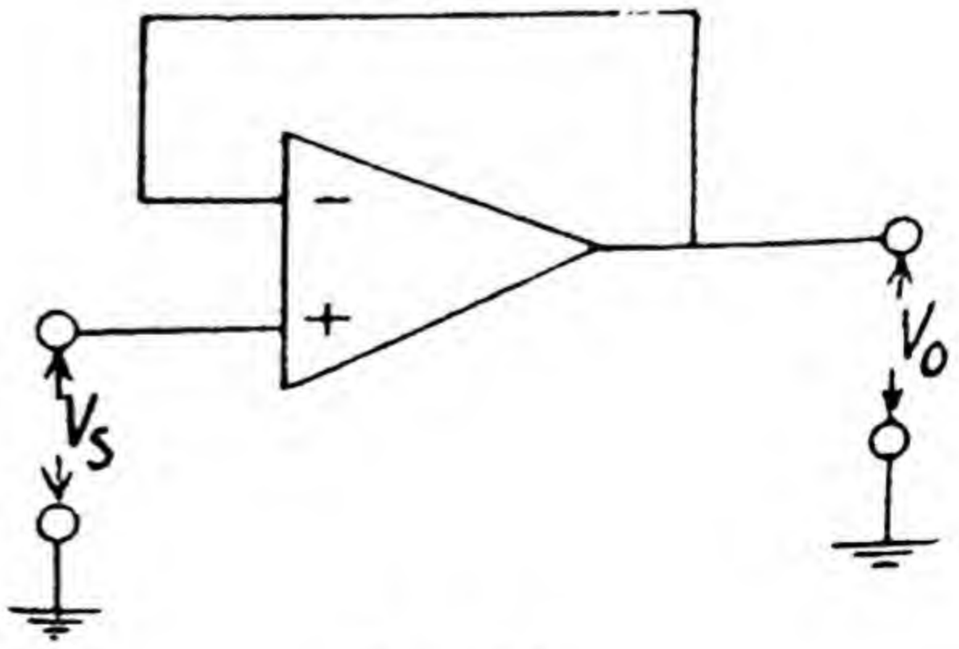


Fig. 18.

(3) Voltage to Current Converter : For magnetic deflection in TV picture tube, it is required to convert voltage signal into a proportional current signal. Circuit for such a converter is shown in

fig 19. The load impedance, Z_L, is floating (neither side grounded). Since virtual ground exists at the input terminals, no current flows into input terminals and therefore the current, I, which flows through R_1 also flows through Z_L. For a single input, V_S, we have

$$I = \frac{V_S(t)}{R_1}$$

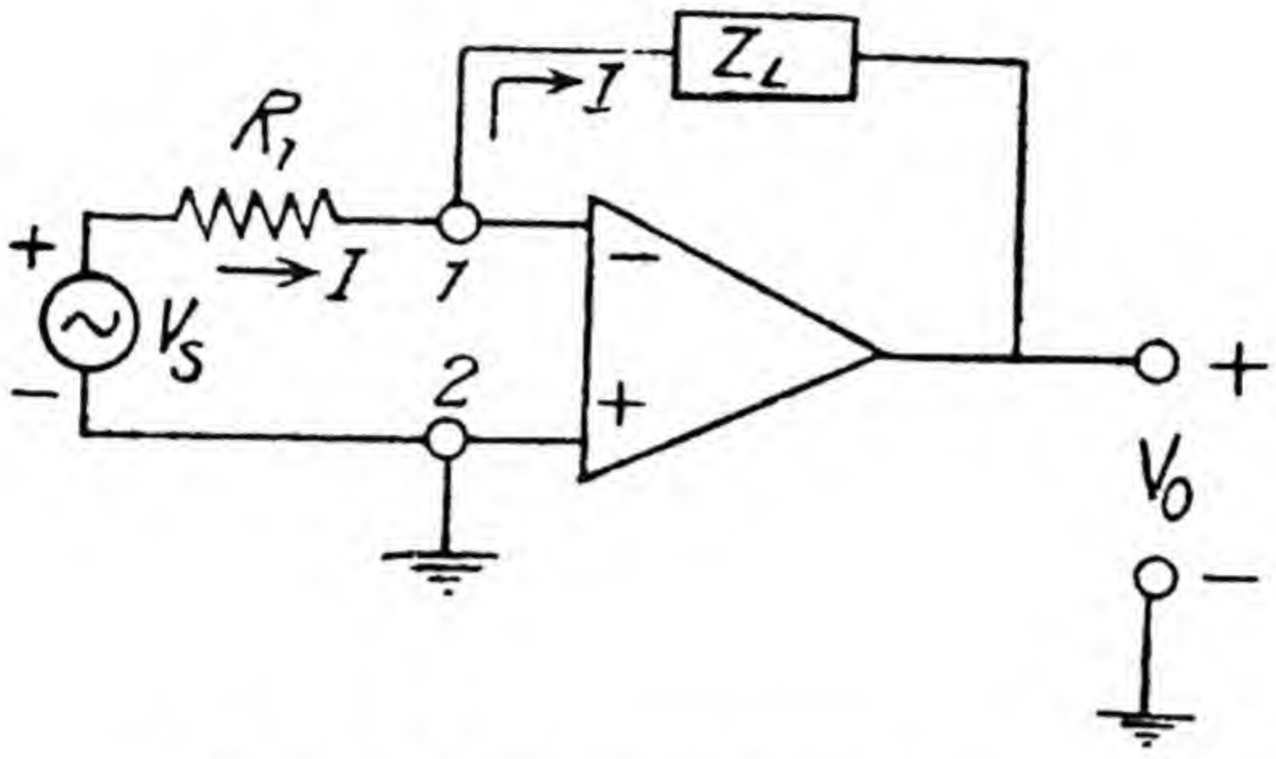


Fig. 19. Voltage to current converter.

Thus load current, I, is independent of load impedance and is proportional to the input signal. It can be used for deflection coils.

(4) Current to Voltage Converter : Circuit is shown in fig. 20. In this circuit output voltage is proportional to the input current. As there is virtual ground at the input terminals, current in R' is zero and input current, I_S, flows through feedback resistor R_2. Output voltage is given by

$$V_0 = -I_S R_2,$$

Thus output voltage is proportional to the input current. C_2 is connected to reduce high frequency noise. Such an input current can be provided by a photocell or photomultiplier tubes as they give an output current which is independent of load.

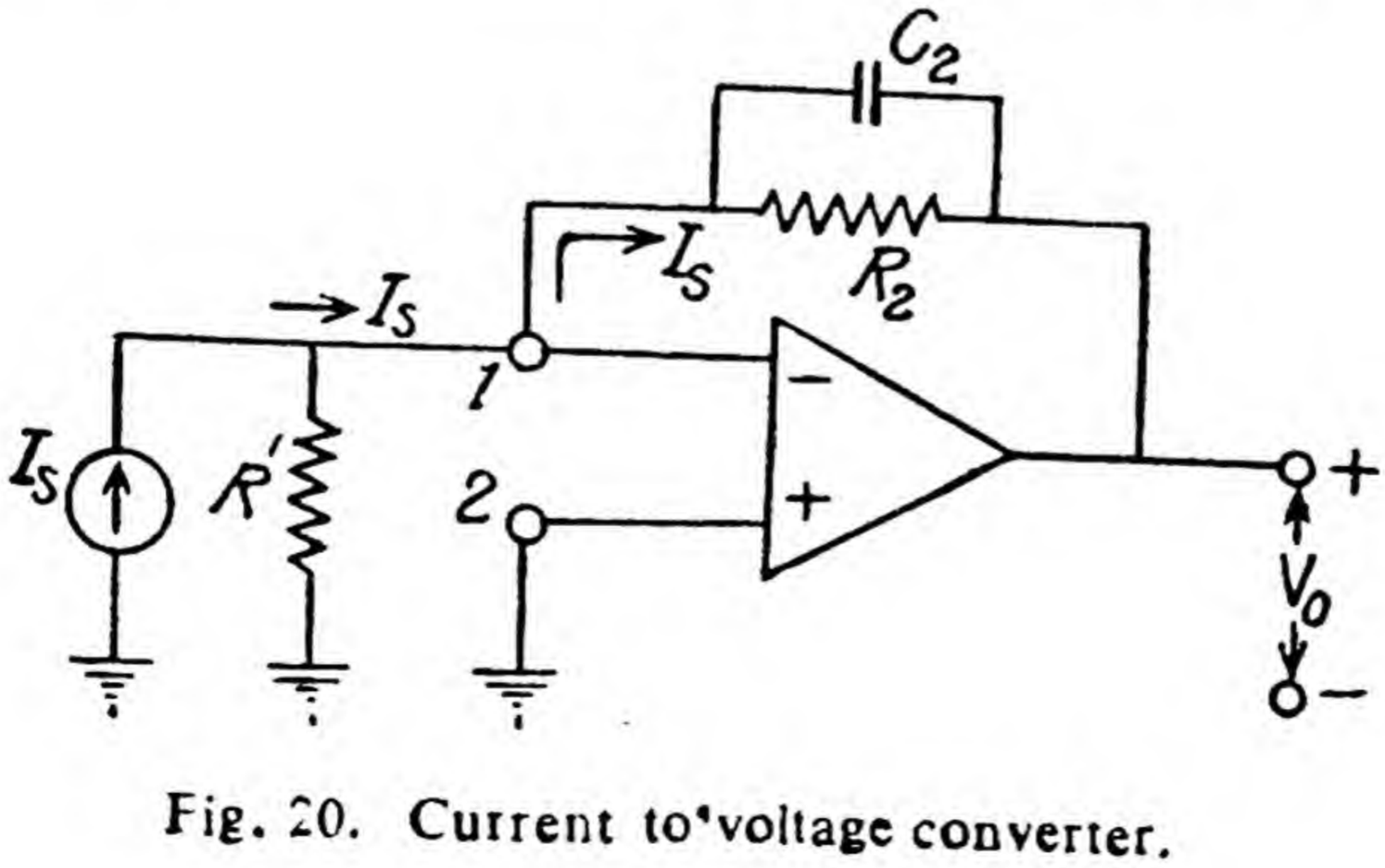


Fig. 20. Current to voltage converter.

(5) Integrator : In fig. 21, an integrator is shown. It is an inverting OP-AMP in which feedback resistor R_2 has been replaced by a capcitor, C. Feedback through the capacitor forces a virtual ground to exist at the inverting input terminal. It means the voltage across, C, is simply the output voltage, V_0. We can write

$$V_0(t) = -\frac{q}{C} + V_0(0)$$

$$= -\frac{1}{C}\int_0^t I dt + V_0(0),$$

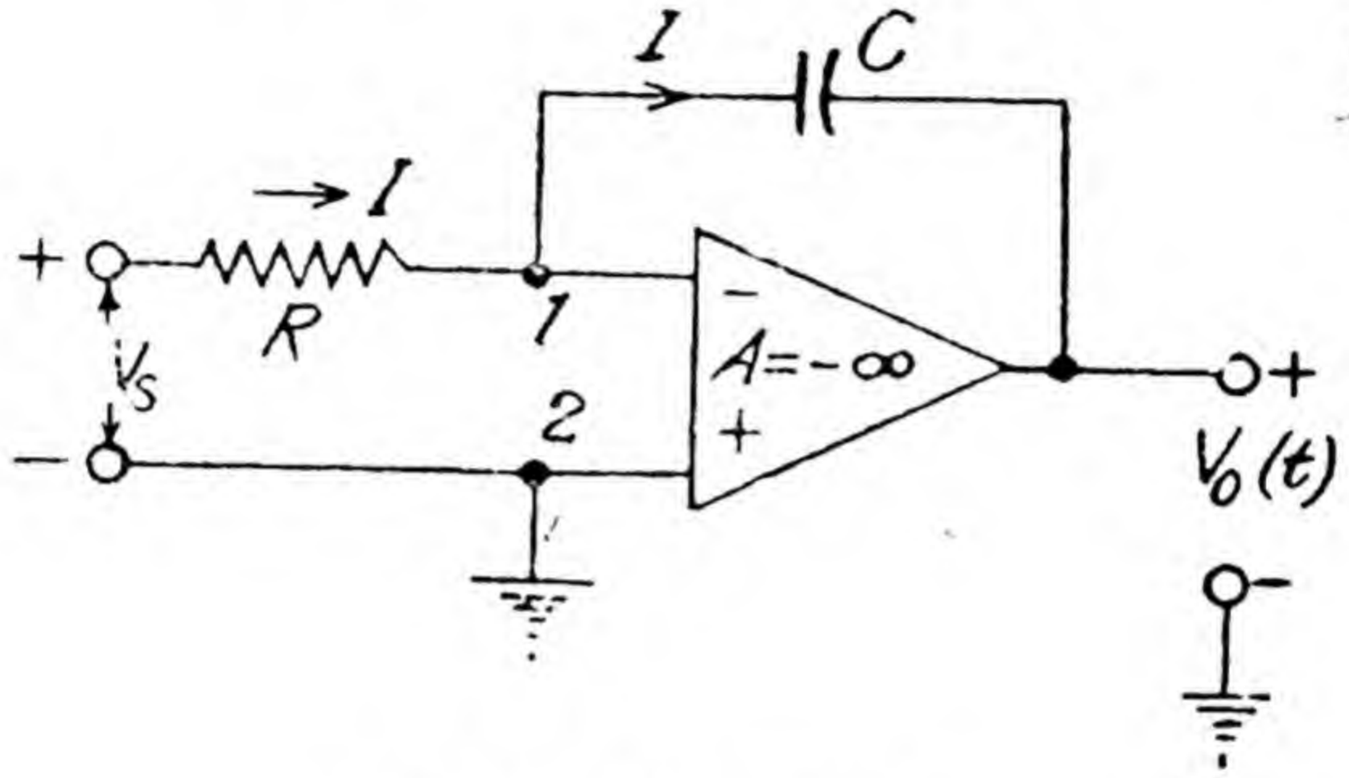


Fig. 21. Integrator

where $V_0(0)$ is the initial voltage. Since input current to the ideal amplifier is zero, we put

$$I = \frac{V_S(t)}{R}$$

so that

$$V_0(t) = -\frac{1}{C}\int_0^t \frac{V_S(t)}{R}\, dt + V_0(0)$$

Thus output voltage is proportional to the integral of the input voltage. Initial voltage $V_0(0)$ can be set to any desired value by introducing additional circuitry. If input voltage is constant then

$$V_0 = -\frac{V_S(t)}{RC}.$$

Thus output rises linearly with time $-a$ ramp. This is called Miller sweep (see chapter on Voltage and Current Sweep Generators).

(6) Differentiator : In inverting OP-AMP, we replace input resistance by a capacitor to design a differentiator (fig. 22). Because of virtual ground at the inverting terminal, we have

$$I = \frac{dq}{dt} = \frac{d}{dt}(CV_S)$$

$$= C\frac{dV_S}{dt}$$

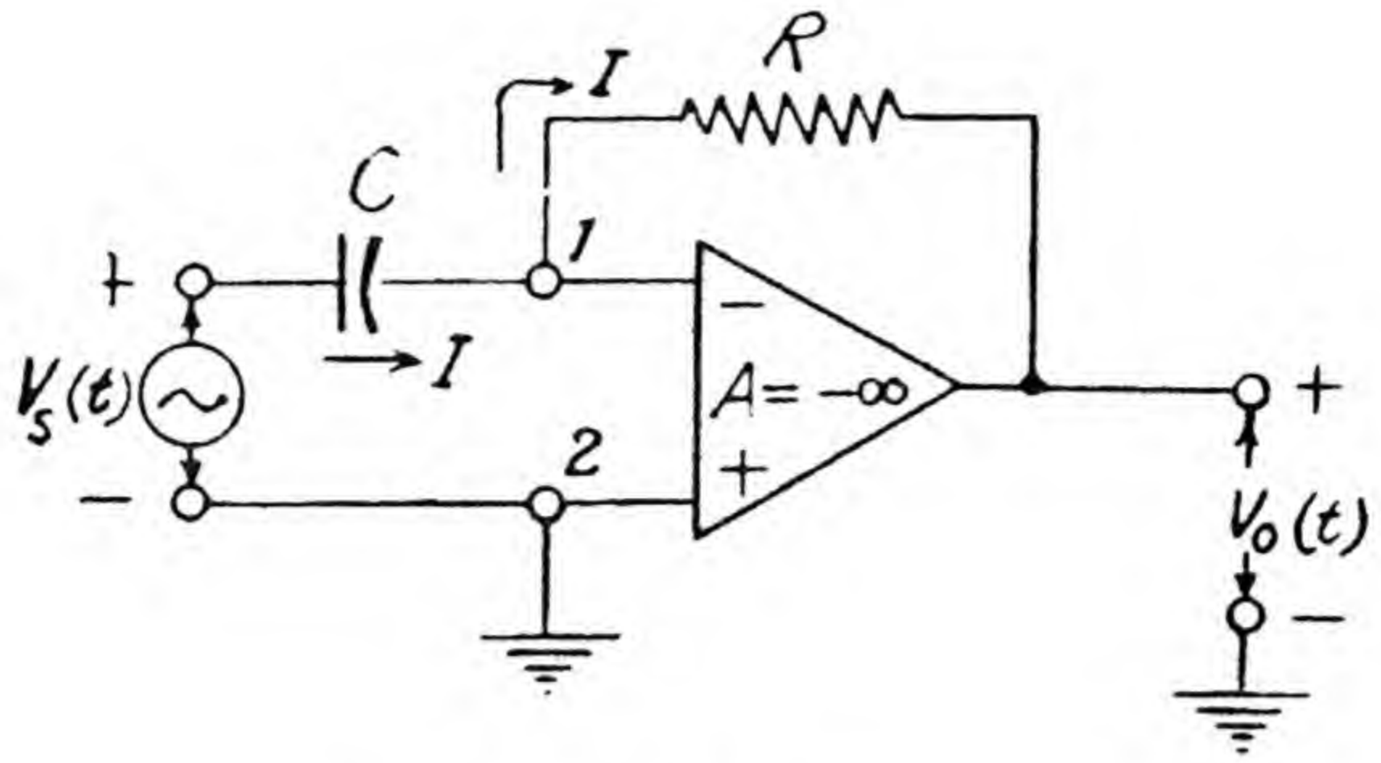


Fig. 22. Differentiator.

The output voltage is

$$V_0(t) = -RI_1$$
$$= -RC \frac{dV_S}{dt}(t)$$

Thus output is proportional to the time derivative of the input voltage. If input signal is

$$V_S(t) = |V_S| \sin \omega t$$

then output is

$$V_0(t) = -RC\,|V_S|\,\omega \cos \omega t$$

which increases linearly with frequency and the differentiator circuit has high gain at high frequencies. It thus also amplifies high frequency components of amplifier noise. Due to this reason it is rarely used in analogue computers.

(7) Analog Computation : Let us solve the quadratic equation

$$\frac{d^2v}{dt^2} + B\frac{dv}{dt} + Cv - v_1(t) = 0 \qquad \ldots(1)$$

where B and C are real positive constants and $v_1(t)$ is a given function of time. We can put this equation as

$$\frac{d^2v}{dt^2} = -B\frac{dv}{dt} - Cv + v_1(t) \qquad \ldots(2)$$

Refer to fig. 23. We know that output of an integrator is

$$V_0 = -\frac{1}{C}\int_0^t \frac{V_S(t)}{R}\,dt + V_0(0)$$

If time constant of the integrator, $RC = 1$, then

$$V_0 = -\int_0^t V_S(t)\,dt + V_0(0).$$

Let us assume that input voltage at the integrator, I_1, is

$$V_S = \frac{d^2v}{dt^2}$$

so that output of this integrater, available at terminal 2, is

$$V_0' = -\frac{dv}{dt} \qquad \text{(taking } V_0(0) \text{ as zero)}$$

But this is the input of second integrator, I_2, and therefore output of it will be

$$V_0'' = -\int_0^t \left(-\frac{dv}{dt}\right) dt = v$$

and will be available at terminal–3.

Further A_1 is a summing amplifier. At its input two voltages $v_1(t)$ and $\left(-\frac{dv}{dt}\right)$ are fed. The output voltage will be

$$= \frac{R}{R_1}\left(\frac{dv}{dt}\right) - v_1(t).$$

This voltage alongwith voltage, v, of terminal–3 are fed to another summing amplifier A_2 so that finally we get as output of terminal–4, a voltage

$$= -\frac{R}{R_2}v - \frac{R}{R_1}\left(\frac{dv}{dt}\right) + v_1(t)$$

$$= -Cv - B\frac{dv}{dt} + v_1(t)$$

But according to eq. (2), it must equal $\frac{d^2v}{dt^2}$, which is the voltage assumed to have been fed at the input terminal–1. Hence the operation is completed by connecting terminal–4 to the input terminal–1. Solution of eq. (1) *i.e.* v is obtained by opening switch

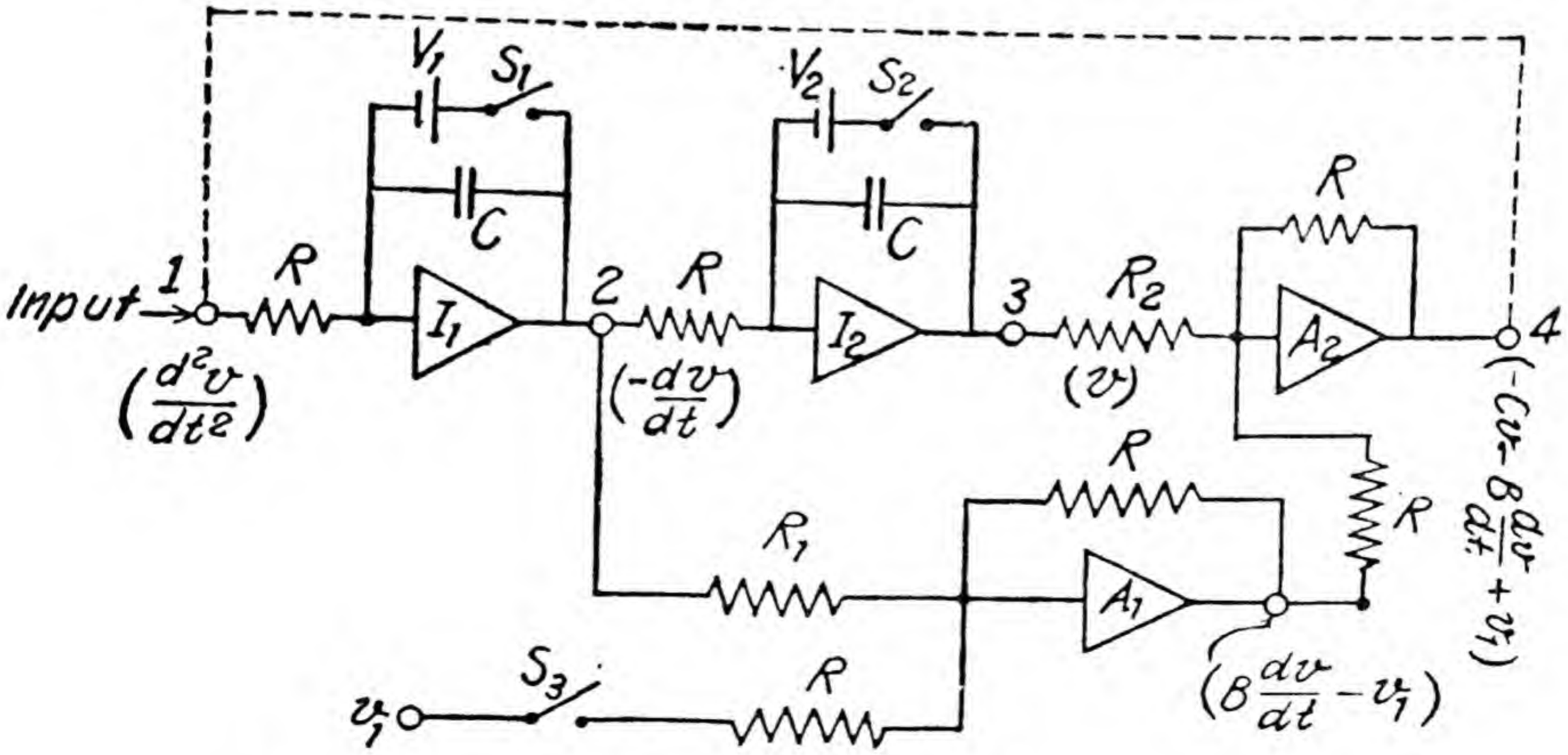


Fig. 23. I_1, I_2 integrators; A_1, A_2 adders. At $t=0$ S_1 and S_2 are opened and S_3 is closed.

S_1 and S_3 and closing S_3 simultaneously by means of relays at time $t=0$ and noting the voltage waveform at terminal–3.

(8) Logarithmic Amplifier : In this amplifier (fig. 24) output voltage is proportional to the logarithmic of the input voltage. Such an amplifier uses the non-linear volt-ampere relationship of a *p-n* juction. This relationship is given by

$$I_f = I_0 \left[e^{V_f/\eta V_T} - 1\right] \qquad \ldots(1)$$

where I_f=forward current, V_f=forward voltage drop, I_o=saturation current and

$$V_T = \frac{kT}{e},\ \eta = 1 \text{ for } G_e, \text{ and } \eta = 2 \text{ for } S_i.$$

If the operating range of V_f is restricted so that

$$e^{V_f/\eta V_T} >> 1$$

then $$I_f = I_0\, e^{V_f/\eta V_T}$$

or $$V_f = \eta V_T \,(\log_e I_f - \log_e I_0)$$

Now $I_f = I = \frac{V_S}{R}$, and $V_0 = -V_f$, then

$$V_0 = -\eta V_T \left(\log_e \frac{V_S}{R} - \log_e I_0\right) \qquad \ldots(2)$$

Thus input voltage, V_0. is proportional to the logarithmic of the input voltage, V_S, provided η, V_T and I_0 are constants (ignoring temperature effects).

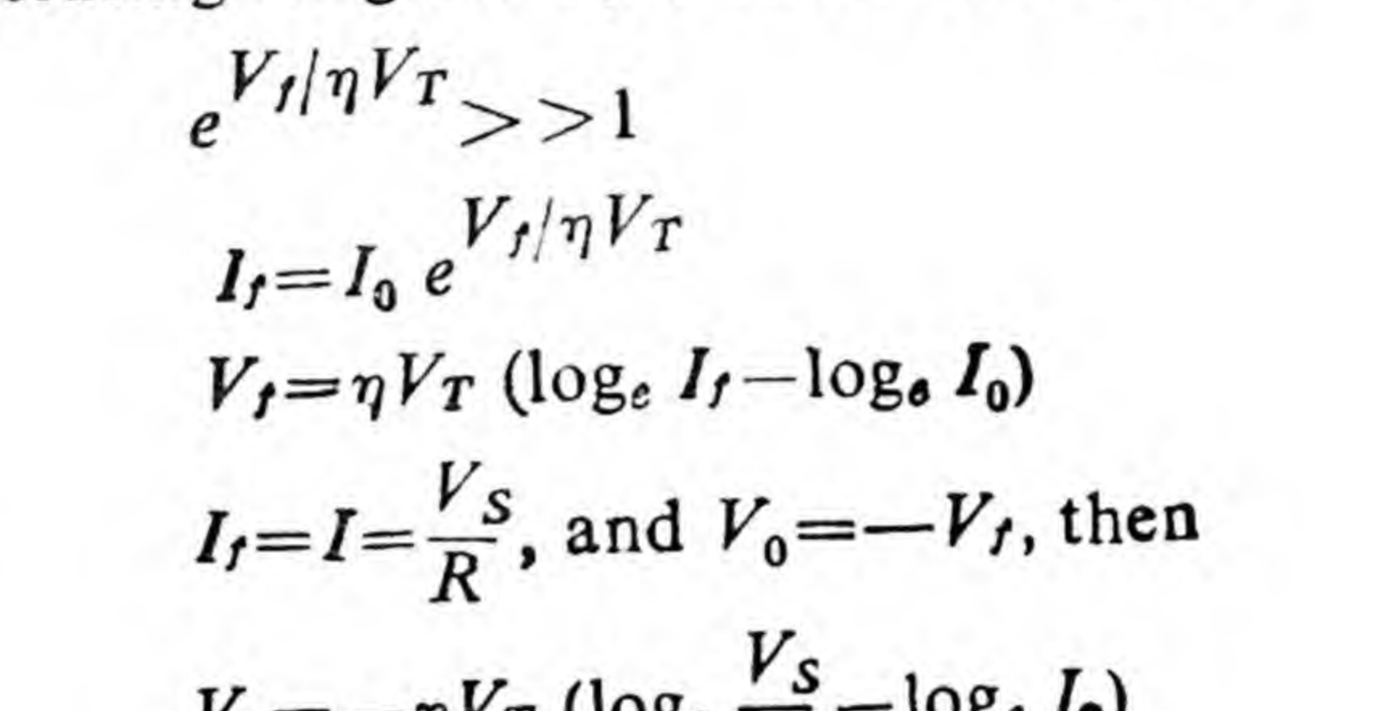


Fig. 24. Log amplifier.

(9) Antilogarithmic Amplifier : Refer to fig. 25. The voltage, V_1, at the noninverting terminal of OP-AMP–1 is given by

$$V_1 = \left(\frac{R_1}{R_1 + R_2}\right) V_S \qquad \ldots(1)$$

Voltage, V_2, at the output of OP AMP–1 is

$$V_2=\left(\frac{R_1}{R_1+R_2}\right)V_S-\eta\ V_T\ (\log_e I_f-\log_e I_0) \quad ...(2)$$

But also

$$V_2=-\eta V_T\ (\log_e I_2-\log_e I_o) \quad ...(3)$$

so that equating eqs. (2) and (3), we get

$$\left(\frac{R_1}{R_1+R_2}\right)V_S-\eta V_T\ (\log_e I_f-\log_e I_o)=-\eta V_T\ (\log_e I_2-\log_e I_0)$$

or

$$\left(\frac{R_1}{R_1+R_2}\right)V_S=\eta V_T\log_e (I_f/I_2) \quad ...(4)$$

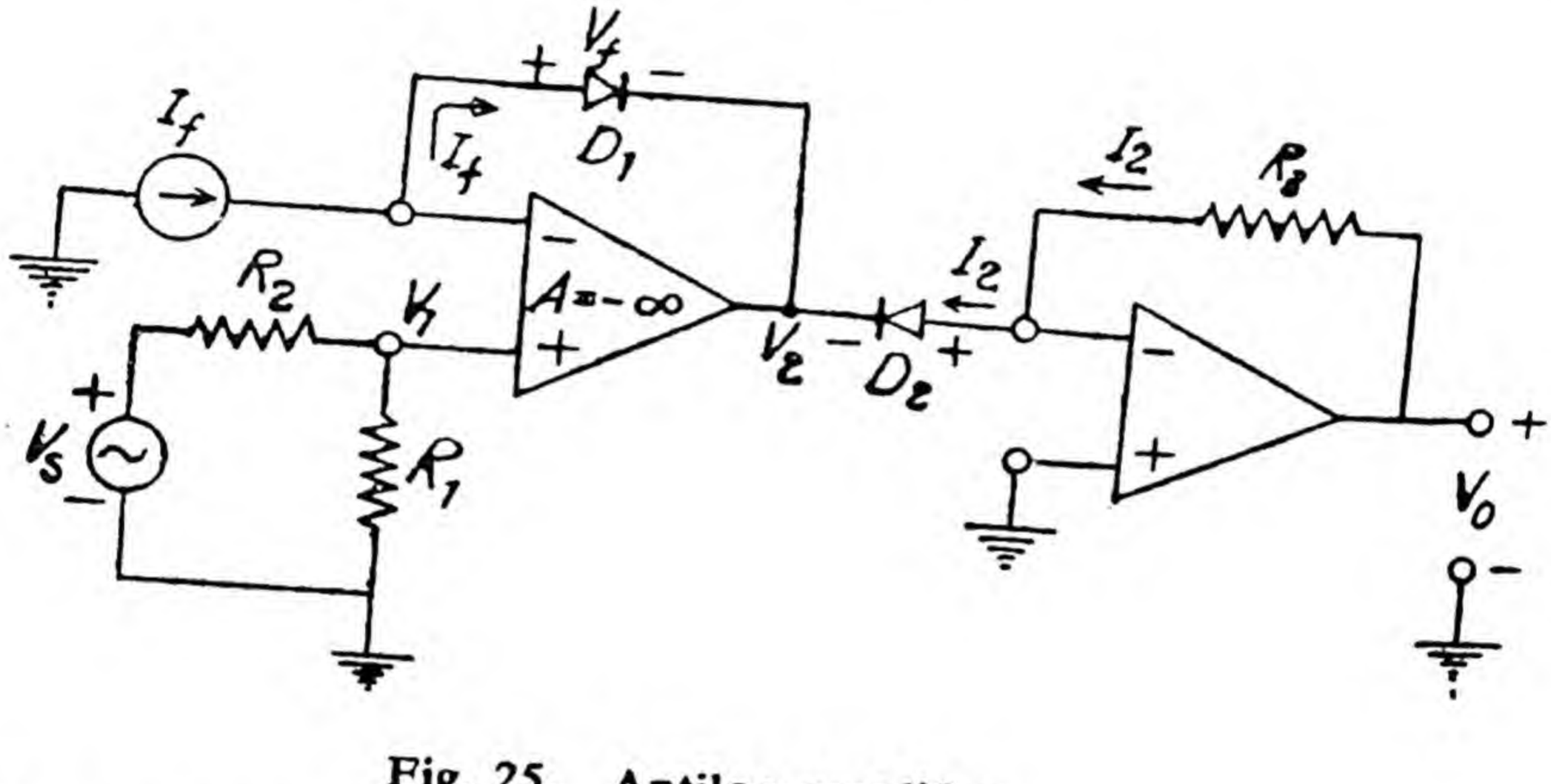


Fig. 25. Antilog amplifier.

The output voltage at OP-AMP–2 is

$$V_0=I_2R_3 \quad ...(5)$$

Putting in eq. (4) for I_2 from eq. (5) we get

$$\left(\frac{R_1}{R_1+R_2}\right)V_S=\eta\ V_T\log\ (I_f\ R_3/V_0)$$

or

$$\log\left(\frac{I_f\ R_3}{V_0}\right)=\left(\frac{R_1}{R_1+R_2}\right)\frac{V_S}{\eta V_T}$$

or

$$\log\left(\frac{V_0}{I_f\ R_3}\right)=-\left(\frac{R_1}{R_1+R_2}\right)\frac{V_S}{\eta V_T}$$

or

$$V_0=I_f\ R_3\ \text{antilog}\left[-\left(\frac{R_1}{R_1+R_2}\right)\frac{V_S}{\eta V_T}\right]$$

$$=k_1\ \text{antilog}\ (k_2\ V_S)$$

where k_1 and k_2 are constants. Thus output of OP-AMP–2 is proportional to antilog of input voltage, V_S.

(10) Astable Multivibrator : The operational amplifier operating as free running (astable) multivibrator is shown in fig. 26.

In free running multivibrator, there are two states which remain momentarily stable and the circuit switches repetitively between these two states. OP-AMP enjoys both positive and negative feedbacks and has a timing capacitor, C, at its inverting input terminal. *The output voltage v_o, is limited by the break down voltage $+V_{0s}$* and $-V_{0s}$ *of the two zener diodes* Z_1 *and* Z_2 connected back to back across the output terminals of the OP-AMP. *Thus v_o will be either* $+V_{0s}\left(\text{or } V_{0S}^{+}\right)$ *or* $-V_{0s}\left(\text{or } V_{0S}^{-}\right)$. A fraction $\beta = R_3/(R_2+R_3)$ of v_0 is fedback to the noninverting input. Thus in one state the amplifier output reaches a positive saturation level

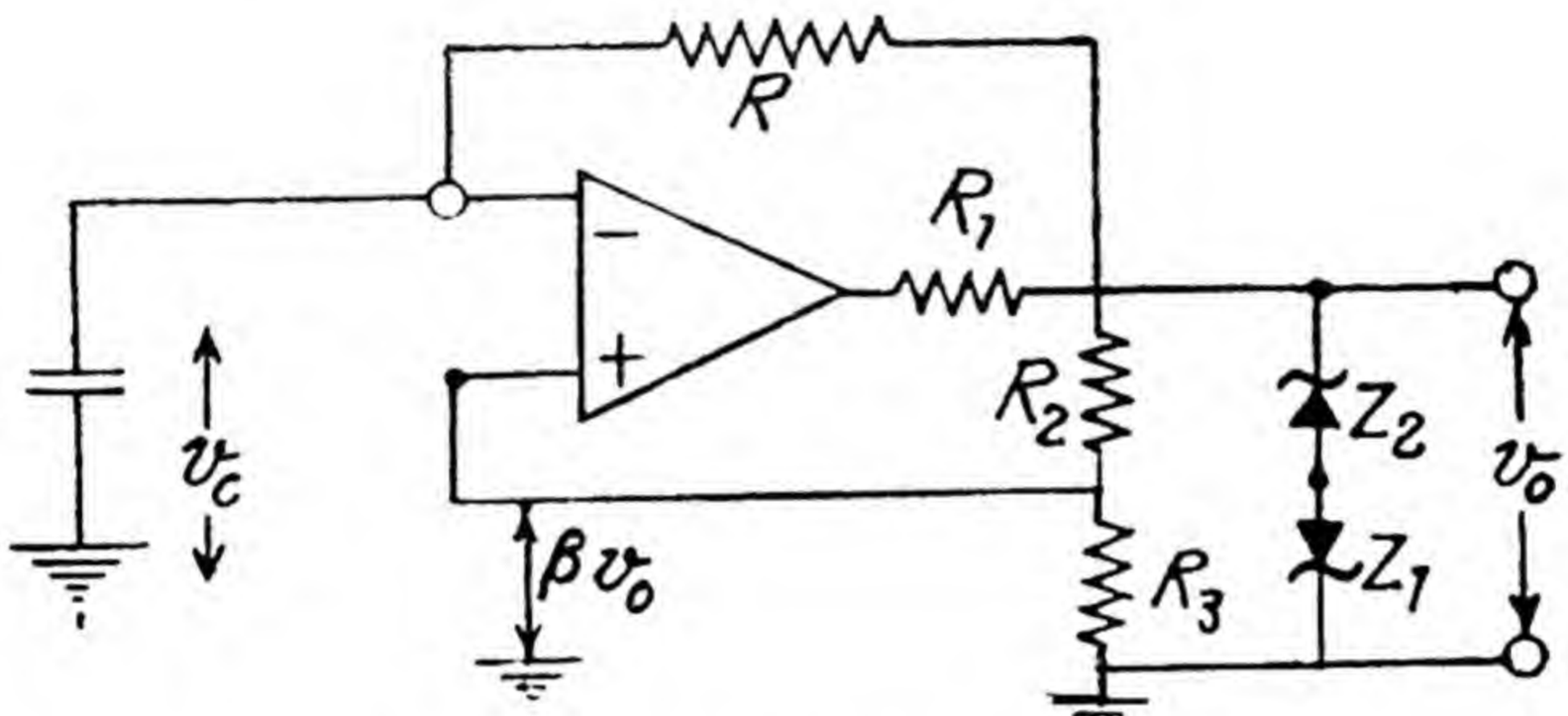


Fig. 26 Astable multivibrator.

($v_o = +V_{0s}$, diode Z_2) and in other state the amplifier output reaches a negative saturation level ($v_0 = -V_{0s}$, diode Z_1). The output waveform is thus a square wave (fig. 27).

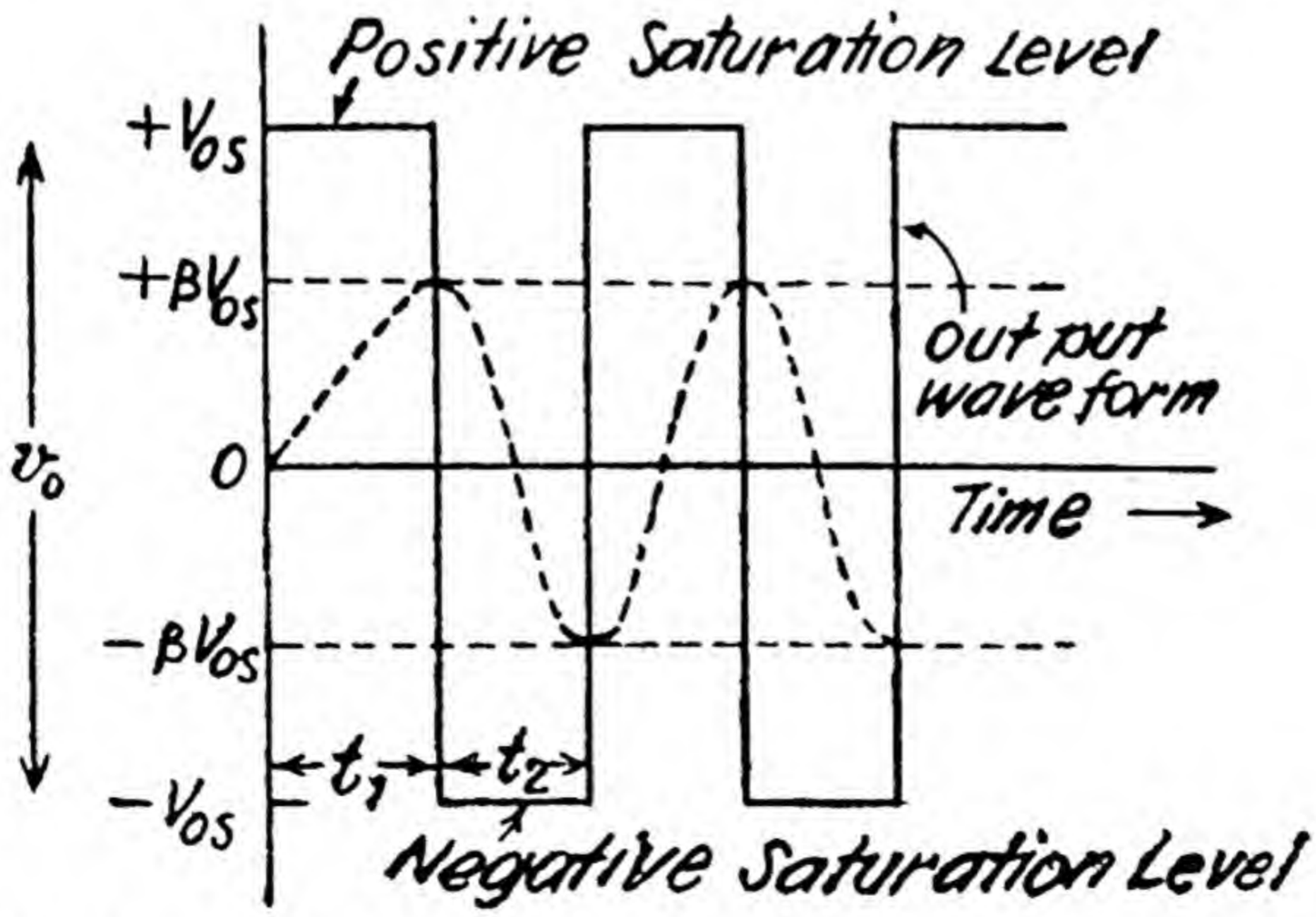


Fig. 27. Output waveform of free running multivibrator.

The input voltage, v_i, to the amplifier is

$$v_i = v_c - \beta\, v_0$$

when $v_i < 0$, or $v_c \lhd \beta\, v_0 = \beta\, V_{oS}^{+}$ capacitor, C charges exponentially towards V_{0S}^{+} through a time constant RC. The output remains constant at V_{0S}^{+} until $v_c = +\beta V_{0S}^{+}$ so that $v_i = v_c - \beta\, V_{oS}^{+} = 0$ *i e.*, the potential difference between the two input terminals approaches zero, and the amplifier output reverses to V_{0S}^{-}. Now C charges exponentially towards V_{0S}^{-}. Then $v_i = v_c - \beta\ v_o = v_c + \beta\, V_{oS}^{+}$. The output remains constant at V_{0S}^{-} until $v_c = -\beta\ V_{0S}^{-}$ at which $v_i = 0$ and the reversal of state takes place. These waveforms are shown in fig. 27.

Charging of capacitor starts from an initial voltage βV_{0S}^{-} and continues upto a voltage level βV_{oS}^{+}. If the charging could have continued, it would have reached a final level of V_{oS}^{+}. But as the charging terminates at βV_{oS}^{+}, the charging period t_1. is given by

$$t_1 = RC \log_e \frac{V_{0S}^{+} - \beta\, V_{0S}^{-}}{V_{0S}^{+} - \beta\, V_{0S}^{+}}$$

$$= RC \log_e \frac{V_{0S}^{+} - \beta\, V_{0S}^{-}}{V_{0S}^{+}\,(1-\beta)}$$

The second charging time from βV_{0S}^{+} to $-\beta V_{0S}^{-}$ will be

$$t_2 = RC \log_e \frac{V_{0S}^{-} - \beta V_{0S}^{+}}{V_{0S}^{-}\,(1-\beta)}$$

If $V_{0S}^{+} = V$, and $V_{0S}^{-} = -V$ then $t_1 = t_2$ so that we shall have period of one oscillation or time period as

$$T = t_1 + t_2$$

$$= RC \log_e \frac{V+\beta V}{V(1-\beta)} + RC \log_e \frac{-V-\beta V}{-V(1-\beta)}$$

$$= RC \log_e \frac{(1+\beta)}{(1-\beta)} + RC \log_e \frac{(1+\beta)}{(1-\beta)}$$

$$= 2RC \log_e \frac{(1+\beta)}{(1-\beta)} \qquad \ldots(1)$$

$$= 2RC \log_e \frac{1+R_3/(R_2+R_3)}{1-R_3/(R_2+R_3)}$$

$$= 2RC \log_e \left[1+\frac{2R_3}{R_2}\right]$$

If β is chosen to be 0·473, then

$$T = 2RC \qquad \ldots(2)$$

so that frequency of oscillations is

$$f = \frac{1}{T} = \frac{1}{2RC}$$

From eq. (1), we observe that time period is independent of saturation levels V_{oS}^{+} and V_{oS}^{-} and depends only on time constant RC and feedback factor β.

Astable multivibrator is very useful for fixed frequency applications in audio frequency range (10 c/s — 10 kc/s). At frequencies greater than 10 kc/s, the delay time of the amplifier for going from one state of saturation to another state of saturation becomes significant. Further, slew rate of OP-AMP also sets a limit on the rise and fall times of square wave output waveform.

(11) Bridge Amplifier : Circuit is shown in fig. 28. Operational amplifier is used here to amplify the output from a Wheatstone bridge, circuit which consists of a transducer (a device that con-

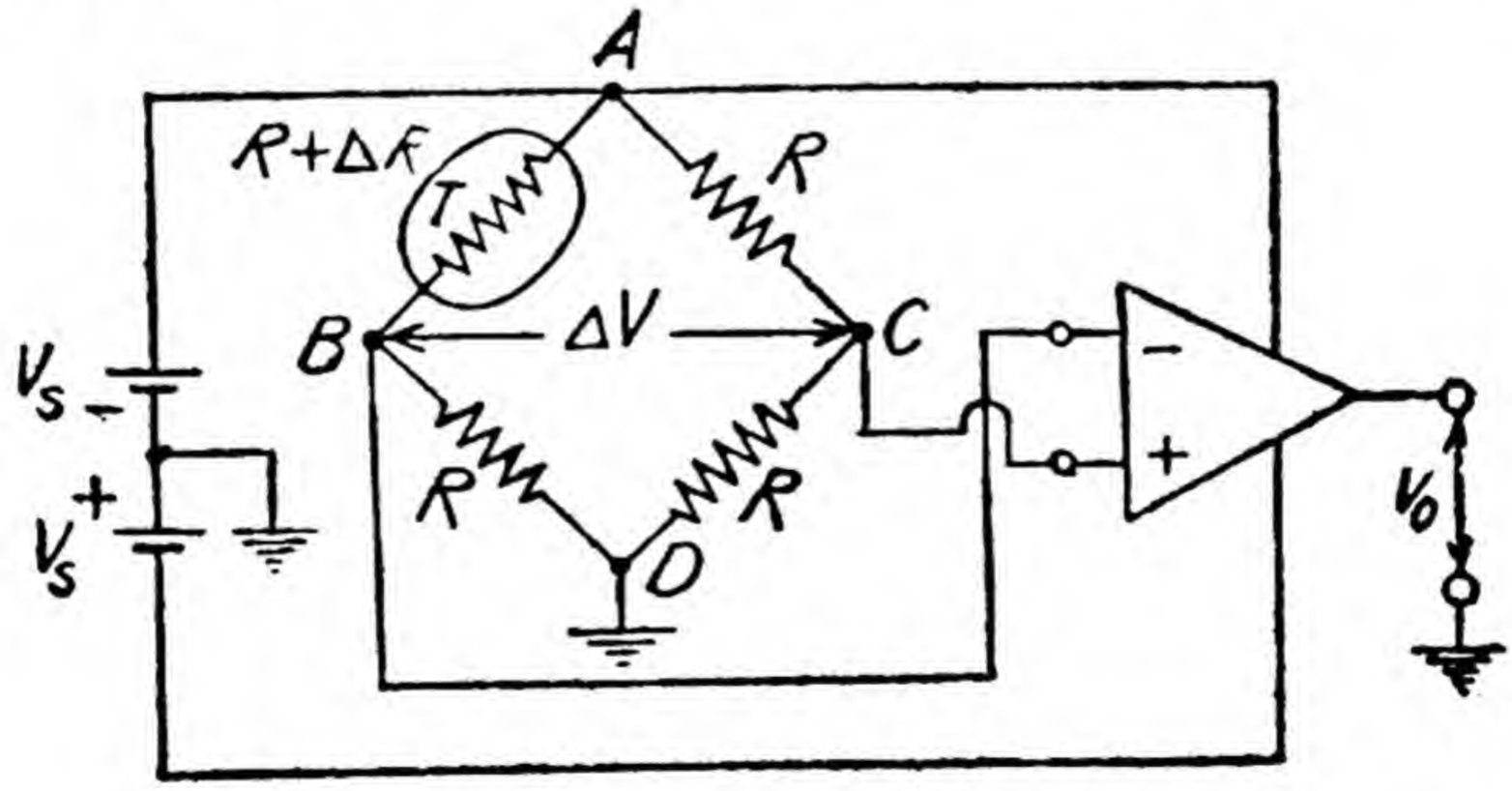


Fig. 28. A bridge amplifier.

verts a physical variable being measured into an electrical signal *e.g.*, a thermistor) is one of its arms. In fig. 28, T is a thermistor. Let the change in resistance of T with temperature be ΔR. Let a

constant voltage, V_S, be applied to the bridge. Then potential at point, B, will be

$$V_B = \left[\frac{R+\triangle R}{(R+\triangle R)+R}\right] V_S = \frac{1}{2R}\frac{(R+\triangle R)\,V_S}{(1+\triangle R/2R)}$$

and at C will be

$$V_C = \left(\frac{R}{R+R}\right) V_S = \frac{V_S R}{2R}$$

so that

$$\triangle V = V_B - V_C = \frac{V_S}{2R}\left[\frac{R+\triangle R}{(1+\triangle R/2R)} - R\right]$$

$$= \frac{V_S}{2R}\left[\frac{R+\triangle R-R-\triangle R/2}{(1+\triangle R/2R)}\right]$$

$$= \frac{V_S}{2R}\left[\frac{\triangle R/2}{1+\triangle R/2R}\right]$$

$$= \frac{\triangle R}{4R}\left[\frac{1}{1+\frac{\triangle R}{2R}}\right] V_S$$

If $\triangle R \ll 2R$, then

$$\triangle V = \frac{\triangle R}{4R} V_S.$$

Fig 29 shows that $\triangle R$ is large for most thermistors. Thus output voltage, $\triangle V$, is *not linear* with $\triangle R$ or temperature. The purpose of operational amplifier here is to provide a usable signal

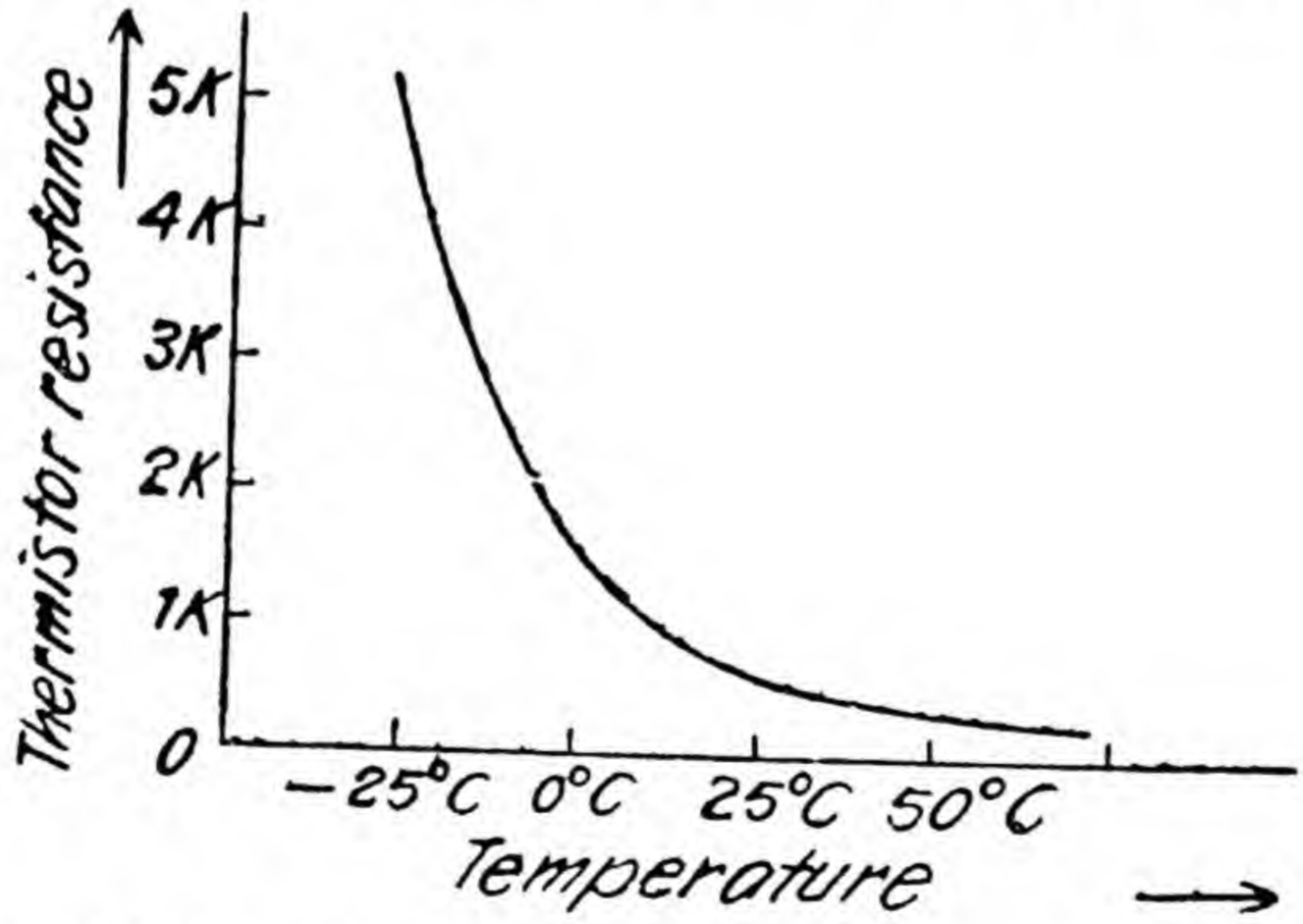


Fig. 29. Variation of thermistor resistance with temperature.

to drive an indicating instrument (a meter) *without introducing any additional nonlinearity in* $\triangle V$. This is possible (though not possible for any milliammeter) for OP-AMP as its high input impedance about 50M ohms) will not load down the bridge and so will not introduce additional non-linearity.

(12) Voltage Comparator : The function of voltage comparator is to compare the time varying voltage at one input with a

fixed reference voltage on the other. A differential amplifier (fig. 30*a* and 30*b*) can be used as a voltage comparator (fig. 31*a*, 31*b*)

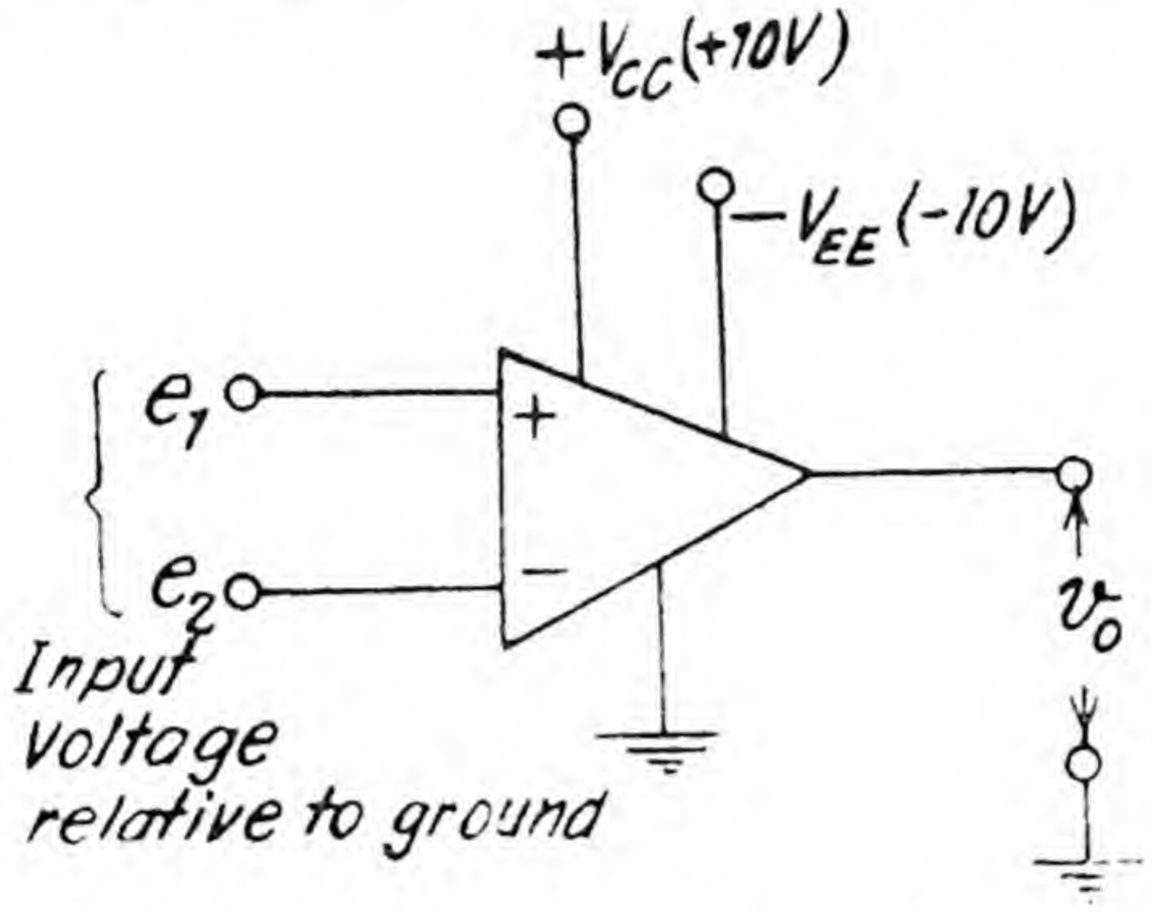


Fig. 30 (*a*) Symbol of differential amplifier.

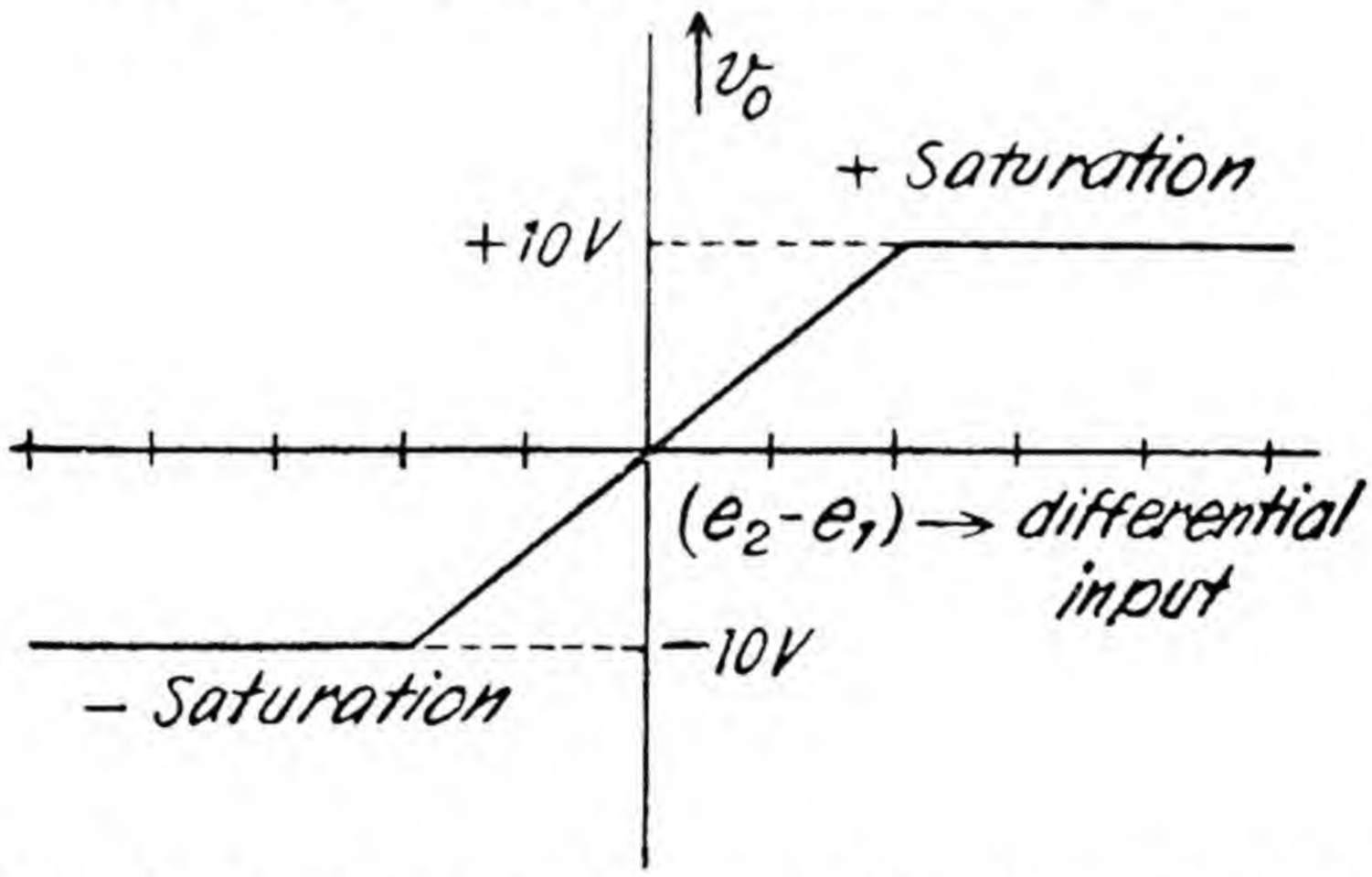
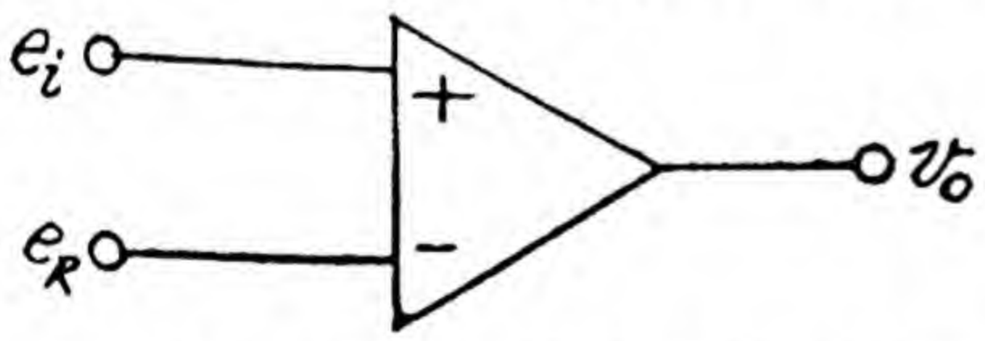
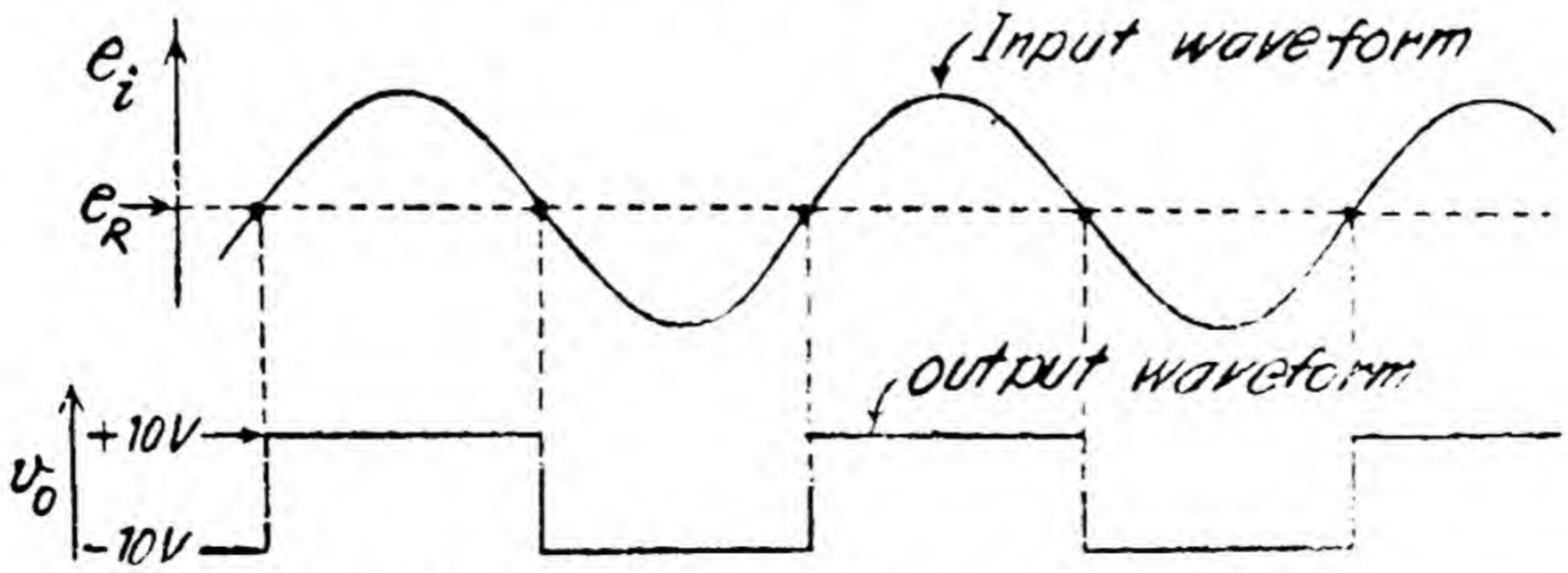


Fig. 30 (*b*) Differential amplifier input/output graph.



$$\begin{cases} +10V \text{ for } e_i > e_R \\ -10V \text{ for } e_i < e_R \end{cases}$$

Fig. 31 (*a*) A comparator; e_I is input voltage, e_R is reference voltage. Output v_0 is $+10V$ $(+V_{CC})$ when $e_i > e_R$ and v_0 is $-10V$ when $e_i < e_R$.



Differential amplifier of fig. 30 (*a*) has been shown as comparator in fig. 31 (*a*), with a difference that here e_1 is shown as e_i (input voltage), e_2 as e_R (reference voltage, normally a constant value). Input signal e_i will be compared with other input a reference voltage, e_R.

Since voltage gain of OP-AMP is very high ($A \rightarrow \infty$) output, v_0, will reach its *positive* saturation value (a d.c. level, shown as $+10V$) whenever e_i becomes slightly greater than e_R (few millivolts), while output, v_0, will reach its *negative* saturation value (a dc level, shown as $-10V$) whenever e_i becomes slightly lower than e_R (fig. 31*b*). Thus output will *quickly jump* from one saturation to the other as e_i varies above and below e_R. Thus output waveform which is a square one, though input is a sine waveform, can be explained as follows :

The moment e_i is greater than e_R by a few millivolts (say even $2mV$), v_0 reaches atonce the *positive* saturation ($+10V$ here) value. It will remain at this level so long as e_i continues to be greater than e_R. But when $e_i < e_R$, then as e_i crosses zero, it becomes slightly less than e_R and therefore output jumps to *negative* saturation value ($-10V$) and continues to remain at this level so long as $e_i < e_R$. When e_i rises again to become greater than e_R, state of affairs is repeated. Thus we get a square wave output.

If $e_R = 0$, then output will jump from zero to the d.c. level ($+10V$ and $-10V$) every time the input signal e_i passes through zero. The OP-AMP then acts as a *zero-crossing detector*.

EXERCISES

1. Draw schematic block diagram of the basic OP-AMP with inverting and non-inverting inputs. Sketch their equivalent circuits.
2. Explain the significance of virtual ground in a basic inverting amplifier. How would you explain its existence ? What do you understand by closed loop and open loop gain of an OP-AMP. When a noninverting OP-AMP acts as a voltage follower ?
3. What is meant by an ideal differential amplifier ? What are differential gain and common mode gain of a differential amplifier ? Define the term common mode rejection ratio (CMRR). Discuss an emitter coupled differential amplifier.
4. Define the terms input bias current, input offset voltage, input offset current, full power bandwidth and slew rate. What are the factors which determine slew rate.
5. Describe the function of an OP-AMP as (i) an inverter (ii) a scale changer, (iii) a phase shifter, (iv) an adder, (v) an integrator, (vi) a differentiator.
6. Draw the circuit diagram of a bridge amplifier using of OP-AMP. Explain its function.
7. Discuss how an OP-AMP can be made to act as an astable multivibrator. Obtain an expression for frequency of oscillations.
8. Discuss how OP-AMP can be used as a comparator.

25 NON-LINEAR INTEGRATED CIRCUITS : LOGIC FUNDAMENTALS

Combinational and Sequential Circuits

We shall take up the subject starting from basic principles and definitions. First combinational logic will be discussed in detail upto 25·10 and sequential logic in brief there-after.

25·1. DIGITAL AND ANALOG METHODS :

The term digital refers to any process that is accomplished using discrete units, *e.g.* fingers, toes, (digits) etc. Each of these can be used as a unit or group of units to express a whole number. On the other hand, analog numbers are represented as directly measurable quantities such as volts, rotations and distances etc. Thus in analog method, the number 20 can be represented as a 200 degree rotation of the needle on a meter. This method has been used extensively in electronics to represent such phenomenon as intensity, frequency and time etc. For communication purposes both digital and analog methods are used. In digital method, an analog waveform is sampled, and the digital representation of the amplitude is transmitted at regular intervals so that nothing useful of the sampled waveform is left. Both the accuracy and economy have turned the preference of the people towards digital readout devices. In this chapter we shall be concerned only with digital methods.

25·2. NUMBER SYSTEMS USED IN DIGITAL ELECTRONICS:

As the term digital implies a system of coupling using discrete units, there are four systems of arithmetic :

(*i*) Decimal

(*ii*) Binary

(*iii*) Octal

(*iv*) Hexadecimal.

Digital systems use binary system extensively. We shall discuss the first two systems.

The Decimal Number System : This is frequently used number system in our daily life. It contains 10 unique symbols : 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9. Because it has symbols its base is used to be 10. To indicate digits greater than 9 the digits are arranged by columns on the left of decimal point, each column having a different weight or multiplying factor, *e.g.*, they are named as hundreds, tens and units. Similarly to represent digits less than 1, the position to the right of the decimal point are named as tenths, hundredths,

thousandths etc. For example number :

$$432{\cdot}45_{10}=4\times10^{3}+3\times10^{2}+2\times10^{1}+4\times10^{-1}+5\times10^{-2}.$$

The Binary Number system : The binary number system uses only two digits 0 and 1, as contrasted to the ten digits of the decimal system. The base (radix) for this system is 2 and the positions to the left or right of the binary point carry weights increasing or decreasing in powers of 2. For example, number :

$$\begin{aligned}101011_2&=1\times2^5+0\times2^4+1\times2^3+0\times2^2+1\times2^1+1\times2^0\\&=32+0+8+0+2+1\\&=43_{10}\text{ in decimal system}\end{aligned}$$

and the number

$$\begin{aligned}0{\cdot}1101&=1\times2^{-1}+1\times2^{-2}+0\times2^{-3}+1\times2^{-4}\\&=\frac{1}{2}+\frac{1}{4}+0+\frac{1}{16}\\&=0{\cdot}500+0{\cdot}250+0{\cdot}062\text{ in decimal system}\\&=0{\cdot}812_{10}.\end{aligned}$$

Converting Binary to Decimal : This process is easy and has been illustrated above. The column weights for each 1 appearing in the number are noted and are added to all other column weights containing a 1. For example to convert 10111_2 to decimal, we write

$$\begin{aligned}10111_2&=1\times2^4+0\times2^3+1\times2^2+1\times2^1+1\times2^0\\&=16+0+4+2+1\\&=23_{10}.\end{aligned}$$

Converting Decimal to Binary : For this, number is successively divided by 2 and its remainders recorded. The final binary result is obtained by assuming all the remainders, with the last remainder being the most significant bit (*MSB*). For example, to convert 43_{10} to binary, we proceed as follows :

2)	43	Remainder
2)	21 — 1	
2)	10 — 1	
2)	5 — 0	
2)	2 — 1	
2)	1 — 0	
	0 — 1	

Reading the remainders from the bottom to top,

$$43_{10}=101011_2.$$

Another number 200_{10} can be converted as follows :

2)	200	*R*
2)	100	−0
2)	50	−0
2)	25	−0
2)	12	−1
2)	6	−0
2)	3	−0
2)	1	−1
2)	0	−1

Reading the remainders from the bottom to the top, the result is

$$200_{10}=11001000_2.$$

Binary Addition and Subtraction : The process of addition using binary numbers is basically identical to that using decimal notation. Prior to addition and subtraction, let us explain carry and borrow in binary system.

When a given binary column is filled (contains a 1) and when the digit is to be implemented by 1 (added to by 1) then the column in question is returned to zero and a carry is placed in the next higher order column.

For example

		cr	cr		
0	0	. 1	1		(number)
			1		(add 1)
		1	0	0	

OR in more detail

	column numbers			
	8	4	2	1
	0	0	1	1
				1
		1	0	0

In the 1's column, $1+1=0$, with carry of 1

In the 2's column, $1+1=0$, with carry of 1.

In the 4's column, $0+1=1$.

For borrow, a binary digit moved from a more significant column becomes a borrow, and in the new column it carries a weight greater than before by the base of the system. For example, a 1 in the third column (or 4's column) is to be moved to the second column

$$0100 \rightarrow \overset{1}{0}0\overset{1}{+}0$$

This peculiar notation is neceesary because a 1 moved one column to the right is now weighted twice its previous value. Twice 1 is 2, but there is no digit 2 in the binary system. These

are then treated as individual digits as illustrated in forthcoming subtraction

Addition : Add 1011_2 and 110_2, concept of carry is to be used.
In the 1's column, $1+0=1$.
In the 2's column $1+1=0$, with a carry of 1.
In the 4's column $1+0+1=0$, with a carry of 1.
In the 8's column $1+1=0$ with a carry of 1 to next higher column, but this is the last column of the number and therefore it is written as $1+1=10$, which is read as "zero and a carry of 1".

	column number			
	8	4	2	1
	1	0	1	1
+		1	1	0
	1	0	0	0	1

Subtraction : Subtract 01111_2 from 11011_2.
concept of borrow is to be used in association with rule that

0 minus 0 equals 0
1 minus 0 equals 1
1 minus 1 equals 0.

As stated earlier 0 minus 1 will require a borrow from next higher-order column which is to be written as $\overset{1}{+}$ on the column demanding a borrow. Let us depict them in the given problem.

$$\begin{array}{cccccc} & & 1 & & & \\ & & + & 1 & & \\ 0 & \nearrow & 1 & + & & \\ & & 0 & \nearrow & 1 & \\ 1 & & 1 & 0 & 1 & 1 \quad =27_{10} \\ - \ 0 & & 1 & 1 & 1 & 1 \quad =15_{10} \\ \hline 0 & & 1 & 1 & 0 & 0 \quad =12_{10} \end{array}$$

As another example let us subtract one from 100_2.

$$\begin{array}{ccccc} & & 1 & & \\ & 1 & + & & \\ & + & 1 & & \\ 0 & \nearrow & 1 & & \\ 1 & & 0 & 0 & =4_{10} \\ - & & & 1 & =1_{10} \\ \hline 0 & & 1 & 1 & =3_{10} \end{array}$$

25·3. BINARY ARITHMATIC IN COMPUTERS :

Two's complement arithmatic to represent negative numbers (Addition of negative numbers) :

In order to represent both positive and negative numbers in a computer, an additional bit, called the sign bit, can be added to each number. If this bit is a 0, the number is positive. If it is a 1, the number is negative. This system is called a signed magnitude system. Examples are in a six bit hardware,

$$000011 = +3$$
$$100011 = -3.$$

However, under this system a great deal of manipulation is necessary to add a positive number to a negative number and is rather impractical.

In a computer, due to its finite word length for which it is designed, a number is to be represented in a specific number of bits. For example six bit word length, eight bit word length, twelve bit word length etc. Suppose system is only of eight bit and we are to add the following :

(*i*)	1011	1011	(*ii*)	1010	0001
+	0100	0100	+	1101	0111
	1111	1111		0111	1000

Note that carry from eight bit in the second addition is to be discarded because ninth bit is outside the computer capability. It is exactly same as in a two bit decimal system, carry of second bit (or tenths column) is to be discarded *viz.*

$$\begin{array}{r} 64 \\ +\ 73 \\ \hline 37 \end{array}$$

where carry of tenth's column is discarded. This type of arithmatic is called modulus arithmatic. The two's component system is used to represent negative numbers using modulus arithmatic. Assume a bit number length then negative numbers are represented by the transformation $2^n - N$, where N is the positive representation of the number (*i.e.* additional bit or sign is 0). Let us take three examples ;

Ex. 1. *Express the decimal number* -5 *in two's complement form in a eight bit system.*

Thus n=8, and positive representation of -5 in eight bit is 00000101. Thus

$2^8 =$	1	0000	0000
Subtract 5	−	0000	0101
Result	$-5 =$	1111	1011

Ex 2. *Express the decimal number* -15 *in* 12*-bit two's component form*

$2^{12} =$	1	0000	0000	0000
Subtract 15	−	0000	0000	1111
	$-15 =$	1111	1111	0001

Ex. 3. *Express the decimal nnmber* —16,000 *in* 16 —*bit two's complement form.*

	$2^{16}=1$	0000	0000	0000	0000
Subtract	16,000—	0011	1110	1000	0000
Result	—16,000=	1100	0001	1000	0000

There is another method of computing two's complement of a number. It is :

(*i*) write the positive expression of that number,

(*ii*) then complement all the bits of that expression (change all the 1's to 0's and all the 0's to 1's).

(*iii*) then add one and get the result.

Ex. 1. *Express decimal number* —4 *in* 8—*bit two's complement form.*

Positive expression of number	0000	0100
Complement of above	1111	1011
Add one in complement	+0000	0001
Result	—4=1111	1100

Thus —4 equals 1111 11 0.

Ex. 2. *What negative value does* 1001 1011 *represent* ?

Number =1001 1011

Its two's complement =0110 0100

Therefore in decimal system

$$0110\quad 0100=101_{10}$$

Therefore

$$1001\quad 1011=-101_{10}$$

Summarising this article we write some two's complement numbers in the following table.

TABLE-1 : Two's Complement Numbers

Binary	Decimal
0111	+7
0110	+6
0101	+5
0100	+4
0011	+3
0010	+2
0001	+1
0000	0
1111	—1
1110	—2
1101	—3
1100	—4
10 1	—5
1010	—6
1001	—7
1000	—8

We note from the table :

1. In two's complement system, there will always be $2^{n-1}-1$ positive integers, 2^{n-1} negative integers, and one 0, for a total of 2^n unique states, where n is the number of bias in the binary word. In the table $n=4$ so that total $2^4=16$ states, out of which 7 positive integers, 8 negative integers and one 0.

2. The left most bit is a sign bit and do not express quantity. If it is 0, number is positive and if 1, number is negative.

Let us now take few more examples for the completeness of above two's complement method.

Ex. 1. *What negative value in decimal system does* 1001 1011 *represents* ?

Find two's complement for this purpose

Number $=1001\quad 1011$
it complement $=0110\quad 0100$
Add $1=0000\quad 0001$
Result $=0110\quad 0101$

Two's complement of Number $=0\times 2^7+1\times 2^6+1\times 2^5+0\times 2^4$
$+0\times 2^3+1\times 2^3+0\times 2^1+1\times 2^0$
$=64+32+4+1$
$=-101_{10}$

Thus $1001\quad 1011=-101_{10}$

Ex. 2. *Add* -5 *to* $+5$ *in* 8 *bit word length.*

First find two's compliment of -5, and then add with $+5$ with sign bit taking part in the process.

To find two's complement of -5

Positive expression of $-5=0000\quad 0101$
Complement of above $=1111\quad 1010$
Add 1 in complement $+\qquad\quad 1$
Result $=1111\quad 1011$

Which is two's complement of -5.

So addition will be

$+5=\quad 0000\quad 0101$
$-5=\quad 1111\quad 1011$
$+$
Sum $0=\quad 0000\quad 0000,$

since the carry is the ninth bit, it is discarded.

Ex. 3. *Add* -20 *and* $+26$.

Two's complement of $-20=\quad 1110\quad 1100$
$+26=\quad 0001\quad 1010$
$+$
Sum $+6=\quad 0000\quad 0110$

Ex. 4. *Add* -10 *and* $+15$.

Two's complement of $-10 =$ 1111 0110

$+15 =$ 0000 1111

$+$

$+5 =$ 0000 0101

Ex. 5. *Add* $+10$ *and* -15.

Two's complement of $-15 =$ 1111 0001

$+10 =$ 0000 1010

Sum $-5 =$ 1111 1011

Since the two's complement of 1111 1011 are 0000 0101 which is $+5$ and hence

$$1111\ \ 1011 = -5.$$

25·4. BINARY CODES :

A code is simply a system of symbols by means of which meaningful communications can be affected.

A coding method much used is the straight binary code, where all numbers are coded in binary fashion. Hence 1 is 0001 ; 7 is 0111 ; 105 is 1101001. This system has the disadvantage that the large numbers require a great deal of hardware to handle these numbers.

The above difficulty is encountered by using another coding system-called *binary coded decimal* or simply *BCD*. In this system, individual decimal digits are coded in binary notation and are operated upon singly. Thus binary codes representing 0 to 9 decimal digits are allowed. For example 342 in *BCD* is

0011	0100	0010
(3)	(4)	(2)

since 4 bit system is sufficient to handle any decimal digit, only four elements of hardware are needed to operate upon the number because each digit is operated upon at different times. Thus a single four stage register, consisting of four flip-flops can handle any decimal number. This is also called the 8–4–2 –1 code.

Another code used is the *Gray code*. In binary system we find that all four bits are changed when going from 7 to 8 (0111 to 1000), in Gray code only *one bit changes at a time*. Following is the procedure to change a binary number to its gray code equivalent :

(*i*) Record the most significant bit (msb).

(*ii*) Add this bit to the next position, recording the sum and neglecting the carry.

(*iii*) Record successive sums until completed.

Ex. 6. *Convert binary* 1011 *to Gray code.*

	→.→.→	position	8	4	2	1
Binary	1 0 1 1	weightage				
	↓ ↓ ↓ ↓		1	0	1	1
Gray	1 1 1 0					

Procedure :

1. First record the 8's bit which is most significant.
2. Now add 8's bit to 4's bit, $1+0=1$. Record it.
3. Then add the 4's bit to the 2's bit. $0+1=1$. Record it.
4. Then add the 2's bit to the 1's bit, $1+1=10$ (0 with a carry). Record 0 and ignore carry.

TABLE 2. Gray Code

Decimal	Gray
0	0000
1	0001
2	0011
3	0010
4	0110
5	0111
6	0101
7	0100
8	1100
9	1101
10	1111
11	1110
12	1010
13	1011
14	1001
15	1000
16	11000

25·5. BOOLEAN ALGEBRA AND LOGIC GATES :

Boolean algebra is a system of mathematical logic and it differs from both ordinary algebra and the binary number system viz. in Boolean $1+1=1$. In it the variables permitted to have only two values *true* or *false*, usually written as 1 and 0, and the algebraic operations on the variables are limited to those defined as AND, OR, NOT.

Boolean two variables true or false may be represented by *on* or *off* states of electronic switching circuits. On account of such functioning of switching circuits, Boolean algebra becomes of practical significance and is implemented through these switches that act as electronic *logic circuits*.

Thus there are two variables within the Boolean system : 0 and 1. Every number is either *a* 0 or *a* 1. There are no negative or fractional numbers. We write :

If $X=1$ then $X\neq 0$
If $X=0$, then $X\neq 1$

The three operations permitted on these variables are described below :

25·5-1 THE AND OPERATOR :

This operator is represented by the symbol • and the operation can be written as

$$A \text{ AND } B = A \cdot B = AB = C,$$

which means that if A is true **AND** B is true, then C will be true. Under any other condition C will be false. There are four different combinations of A and B that must be considered as shown in table 3. Such a table in which all possible states are examined and a decision made as to whether the output would be true or false is called *truth table*.

TABLE—3. Truth table for AND operation

$A\ B=C$
$0\ 0=0$
$0\cdot 1=0$
$1\cdot 0=0$
$1\cdot 1=1$

AND functions

	0	1	2	3
A	0	0	1	1
B	0	1	0	1
C	0	0	0	1

truth table

We note from truth table that only when both A and B are 1 (or true) will the output, C, be 1 (or true).

AND Gates : Electrical switches can be used to represent AND gate. In fig. 1 (*b*) two switches A and B are connected in series. Lamp glows only when both switch A **AND** switch B are closed.

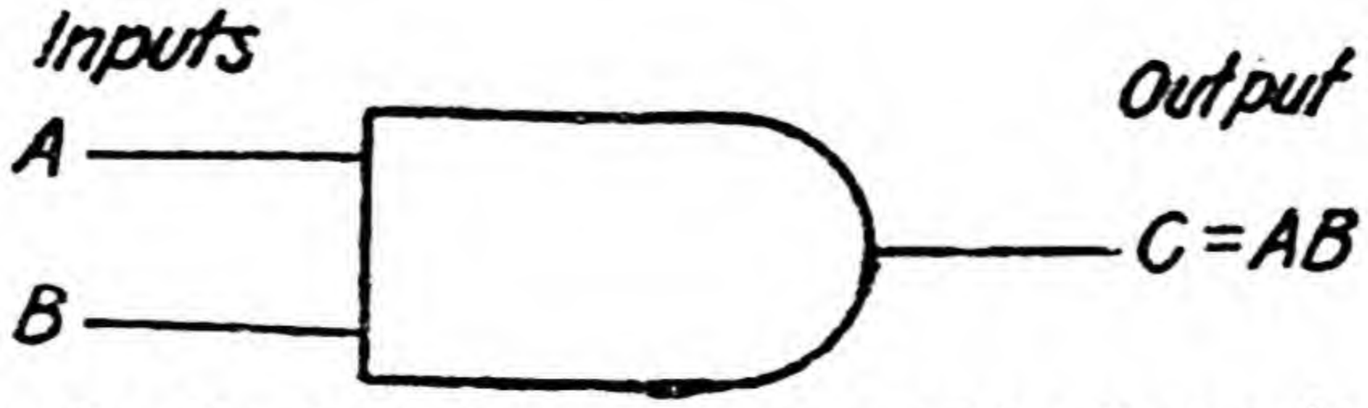
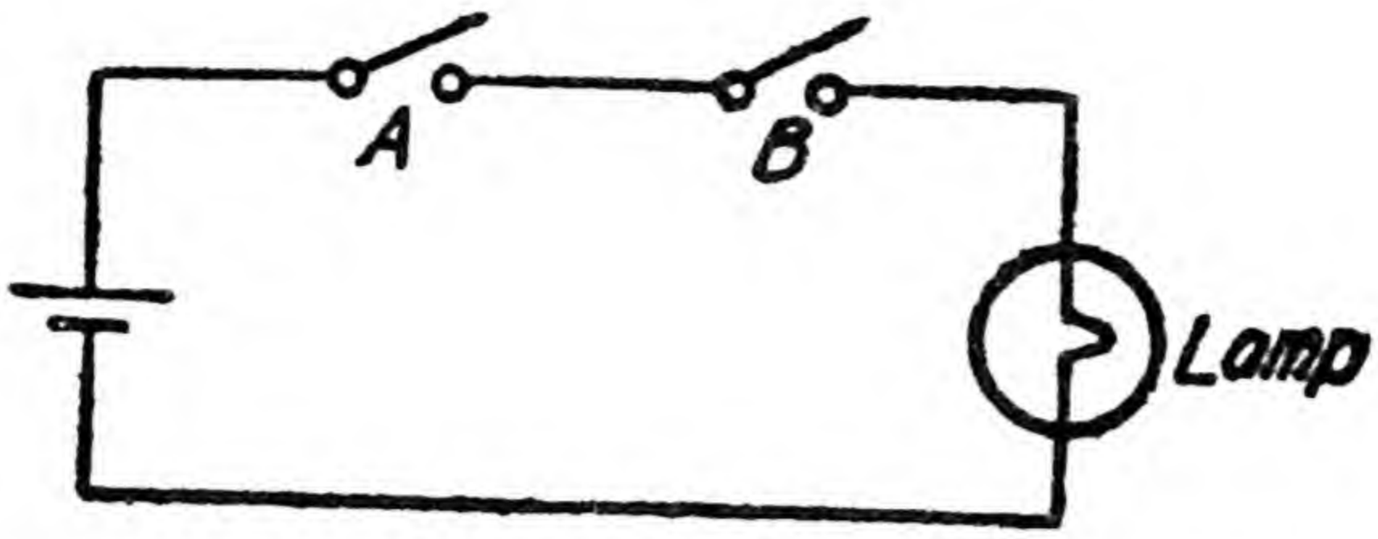


Fig. 1. (*a*) Symbol of AND gate used in electronics.



(*b*) Electrtcal analog o AND gate

Fig. 1. Symbol and electrical analog of AND gate.

Circuit and operation :

Circuit is shown in fig. 1 (*c*). Such a circuit has more than one input circuits and gives an output, as we shall see later, only when the signals are applied to input 1 *and* input 2 *and* input 3 *simultaneously*. On account of this fact, such a circuit is called an **AND** or *coincidence* circuit.

Diodes conduct heavily in the absence of any input signal. Because of low resistance of diode in conduction, high resistance, R_L, and low source resistance, circuit acts like a clamper. The output voltage is thus clamped at approximately ground level.

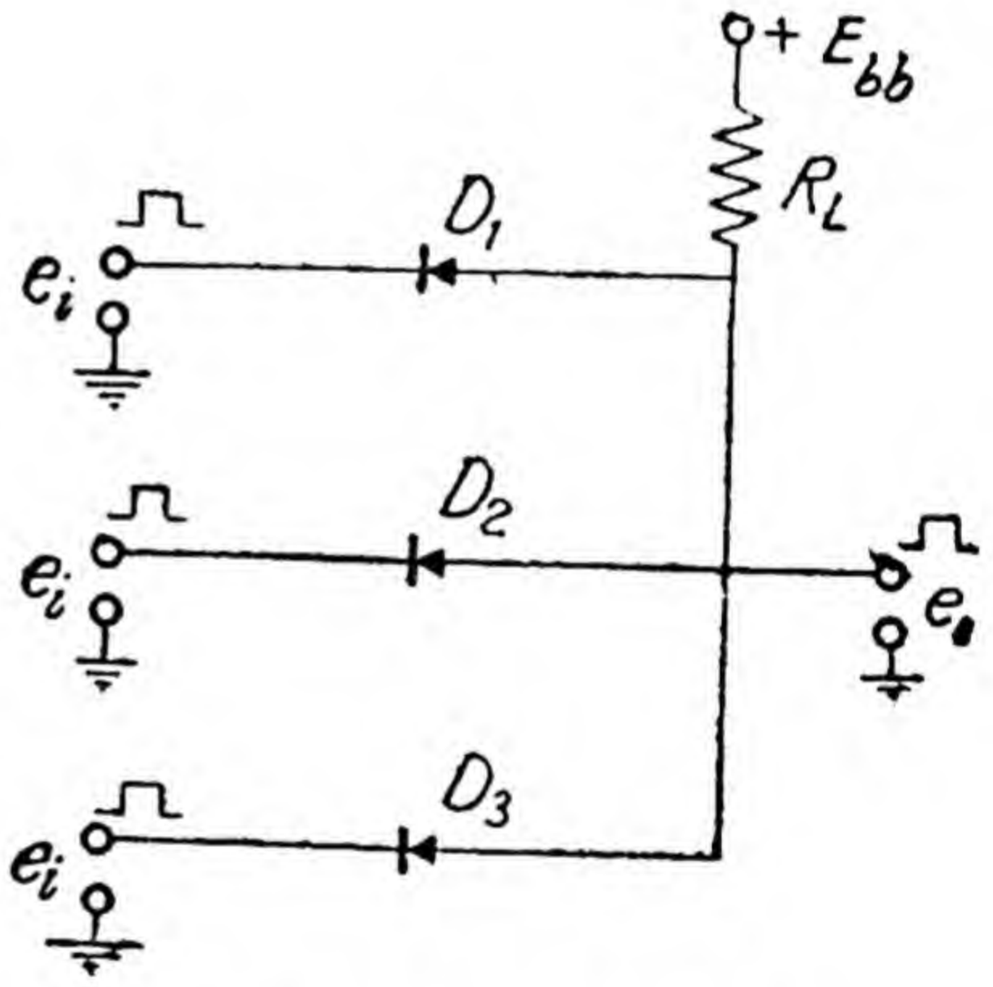


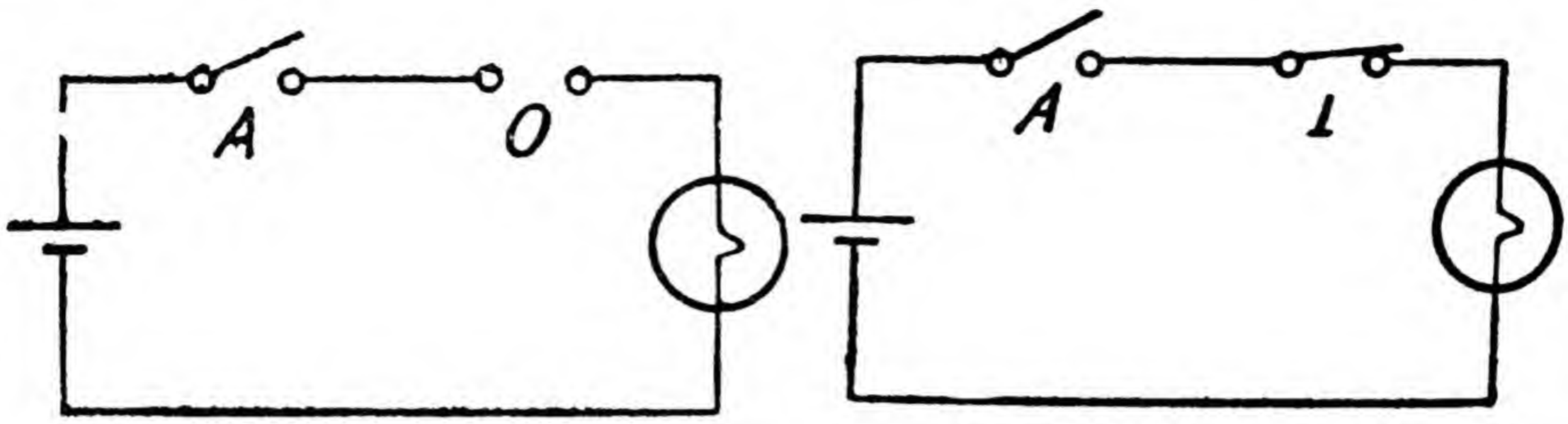
Fig 1. (*c*) Diode AND circuit

Vacuum diodes can also be used.

When a positive going pulse is applied to any one of the input circuits then due to its sufficient amplitude, it will overcome the power supply voltage so as to produce a reverse bias condition. The corresponding diode then does not conduct. Suppose such a pulse is applied to input 1, then diode D_1 will not conduct but because D_2 and D_3 still conduct heavily, they will effectively clamp the output to ground level. Similarly, if the pulses are applied to input 1 and 2 but not to 3, D_3 will remain in conduction and clamp the output at zero level. Therefore an output is possible only when pulses are applied to inputs 1, 2 and 3 *simultaneously*. For further discussion refer to Logic families.

AND laws : The three Boolean algebraic laws can be written as

$$A \text{ AND } 0 = A.0 = 0$$
$$A.1 = A$$
$$A.A = A$$



(*a*) Verifying $A,0=0$ (*b*) Verifying $A.1=A$.

(*c*) Verifying $A\ A=A$.

Fig. 2. Representing AND laws.

Suppose $A=0$ then we get 'off' state in all the three cases :

$$0\cdot0=0$$
$$0\cdot1=0$$
$$0\cdot0=0$$

If $A=1$, then we get 'off' state in one and 'on' state in two cases :

$$1\cdot0=0$$
$$1\cdot1=1$$
$$1\cdot1=1$$

Keeping these values one can easily verify on or off state of the lamp in fig. 2.

25·5-2. THE OR OPERATOR :

This operator is represented by the symbol + and the operation may be written as

$$A \quad \text{OR} \quad B=A+B=C$$

which means that if A is true **OR** B is true, then C will be true. Under any other condition C is false. There are four different combinations that must be considered as shown in table 4. We note from truth table that when either A is 1 or B is 1 (true) the output C will be 1 (true).

TABLE 4. Truth table for OR operation.

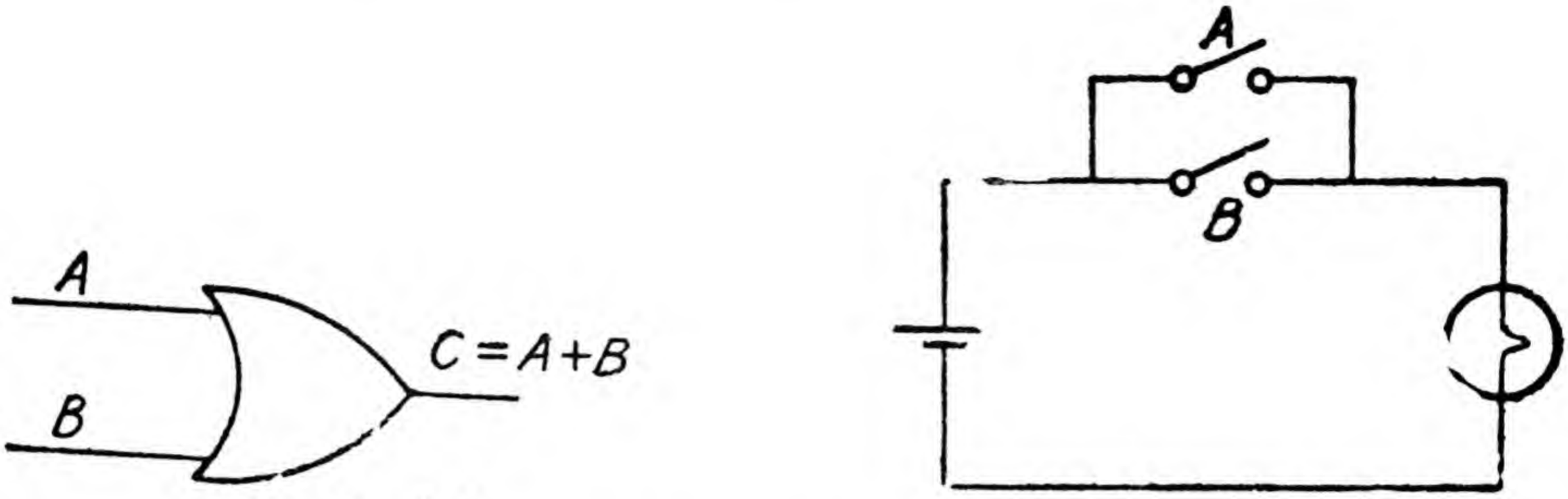
$A+B=C$
$0+0=0$
$0+1=1$
$1+0=1$
$1+1=1$

OR function

	0	1	2	3
A	0	0	1	1
B	0	1	0	1
C	0	1	1	1

Truth table

OR Gates : The symbol for OR gate used in electronics is shown in fig. 3 (*a*) and its electrical analog in fig 3 (*b*).



(a) Symbol (b) Electrical analog

Fig. 3. Symbol and electrical analog of OR gate.

Circuit and operation :

An **OR** circuit is shown in fig. 3 (*c*) with two or more input circuits. Such a circuit produces an output when a signal is applied to *any one* or *all* its input circuits.

In this circuit, if all inputs are at zero level, diodes do not conduct : current through R_L is zero and output level taken across R_L is zero.

When positive going pulse is applied to input 1 *or* input 2 *or* input 3, current flows through R_L developing a positive voltage across it and with R_L being sufficiently large compared to diode forward resistance, output will be equal in magnitude to input pulse Therefere output reaches at maximum amplitude and conduction of rest two diodes will not add any more to the output *i.e.* output voltage is same whether one or all the diodes conduct. For further details, refer to logic families.

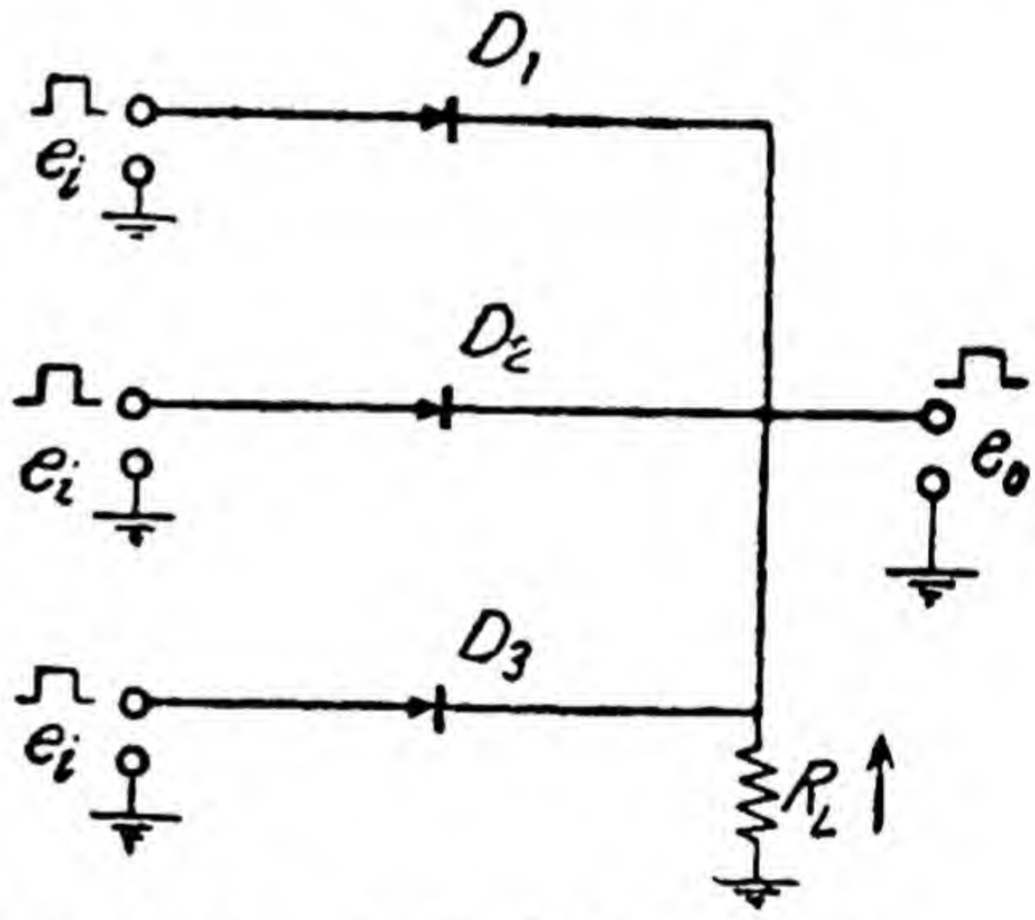
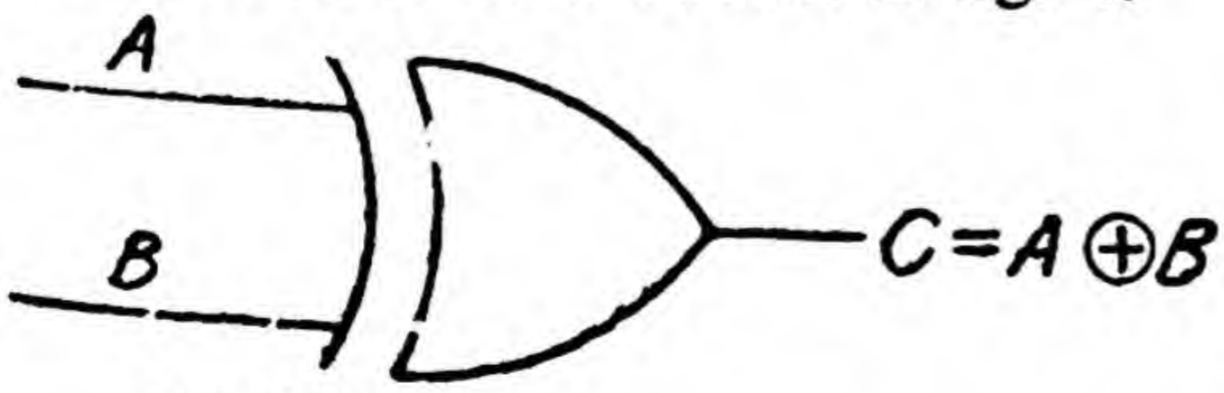


Fig. 3 (*c*). A diode OR circuit.

We note that a positive input pulse produces a positive output pulse which is obviously an *OR* circuit action.

Inclusive OR gate : The above mentioned OR gate in which output is true ($C=1$) when both inputs are true ($A=1$ and $B=1$) is called inclusive OR gate.

Exclusive OR gate : In this case output is true ($C=1$) only when either input ($A=1$ or $B=1$) but *not both*, is true. Symbol and truth table is shown in fig. 4.



symbol

	0	1	2	3
A	0	0	1	1
B	0	1	0	1
C	0	1	1	0

truth cable

Fig. 4. Symbol and truth table of exclusive OR gate

OR laws : We can write three Boolean algebraic laws as follows :

$$A+1=1$$
$$A+0=A$$
$$A+A=A$$

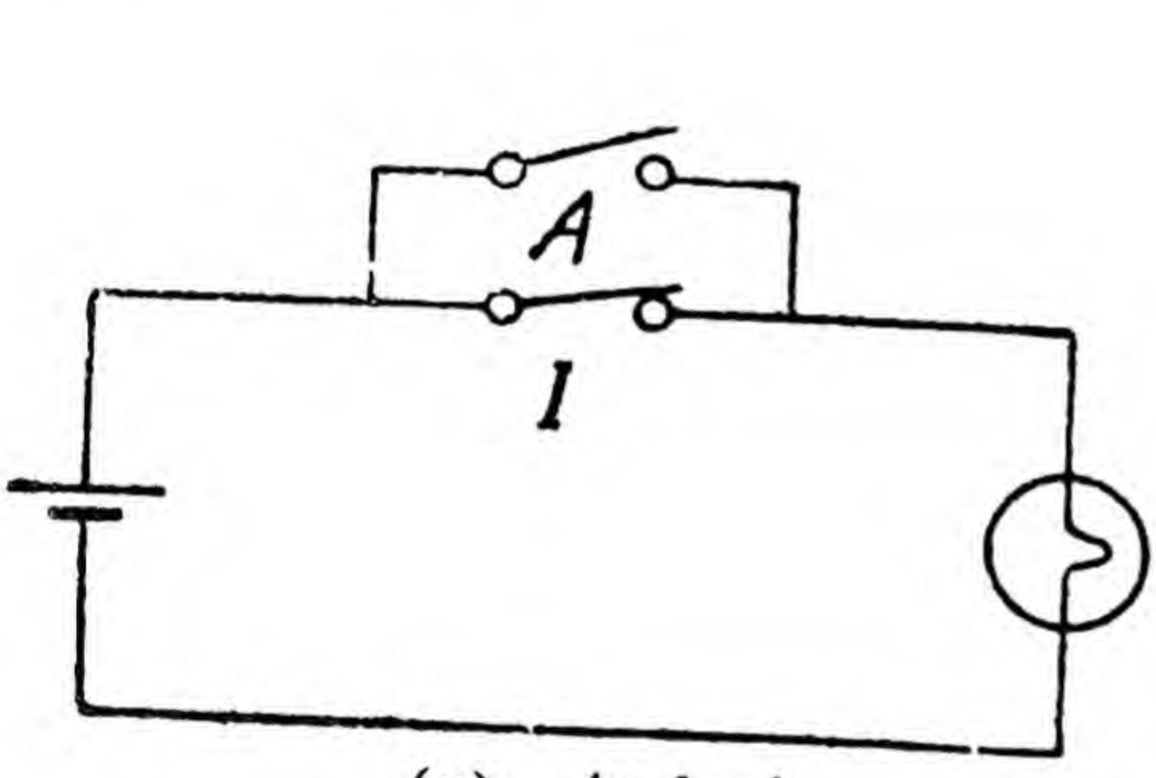
Let $A=0$ then we get

$0+1=1$ on state
$0+0=0$ off state
$0+0=0$ off state

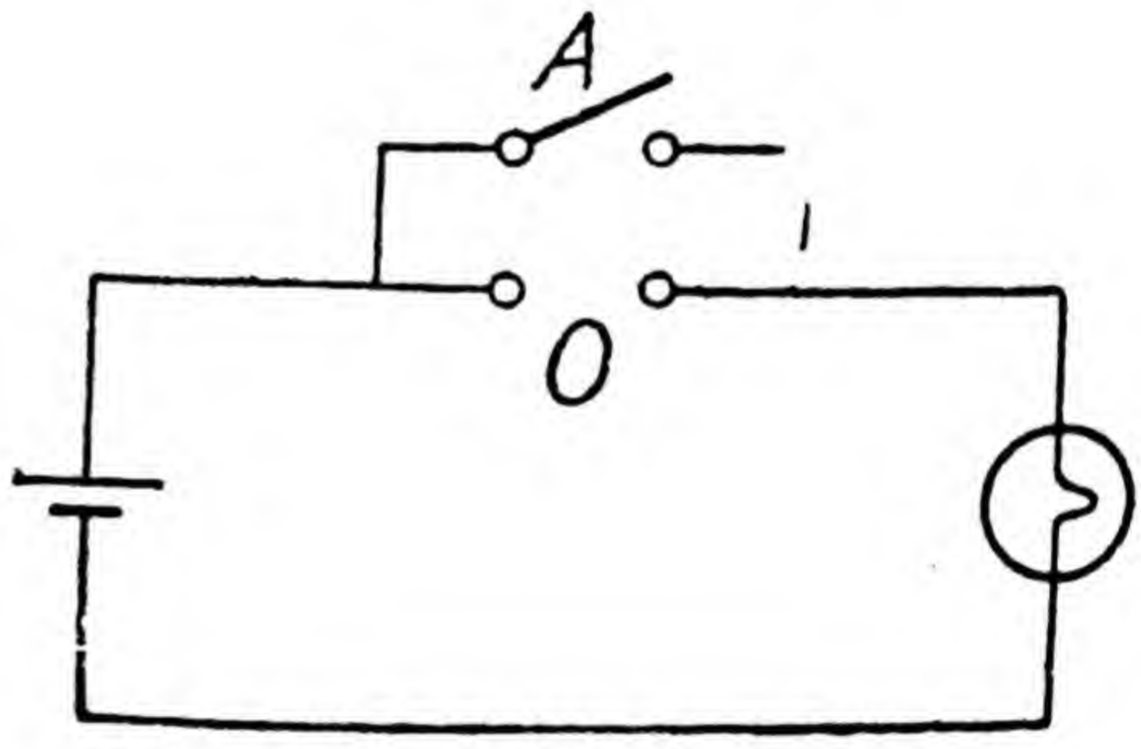
and $A=1$, then

$1+1=1$ on state
$1+0=1$ on state
$1+1=1$ on state

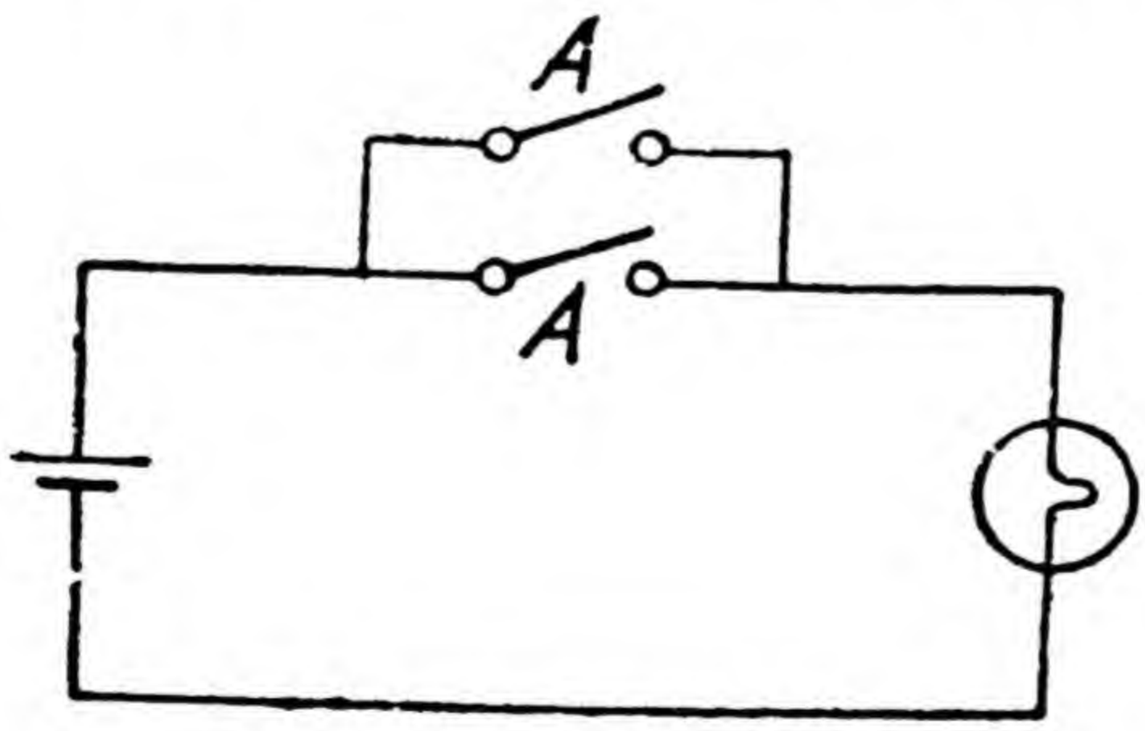
These states can be easily tested from circuits given in fig. 5. In 'on' state the lamp will glow. 1 means connected and 0 means disconnected.



(*a*) $A+1=1$



(*b*) $A+0=A$



(*c*) $A+A=A$

Fig. 5. Verifying OR laws.

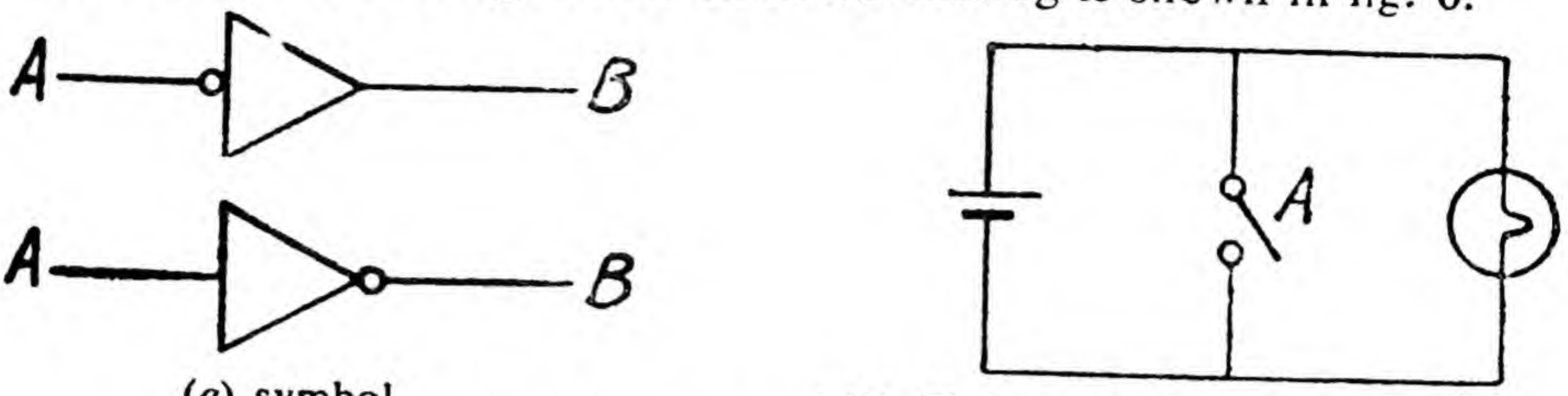
25·5·3. THE NOT OPERATION :

This operator is represented by a bar *e.g.*, $\bar{A}$. The operation is also called as complementation or negation (inversion). The NOT gate is thus called complementary circuit or inverter. It has single input and single output. It inverts or complements the input. Thus if input is in 1 state, output will be in 0 state or *vice-versa*.

The truth table is

	0	1
A	0	1
B	1	0

and the symbol alongwith its electrical analog is shown in fig. 6.

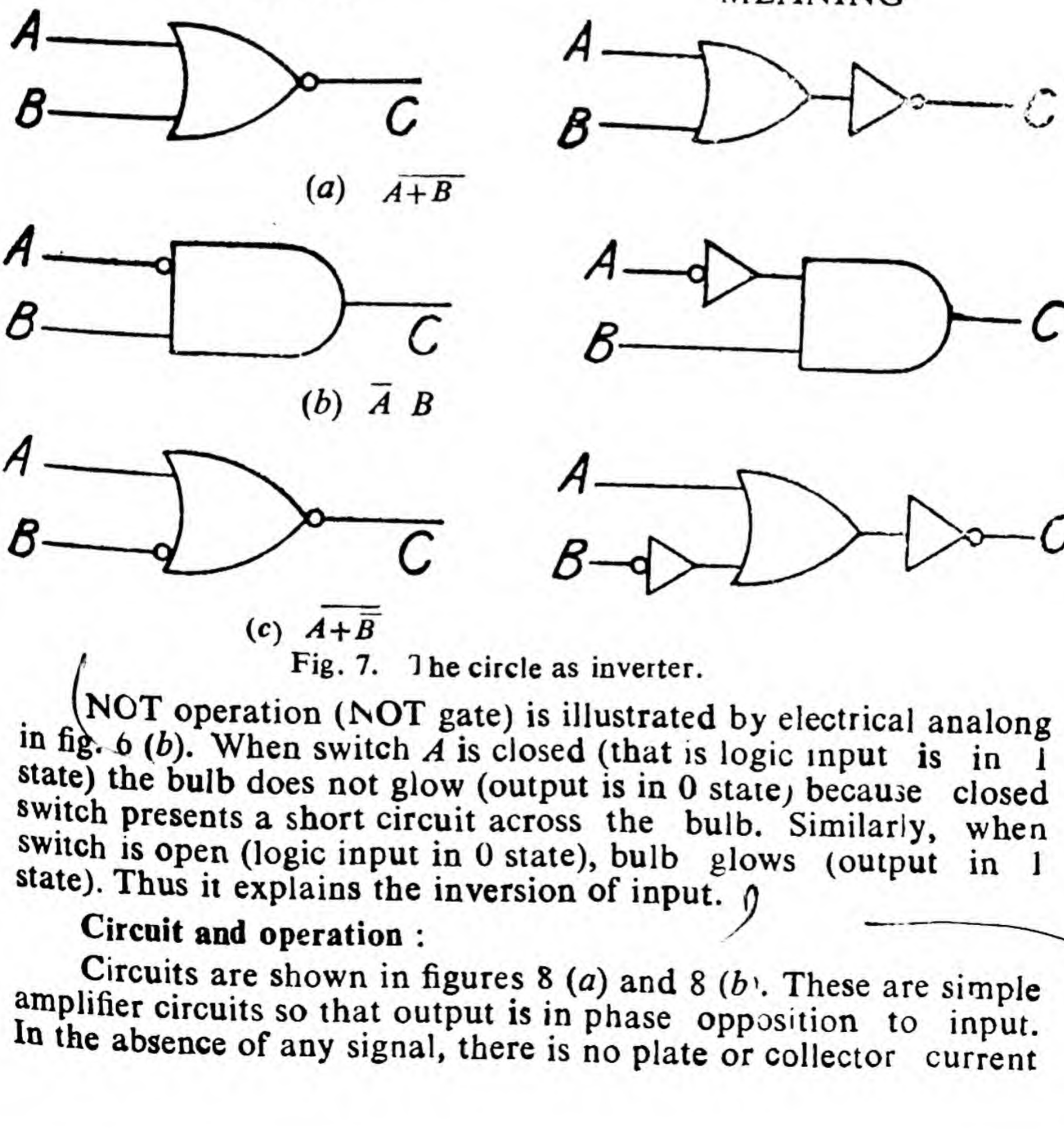


(*a*) symbol (*b*) Electrical analog

Fig. 6. The inverter.

In symbol the *circle* actually represents the *inversion* and the triangle an amplifier. Some examples of this use of circle (of inversion) are shown in fig. 7 alongwith their meaning.

SYMBOL MEANING



(*a*) $\overline{A+B}$

(*b*) $\bar{A}\ B$

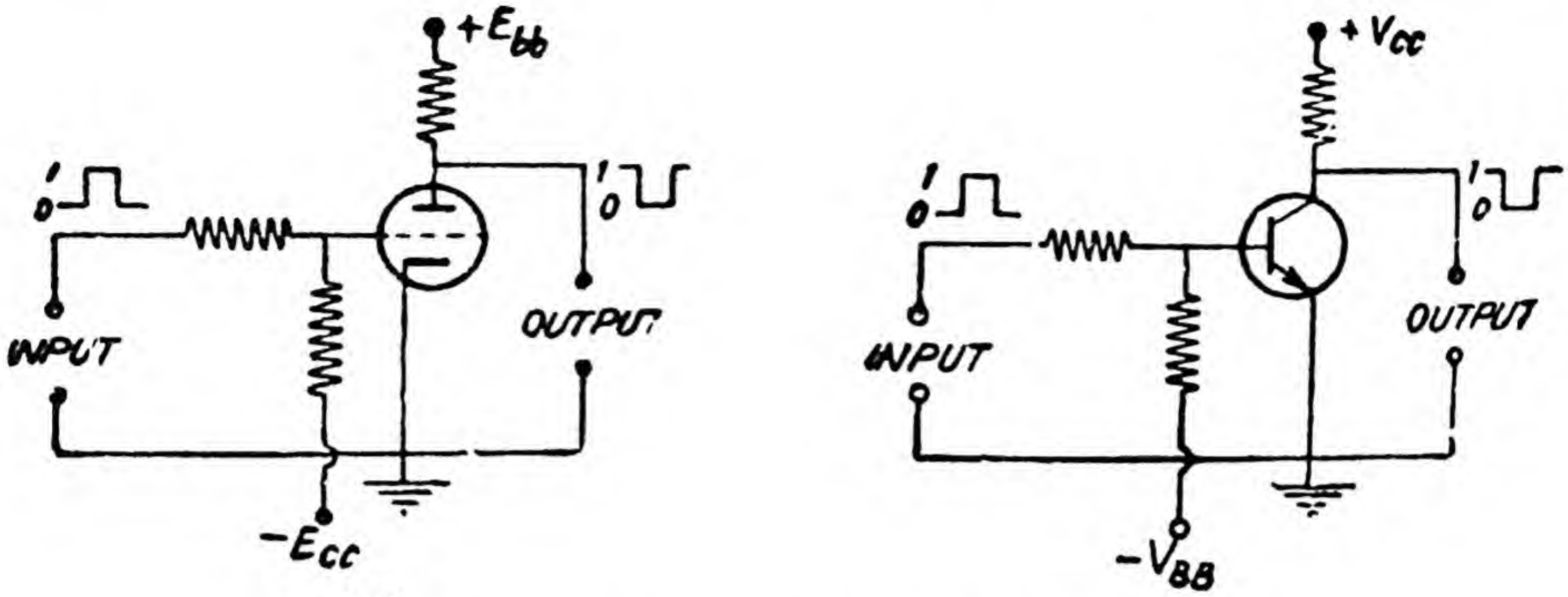
(*c*) $\overline{A+\bar{B}}$

Fig. 7. The circle as inverter.

NOT operation (NOT gate) is illustrated by electrical analong in fig. 6 (*b*). When switch *A* is closed (that is logic input is in 1 state) the bulb does not glow (output is in 0 state) because closed switch presents a short circuit across the bulb. Similarly, when switch is open (logic input in 0 state), bulb glows (output in 1 state). Thus it explains the inversion of input.

Circuit and operation :

Circuits are shown in figures 8 (*a*) and 8 (*b*). These are simple amplifier circuits so that output is in phase opposition to input. In the absence of any signal, there is no plate or collector current



(a) (b)

Fig. 8. The invertor or NOT Circuit.

because tube or transistor is biased beyond cut-off and output reaches the full supply voltage [(stage represented by 1 in fig. 8 (*a*) (*b*) (*c*)]. When a positive going square wave is applied at the input, tube or transistor conducts so that plate of collector potential drops to low voltage level characterised by the symbol 0. Now there is no output.

We thus find that output is present only when input signal is *not* applied. Therefore these are called *NOT* circuits. For details, see Logic families.

NOT laws : Few Boolean algebraic laws can be written as

$$\bar{0}=1$$

$$\bar{1}=0$$

If $A=0$, then $\bar{\bar{A}}=1$

If $A=1$, then $\bar{\bar{A}}=0$.

These laws are already explained by the electrical analog which represents that if switch is closed (1 state) lamp will not glow (0 state) or vice versa.

A fourth law can be written following the definition of an inverter and is

$$\bar{\bar{A}}=A$$

and can be verified by examining fig. 8 (*c*). Obviously, A inverted twice, $\bar{\bar{A}}$, is identical to A.

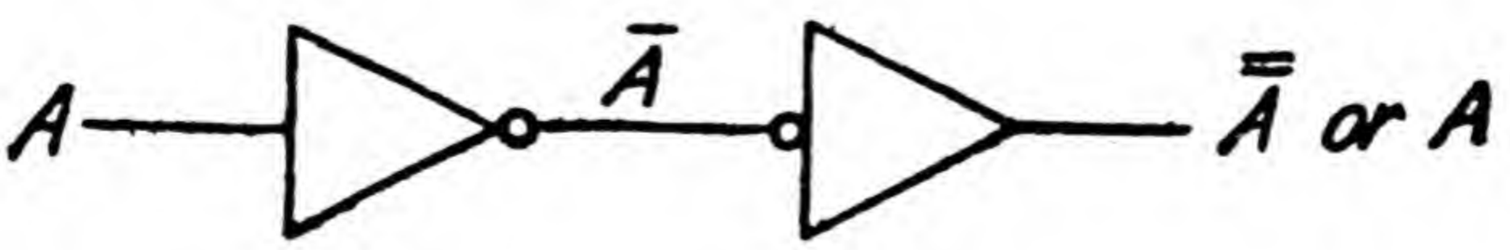


Fig. 8. (*c*) Verifying $\bar{\bar{A}}=A$.

25·6. UNIVERSAL BUILDING BLOCKS :

In previous art. we have discussed AND, OR and NOT gates. There are some units which can perform all the aforesaid three

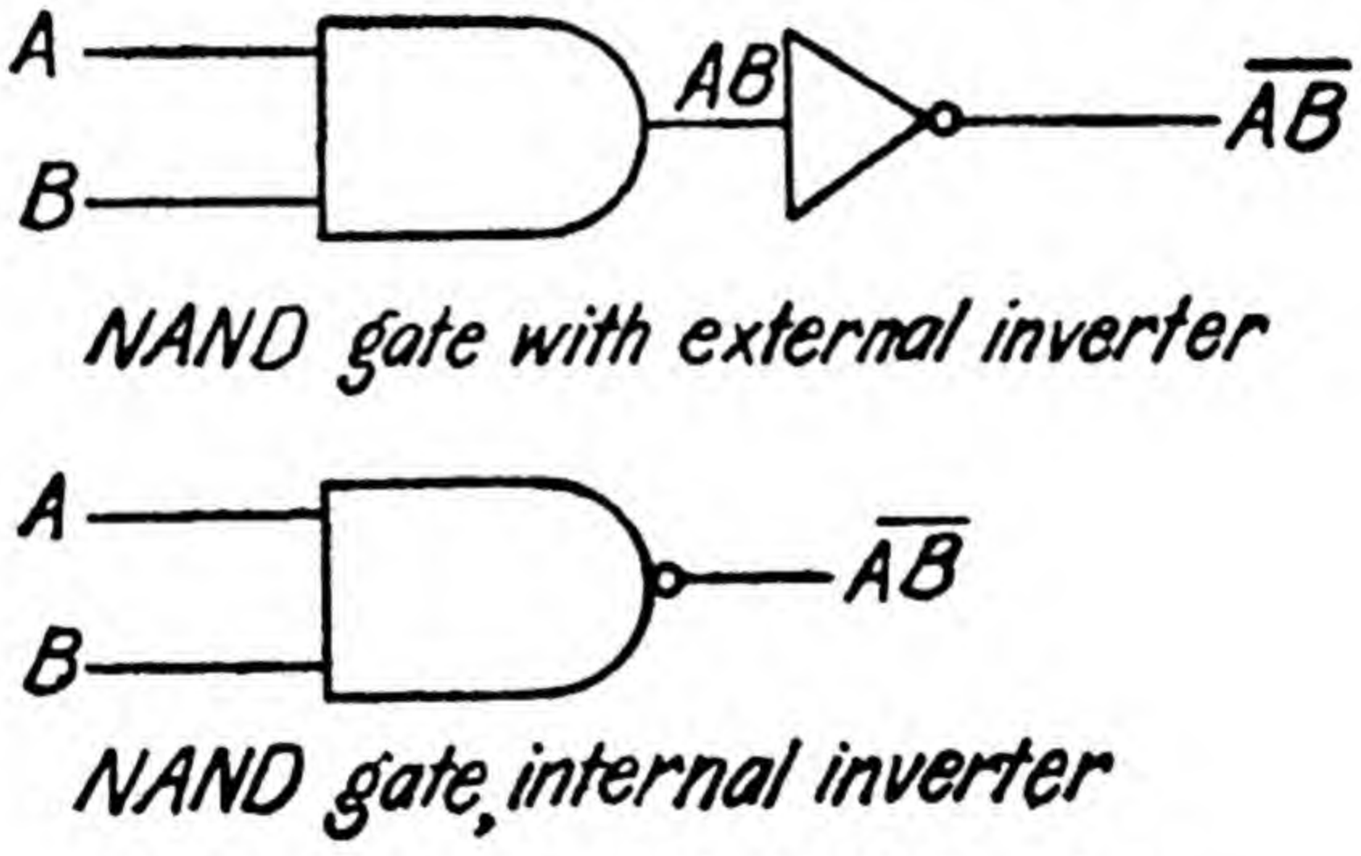
logic functions *e.g.* a NAND gate; a NOR gate can perform both the OR function and the AND function. Thus these two gates are called universal building blocks.

NAND gate : When AND gate is combined with NOT gate in cascade the resultant gate is known as NAND gate. Fig. 9 (*a*) shows a NAND gate symbolically in which a NOT gate follows an AND gate (inverter). The inputs A and B are ANDed, yielding AB and then inverted, giving $\overline{AB}$. The truth table is

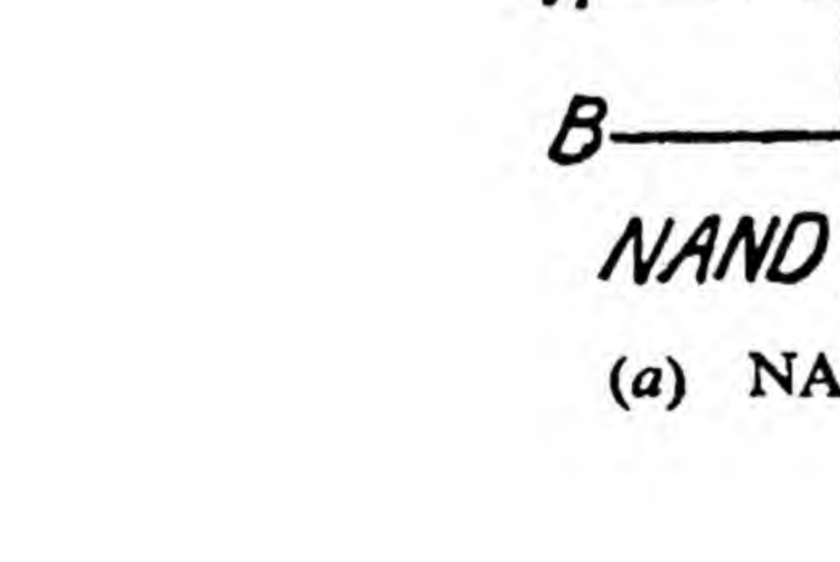
	0	1	2	3
A	0	0	1	1
B	0	1	0	1
AB	0	0	0	1
output, C, $=\overline{AB}$	1	1	1	0

Fig. 9 (*b*) shows the electrical analog of this gate. Obviously the lamp will not glow if both inputs A and B are in logic '1' state *i.e.* switches are closed.

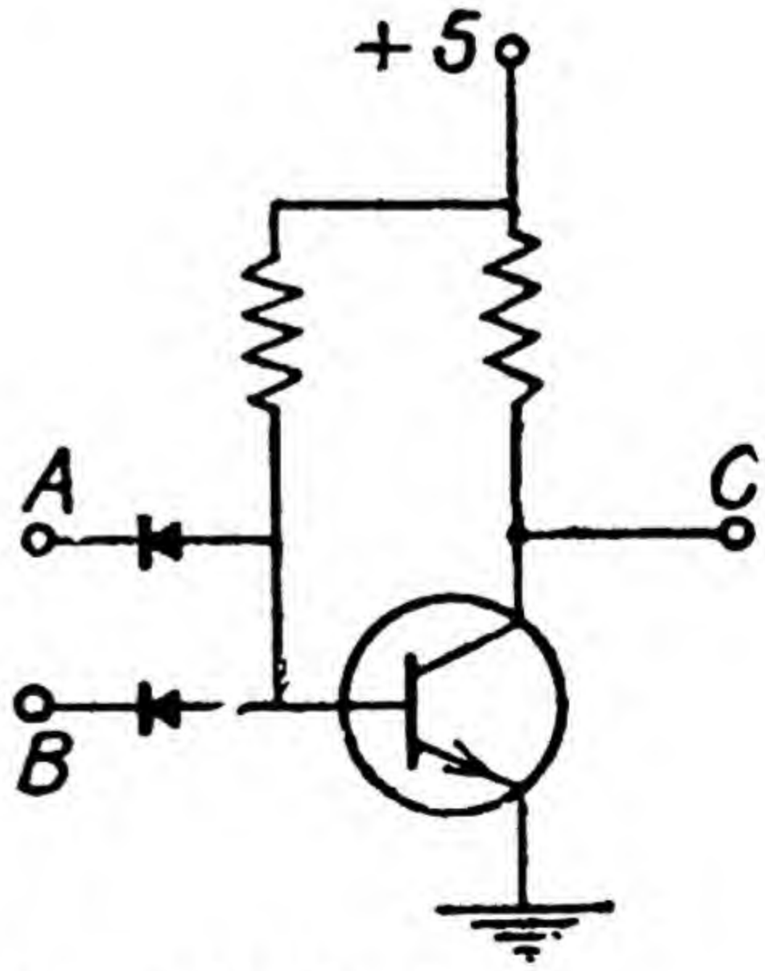
Fig. 9 (*c*) indicates a NAND gate circuit. The logic state 1's represents $+5V$ and state 0 is ground. The output, C, is low only



(*a*) NAND gate used AND function.



(*b*) electrical analog.



(*c*) NAND circuit.

Fig. 9. NAND gate.

when both inputs A and B are at 1 logic state. Thus when A is a1, and B is a1, the output C is a 0.

Another way of looking at the NAND circuit is to observe that its output is a 1 if A is 0 or if B is 0. Thus it can perform the OR function. Fig. 9 (d) depicts this function. First A is inverted giving $\bar{A}$, B is inverted giving $\bar{B}$. Then OR gate forms $\bar{A}+\bar{B}$ at its output. Truth table is

	0	1	2	3
A	0	0	1	1
B	0	1	0	1
$\bar{A}$	1	1	0	0
$\bar{B}$	1	0	1	0
output, $C, =\bar{A}+\bar{B}$	1	1	1	0

While discussing De Morgan's theorem in art. 25·7 we shall conclude that logic circuits of fig. 9 (a) and fig. 9 (b) are equivalent.

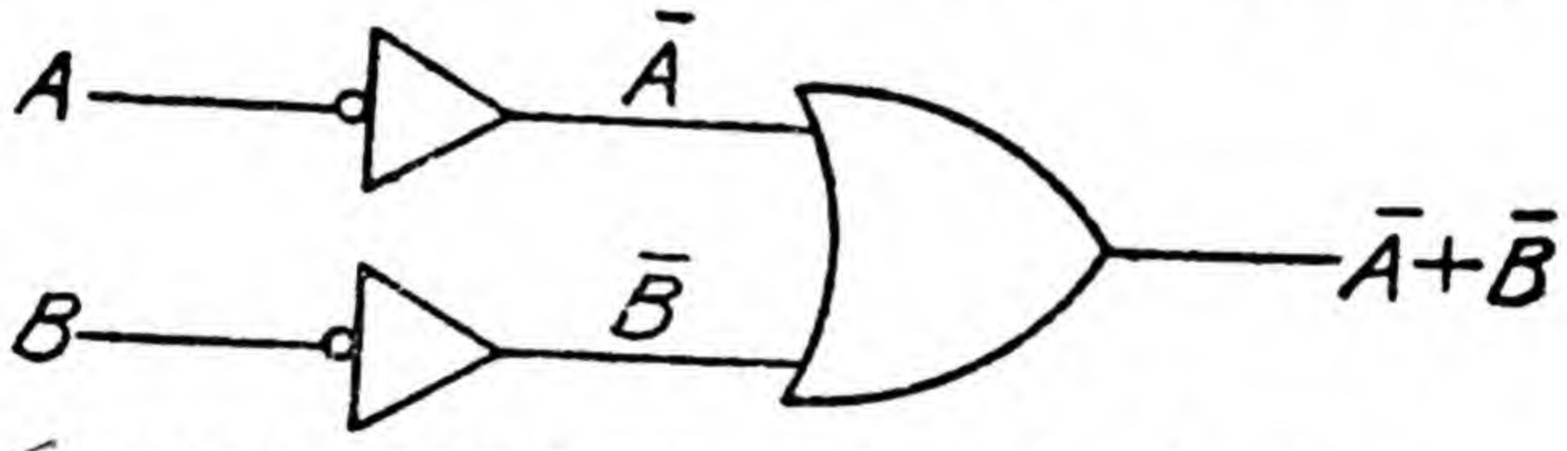


Fig. 9 (d) NAND gate used as OR function.

NOR gate : When a NOT gate is combined with OR gate in cascade, the resultant gate is called as NOR gate. It is represented symbolically in fig. 10 (a), in which a NOT gate follows an OR gate. The logic equation is

$$\overline{A+B}=C$$

and the truth table is

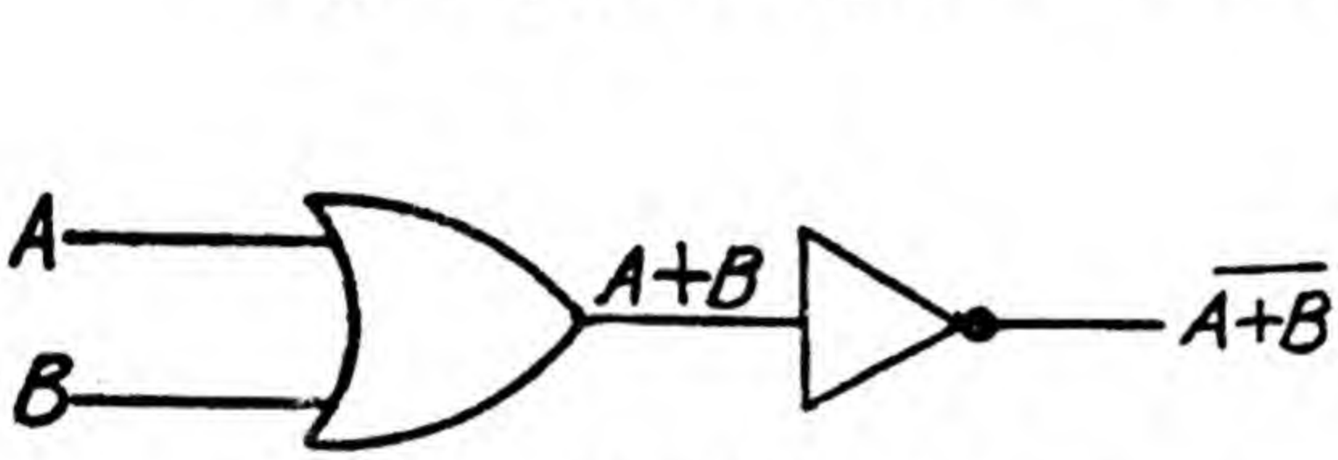
	0	1	2	3
A	0	0	1	1
B	0	1	0	1
$A+B$	0	1	1	1
output, $C=\overline{A+B}$	1	0	0	0

'Not $A+B$' or $\overline{A+B}$ is interpreted as "If all inputs are at logic 0, the output C will be at logic 1. If any input is logic 1, the output C will be logic 0". NOR gate electrical analog is shown in fig. 10 (b), which predicts that if any one of the switches is in state 1 *i.e.* closed, the lamp will not glow (0 state). Circuit for gate is given in fig. 10 (c). If logic 1 appears at A OR B, the OR gate output itself will be at logic 1 which will be inverted to logic 0 at C. Hence, to have a logic 1 at C, there must be logic 0 at the inverter

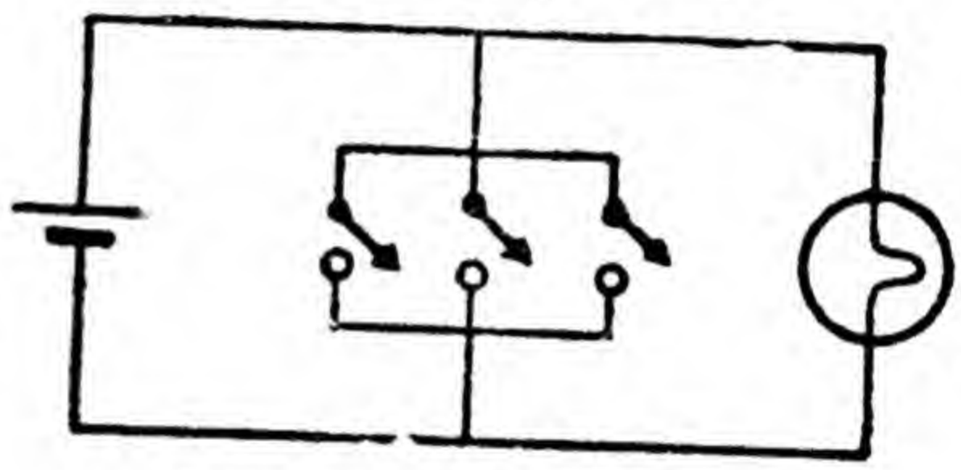
input, which requires that A AND B be at logic 0. (Here we note that the NOR gate has similar characteristics to that of the AND function. The truth table is

	0	1	2	3
A	0	0	1	1
B	0	1	0	1
$\bar{A}$	1	1	0	0
$\bar{B}$	1	0	1	0
output, $C=\bar{A}\bar{B}$	1	0	0	0

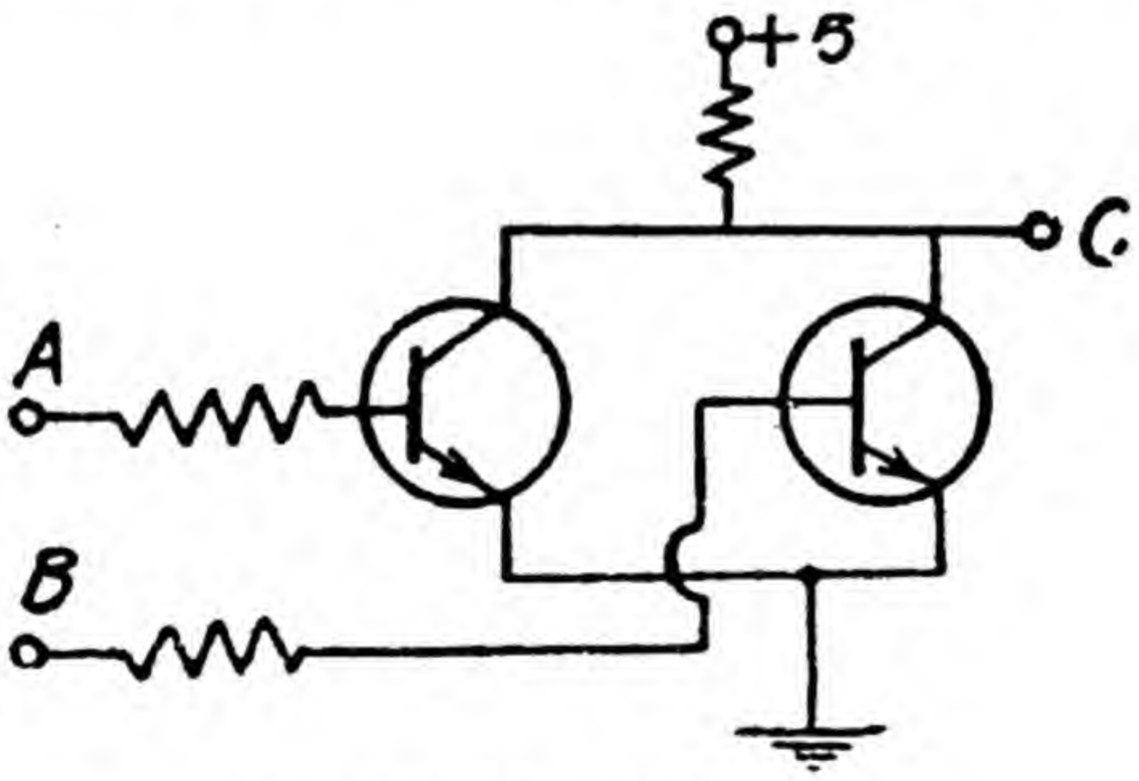
In fig. 10 (*d*) NOR gate used as AND function is depicted.



(*a*) NOR gate used as OR function.



(*b*) Electrical analog.



(*c*) NOR gate circuit.

Fig. 10. NOR gate representation.

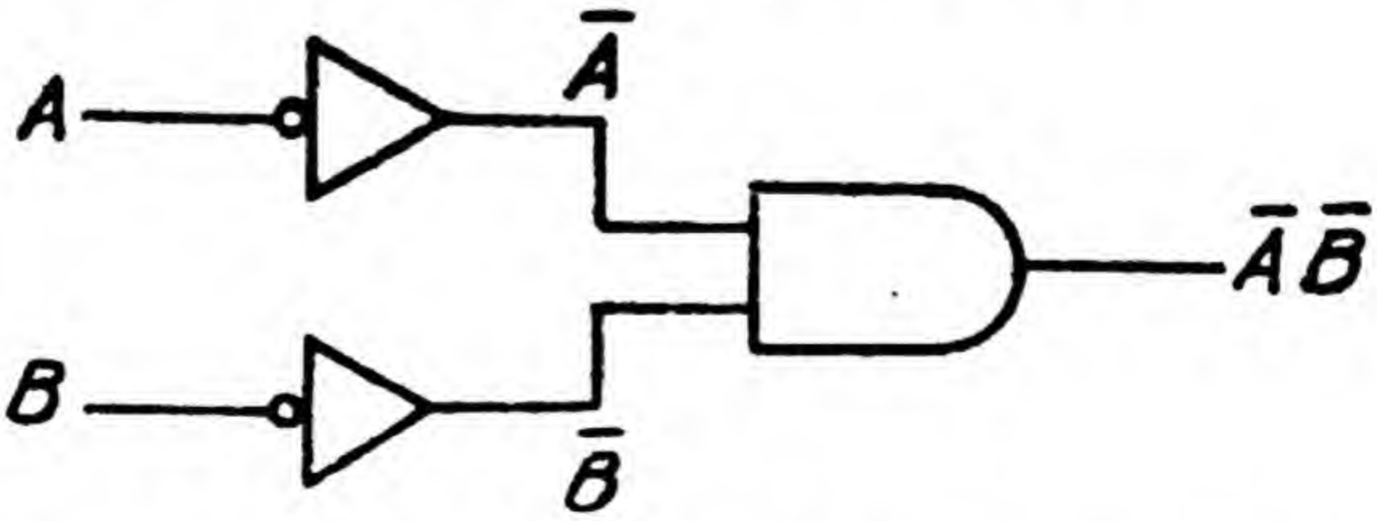


Fig. 10. (*d*) NOR gate used as AND function.

In art, 25·7 while discussing De Morgan's theorem, it will be concluded that logic circuits for NOR gate sketched in fig. 10 (*a*) and that in fig. 10 (*d*) are equivalent.

25·7. LAWS OF BOOLEAN ALGEBRA AND THEIR CONVERSION INTO A LOGIC DIAGRAM :

Though some of the laws are stated in art. 25·5, we shall take up them here again.

Laws of Complementation (NOT laws) :

Law 1 $\bar{0}=1$

Law 2 $\bar{1}=0$

Law 3 If $A=0$, then $\bar{A}=1$

Law 4 If $A=1$, then $\bar{A}=0$

Law 5 $\bar{\bar{A}}=A$.

AND Laws : The four AND laws are

Law 6 $A.0=0$

Law 7 $A.1=A$

Law 8 $A.A=A$

Law 9 $A.\bar{A}=0$.

The last law represents that if at one input of logic is 1, then other input $\bar{A}$ would be 0 so that output is 0 (fig. 11). Also if $A=0$ $\bar{A}=1$ even output will be 0.

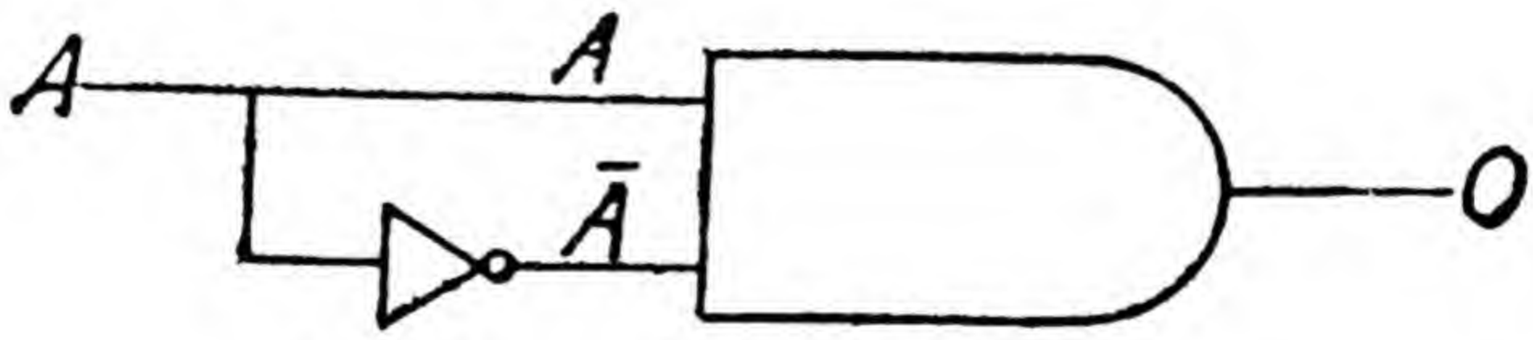


Fig. 11. Verifying $A.A=0$.

OR Laws : The four OR laws are

Law 10 $A+0=A$

Law 11 $A+1=1$

Law 12 $A+A=A$

Law 13 $A+\bar{A}=1$.

The last law can be easily verified by taking $A=1$, then $\bar{A}=0$, so that $1+0=1$.

Commutative Laws :

Law 14 $A+B=B+A$

Law 15 $A\,B=B.A$.

Law 14 predicts that adding A and B one arrives at the same rerult as by adding B and A Similarly Law 15 predicts that

multiplying A and B one gets the same result as by multiplying B and A. It can be seen as follows :

	0	1	2	3
A	0	0	1	1
B	0	1	0	1
$A+B$	0	1	1	1
$B+A$	0	1	1	1

	0	1	2	3
A	0	0	1	1
B	0	1	0	1
$A.B$	0	0	0	1
$B.A$	0	0	0	1

Thus in general, we may state that the logical element may appear in any order, the result is the same.

Associative Law : The two associative laws are

Law 16 $A+(B+C)=(A+B)+C$

Law 17 $A\cdot(B\cdot C)=(A\cdot B)\cdot C$

This law states that the logical elements may be grouped in any combination provided they are connected by the same sign. Law 16 is depicted in fig. 12 using OR gate and Law 17 in fig. 13 using AND gate.

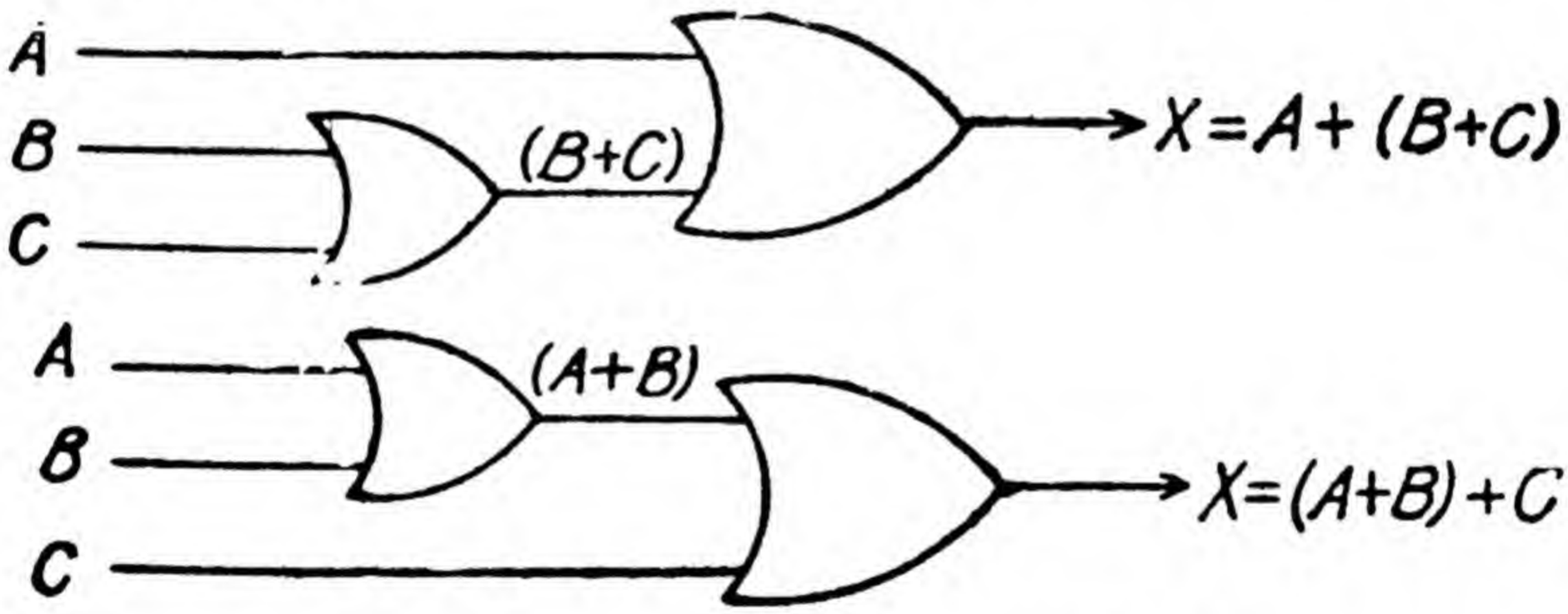


Fig. 12. Illustrating the associative law 16.

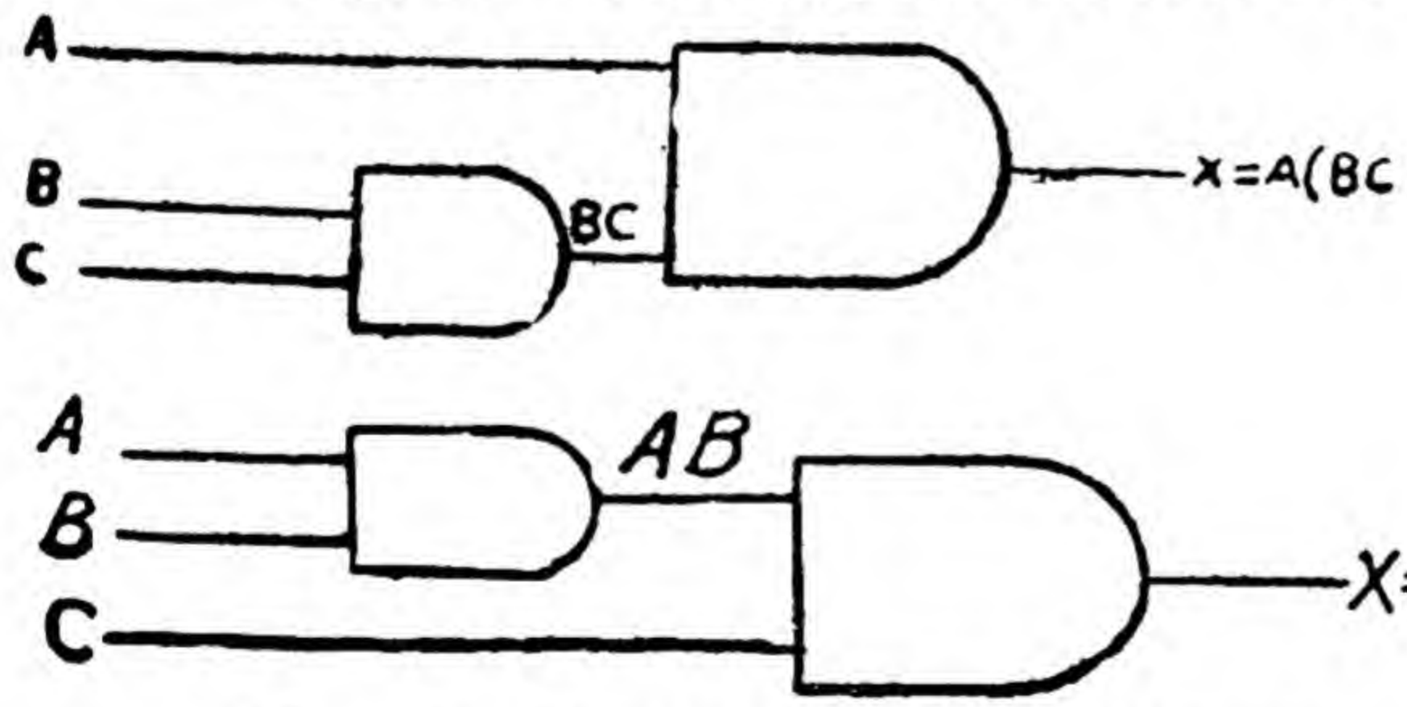


Fig. 13. Illustrating the associative law 17.

Distributive Laws : It states that the expressions can be expanded by multiplying term by term as in ordinary algebra. Thus

Law 18 $A\cdot(B+C)=A\cdot B)+(A\cdot C)$

Similarly, with the help of distributive laws, the terms can be factored *e.g.*

$$A\cdot B+A\cdot C=A\cdot(B+C).$$

We can write other laws as follows :

Law 19 $A+(B\cdot C)=(A+B)\cdot(A+C)$

Law 20 $A+(\overline{A}\cdot B)=A+B$

We can proceed to prove Law 19 in the following way

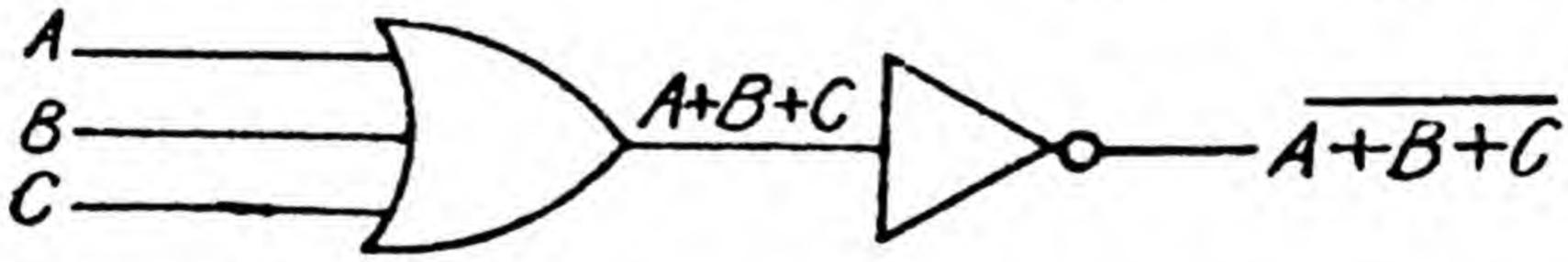
$$A+BC=A\div BC$$
$$=A\cdot 1+BC \quad \text{Law 7}$$
$$=A\ (1+B)+BC \quad \text{Law 11 and 14}$$
$$=A+AB+BC \quad \text{Law 18}$$
$$=A\ (1+C)+AB+BC \quad \text{Law 11}$$
$$=AA+AC+AB+BC \quad \text{Law 8 and 18}$$
$$=A(A+C)+BA+BC \quad \text{Law 18 and 15}$$
$$=A(A+C)+B(A+C) \quad \text{Law 18}$$
$$=(A+C)\ A+(A+C)\ B \quad \text{Law 15}$$
$$=(A+C)\ (A+B) \quad \text{Law 18}$$
$$=(A+B)\ \ A+C) \quad \text{Law 15.}$$

De Morgan's Theorems : These theorems are used in changing Boolean expressions to equivalent forms.

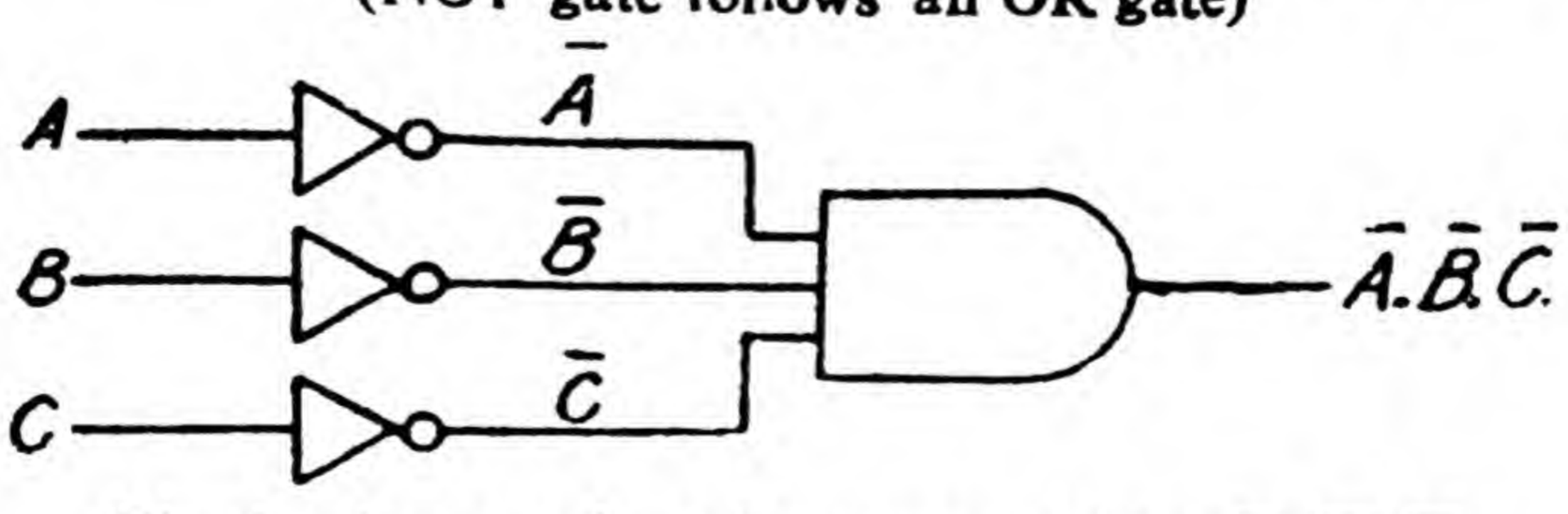
First Theorem : It states that the complement of a sum equals the products *i.e.*

$$\overline{A+B+C+\ ..\ +N}=\overline{A}.\ \overline{B}\ \ \overline{C}....\overline{N}$$

Left hand side is implemented in fig. 14 (*a*) while right hand side is implemented in fig. 14(*b*). With the help of above equality it can be said that these two logic circuits are equivalent or in other words the circuits are interchangeable in regard to logic functions.



(*a*) Implementation of logic expression $\overline{A+B+C}+...$
(NOT gate follows an OR gate)



(*b*) Implementation of logic expression $\overline{A}.\ \overline{B}\ \ \overline{C}.$

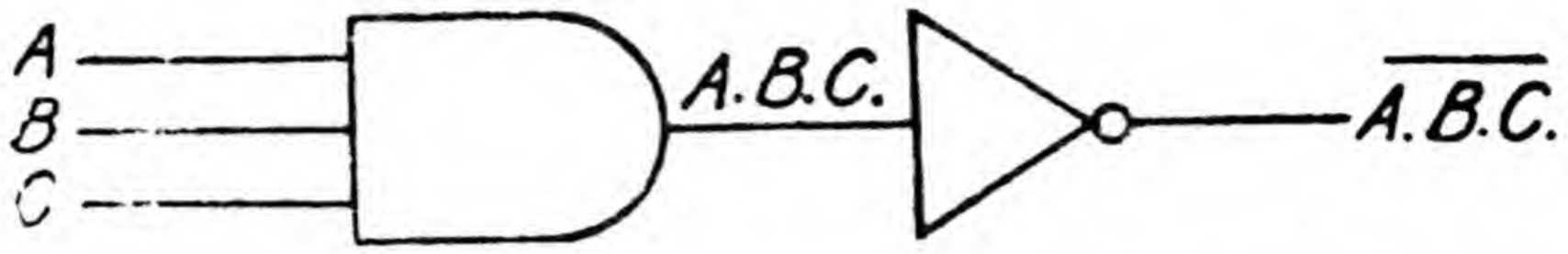
Fig. 14. Equivalent logic circuits (De Morgan's first theorem).

Second De Morgan's Theorem : It states that

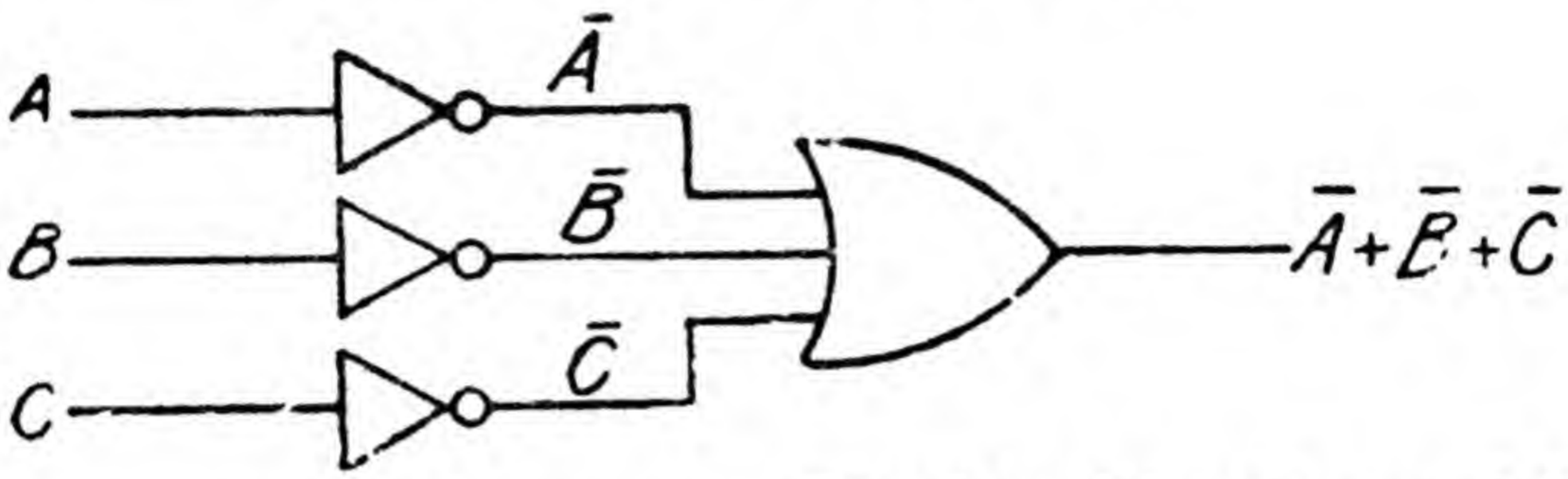
$$\overline{A.B.C.....N}=\overline{A}+\overline{B}+...+\overline{N}$$

Both sides of this equation have been separately implemented through NAND gate (fig. 15*a*) and NOR gate (fig. 15*b*) respectively.

By above equality it is inferred that the two logic circuits (fig. 15*a* and 15*b*) are equivalent.



(*a*) Implementation of logic expression $\overline{A.B.C}$...(NOT gate follows an AND gate *i e.* it is a NAND gate)



(*b*) Implementation of logic expression $\overline{A} + \overline{B} + \overline{C}$ (an OR gate follows a NOT gate *i.e.* it is a NOR gate)

Fig. 15. Equivalent logic circuits (second De Morgan's theorem)

Procedure to apply De Morgan's theorem : Following steps are used to determine equivalent forms using the De Morgan's theorem :

(*i*) complement all variables including output variables.

(*ii*) change all AND to OR gates.

(*iii*) change all OR to AND gates.

(*iv*) complement both sides of equations to obtain the original resultant variables.

The procedure is called *demorganization.*

Ex. 1. *Demorganize the function* $\overline{A.B}$

Given $\overline{AB}$

Step 1. Complement all variables $\overline{\overline{A}.\overline{B}}$

Step 2. Change *AND* gate to OR gate $\overline{\overline{A}+\overline{B}}$

Step 3. No change

Step 4. Complement variables $\overline{A}+\overline{B}$

Thus we get

$$\overline{A.B}=\overline{A}+\overline{B}$$

Ex. 2. *Demorganize the function* $\overline{A\ \overline{B}+C}$

Given $\overline{A\ \overline{B}+C}$

Complement function $A\ \overline{B}+C$

Change operators $(A+\overline{B})\ (C)$

Complement variables $(\overline{A}+B)\ (\overline{C})$

25·8 REDUCING BOOLEAN EXPRESSIONS :

In order to require lesser elements of hardware and therefore for reduction in cost, every Boolean equation should be reduced to as simple a form as possible. The simplified form will have small number of operators and therefore requiring smaller number of elements of hardware. The general procedure is to use the 21 laws discussed in art. 25·7. Some help can be sought from the following points :

(*i*) Multiply all variables necessary to remove parentheses.

(*ii*) Look for identical terms and using Law 12 drop one of them.

(*iii*) Look for a variable and its negation in the same term. This term can be dropped. For example

$$B\bar{B}C = 0.C \quad \text{Law 9}$$
$$= 0 \quad \text{Law 6}$$

(*iv*) Look for pairs of terms that are identical except for one variable. If the one variable is missing, the larger term can be dropped. For example

$$ABCD + ABC = ABC\,(D+1) \quad \text{Law 18}$$
$$= ABC.1 \quad \text{Law 11}$$
$$= ABC \quad \text{Law 7}$$

or
$$ABCD + AB\,\bar{C}\,D = ABC\,(C+\bar{C}) \quad \text{Law 18}$$
$$= ABD.1 \quad \text{Law 13}$$
$$= ABD \quad \text{Law 7}$$

Ex. 1. *Reduce the expression* $\overline{\overline{AB} + \bar{A} + AB}$

Demorganize $\overline{AB}$	$\overline{\bar{A} + \bar{B} + \bar{A} + AB}$	Law 21
Reduce	$\overline{\bar{A} + \bar{B} + AB}$	Law 12
Reduce	$\overline{\bar{A} + \bar{B} + A}$	Law 20
Rearrange	$\overline{A + \bar{A} + \bar{B}}$	Law 14
Reduce	$\overline{1 + \bar{B}}$	Law 13
Reduce	$\bar{1}$	Law 11
Convert	0	Law 2

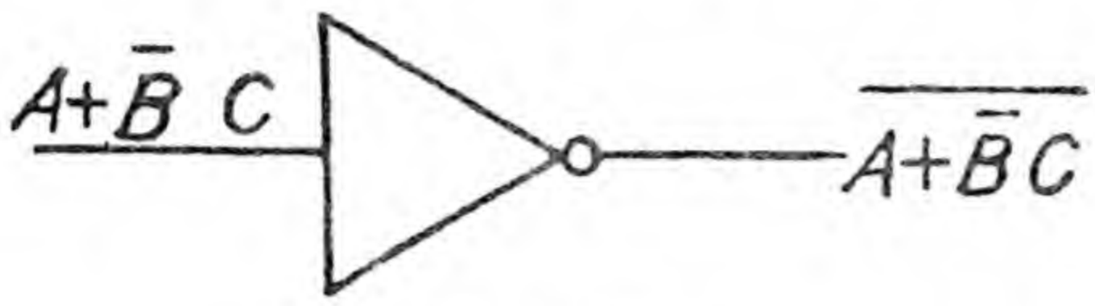
Algebra to Logic : To convert an algebraic expression into a logic diagram easy course is to start with the output and work toward the input. For example, let us convert the expression

$$\overline{A + \bar{B}C}$$

into logic diagram. We note that :

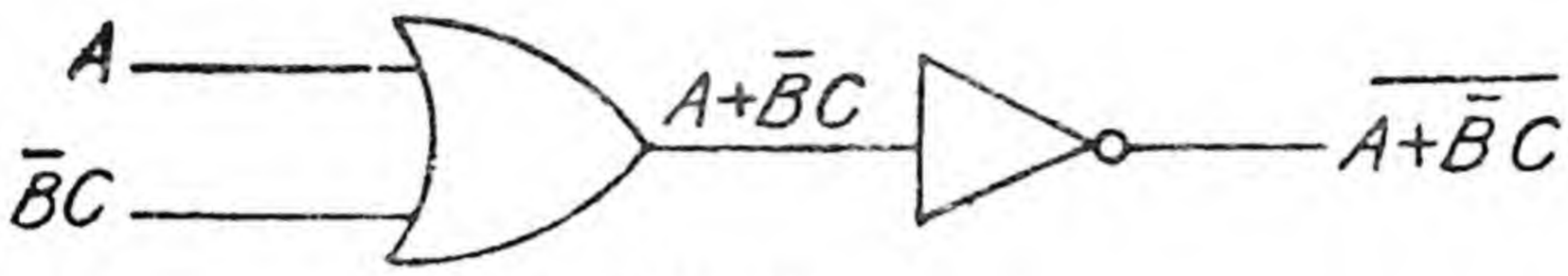
(*i*) It is a NOTed function so it must be an inverse of $A + BC$. Thus at output there is an inverter whose input will be $A + \bar{B}C$. The

circle is at the output end of the triangle indicating that when $A+\bar{B}C$ is high (1 state), $\overline{A+\bar{B}C}$ is low (0 state).



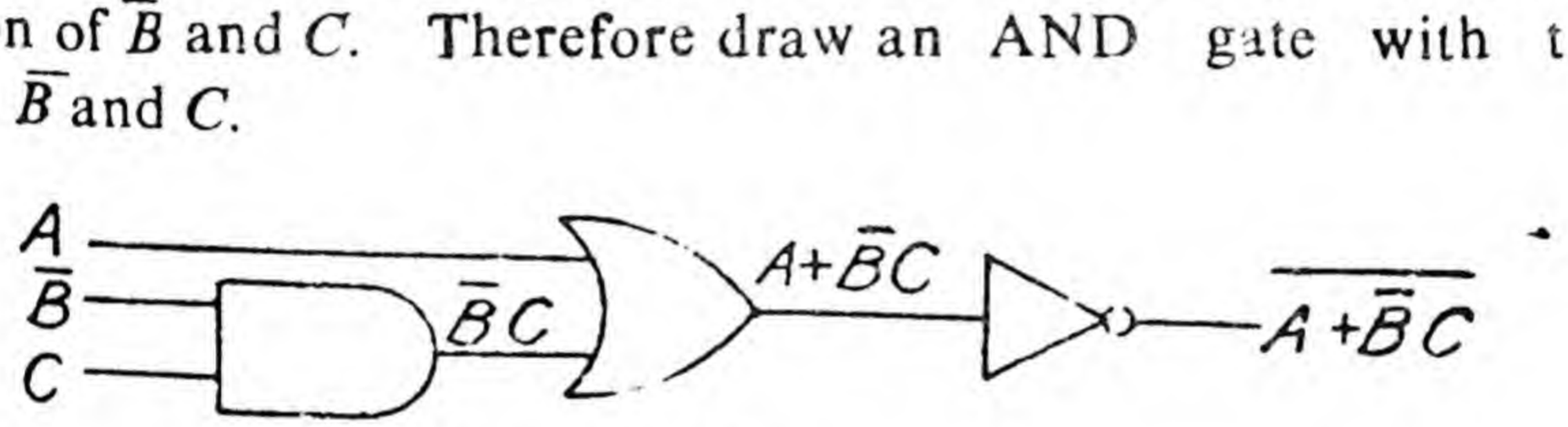
(*a*) Step (*i*)

(*ii*) Now we consider input $A+\bar{B}C$ of invertor. Obviously it is output of an OR gate. Therefore draw on OR gate with two inputs A and $\bar{B}C$.



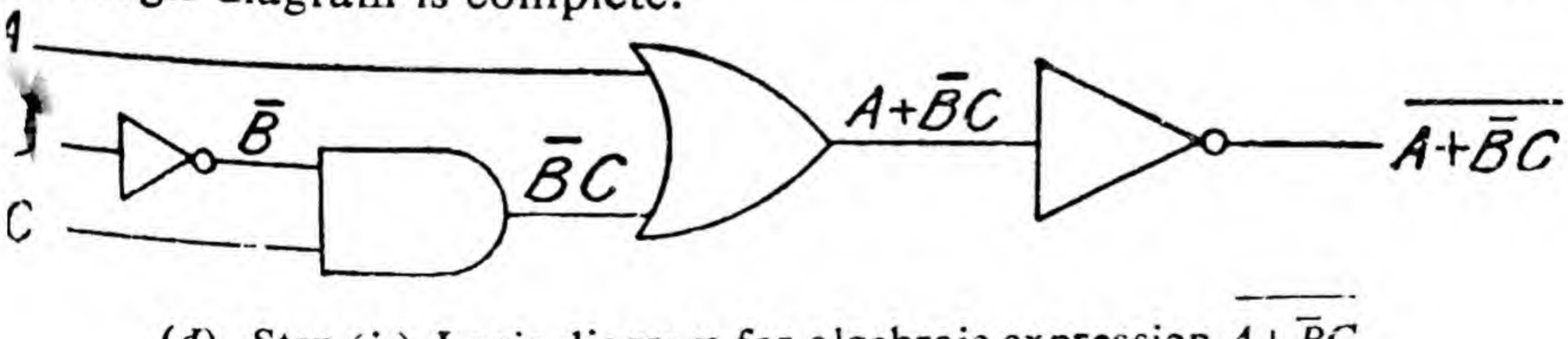
(*b*) Step (*ii*).

(*iii*) Next consider input term $\bar{B}C$ of OR gate. It is an AND function of $\bar{B}$ and C. Therefore draw an AND gate with two inputs $\bar{B}$ and C.



(*c*) Step (*iii*)

(*v*) Next consider input term $\bar{B}$ of AND gate. It can be represented by another inverter with B input. With this inclusion, the logic diagram is complete.



(*d*) Step (*iv*) Logic diagram for algebraic expression $A+\bar{B}C$

Fig. 16. (*a*), (b), (*c*), (*d*).

25·9. ARITHMATIC CIRCUITS : EXCLUSIVE-OR GATES (XOR Gate) :

Some of the fundamental arithmatic circuits, used in digital computers, are exclusive *OR* gate (*XOR* gate), the half and full adders. They are being described below :

The Exclusive OR Gate (XOR gate) · An exclusive *OR* gate is shown in fig. 17 (*a*) with its logic symbol in fig. 17 (*b*). It has two inputs and one output. If the inputs be denoted by A and B then its output is given as

$$Y = A\bar{B}+\bar{A}B$$

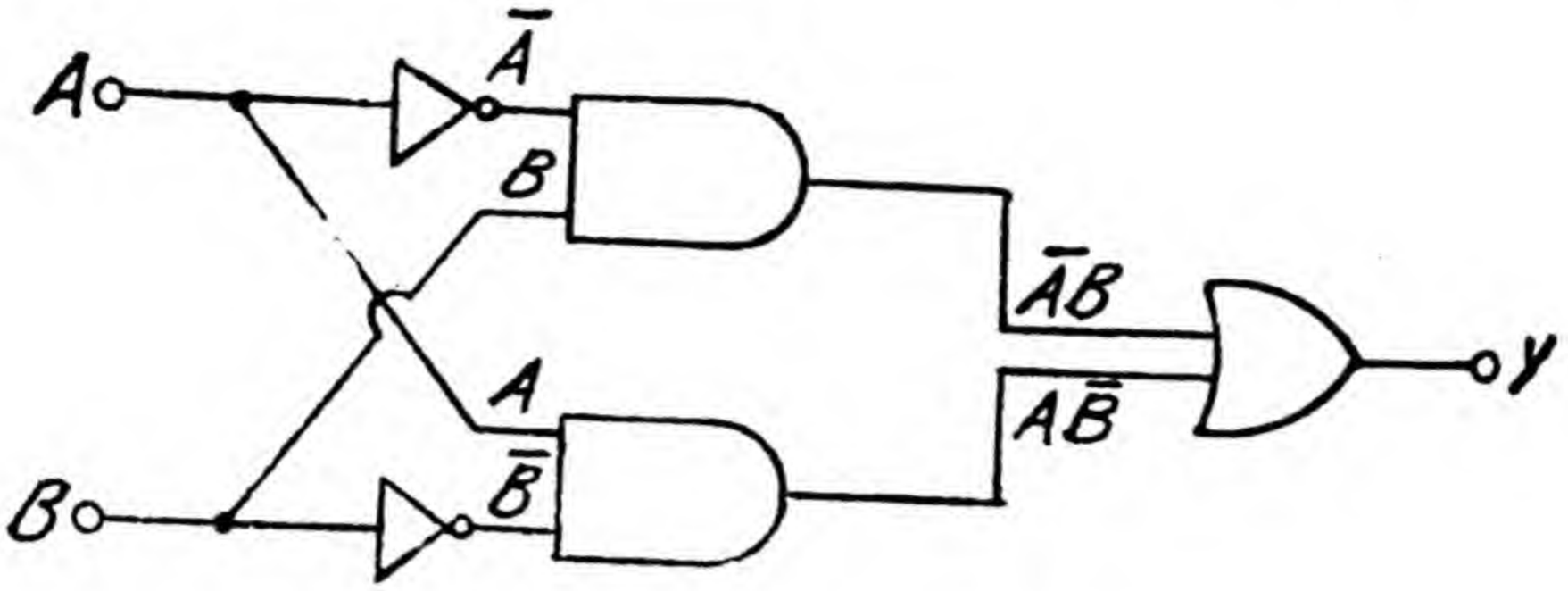


Fig. 17. (*a*) An exclusive OR gate.

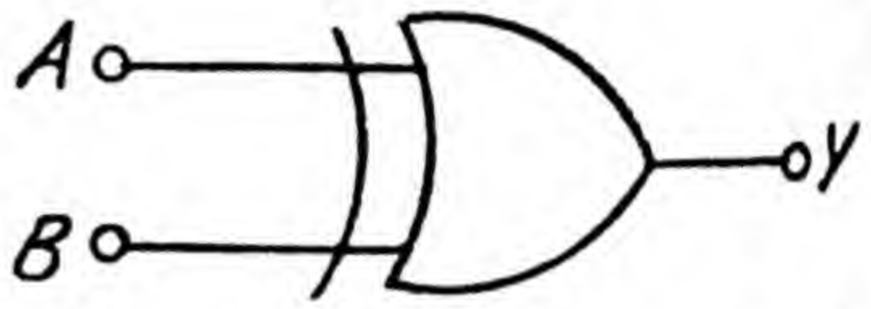


Fig. 17. (*b*) Logic symbol of OR gate.

In order to develop its truth table, we take on the input values as follows :

(*i*) when $A=0$ and $B=0$,

so that $\bar{A}=1$ and $\bar{B}=1$, $Y=0.1+1.0=0+0=0$

(*ii*) when $A=0$ and $B=1$

so that $\bar{A}=1$ and $\bar{B}=0$, $Y=0.0+1.1=0+1=1$

(*iii*) when $A=1$ and $B=0$

so that $\bar{A}=0$ and $\bar{B}=1$, $Y=1.1+0.0=1+0=1$

(*iv*) when $A=1$ and $B=1$

so that $\bar{A}=0$ and $\bar{B}=0$, $Y=1.0+0.1=0+0=0$

We tabulate it as follows :

A	*B*	*Y*
0	0	0
0	1	1
1	0	1
1	1	0

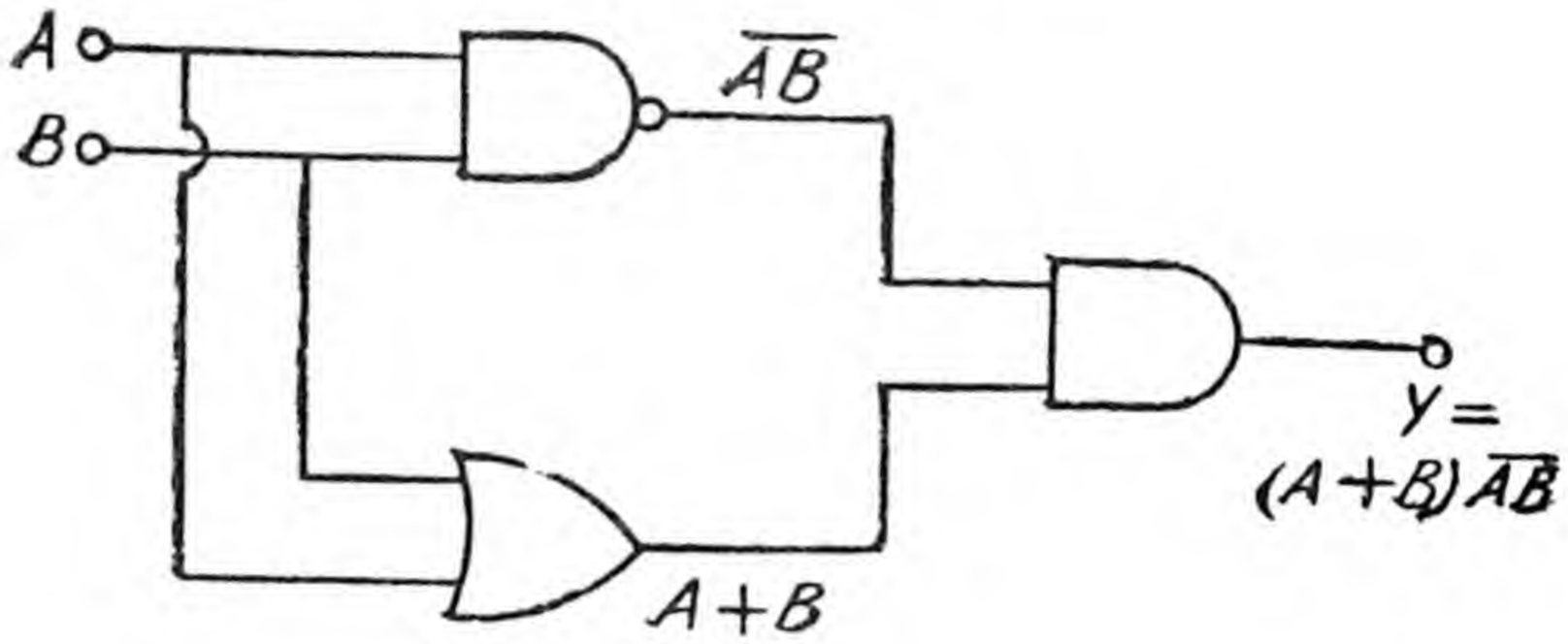
We observe from truth table that *Output Y is a* 1 *only when one of two inputs is* 1 *but not both.* Hence the name exclusive *OR.* Symbolically, we write *XOR* operation as

$$Y=A \oplus B$$

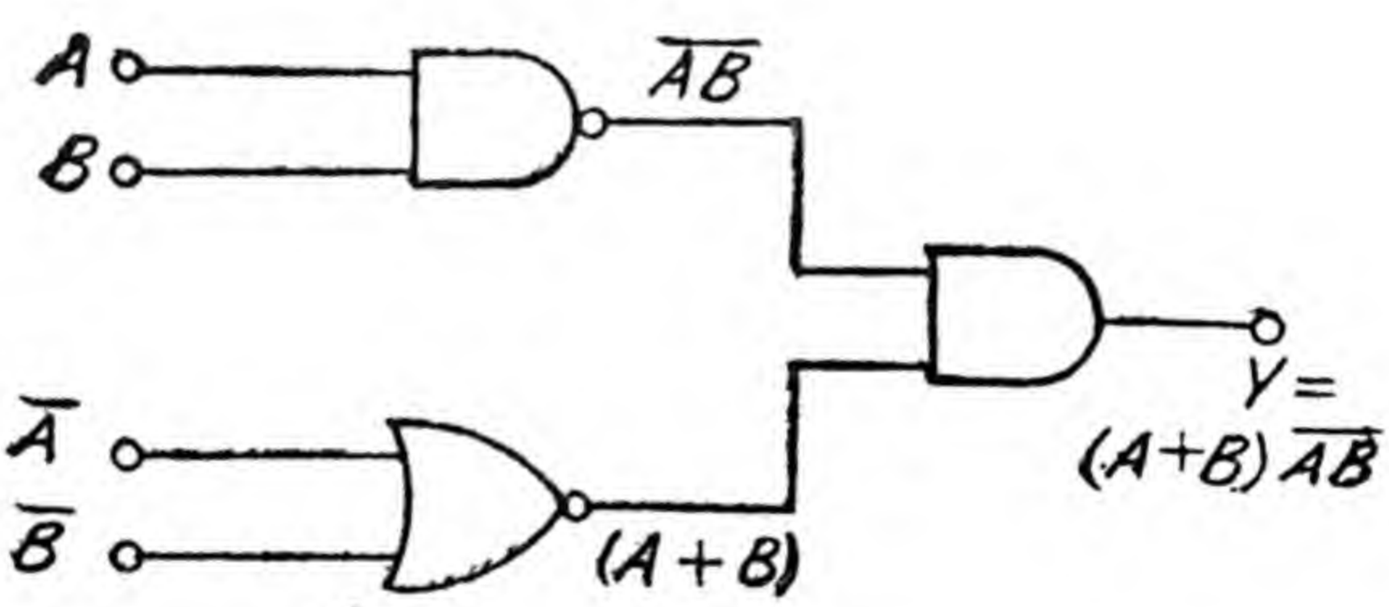
XOR operation is also called *mod-2-addition* and rules of addition are

$$0 \oplus 0 = 0$$
$$0 \oplus 1 = 1$$
$$1 \oplus 0 = 1$$
$$1 \oplus 1 = 0$$

From these rules we conclude that *mod-2-addition* is binary addition if we neglect to take into account the carries. Fig. 18 (*a*) and 18 (*b*) also show the another implementation of *XOR* gate as indicated by the corresponding truth table.



(*a*)



(*b*)

Fig. 18. Circuits having the same logics as of XOR operation.

The truth table is as follows :

A	*B*	*A+B*	$\overline{AB}$	$Y=(A+B)\,\overline{AB}$
0	0	0	$\overline{00}=\bar{0}=1$	$0.1=0$
0	1	1	$\overline{01}=\bar{0}=1$	$1.1=1$
1	0	1	$\overline{10}=\bar{0}=1$	$1.1=1$
1	1	1	$\overline{11}=\bar{1}=0$	$1.0=0$

Thus output is a 1 only when one of the two inputs is a 1 but not both an exclusive *OR* operation.

Applications of XOR gate : *(i) Binary to Gray Code Converter* : An *XOR* gate can be used as a binary to Gray Code converter. Refer to art. 25·4, Ex. 6, where we have converted a

binary to Gray Code. The process of conversion involves addition of each successive bit of binary word and *ignoring* the carry (starting from most significant bit) which is infact a *mod–2–addition i.e.*, an *XOR* gate operation. It has been shown in fig. 19. Binary word 1011 in Gray Code is 1110 (See Ex. 6 art. 25·4). The same is obtained by *mod–2–addition.*

For first from left *XOR* gate input $A=1$, $B=0$ so output $Y=1$
second from left *XOR* gate input $A=0$, $B=1$ so output $Y=1$
third from left *XOR* gate input $A=1$, $B=1$, so output $Y=0$

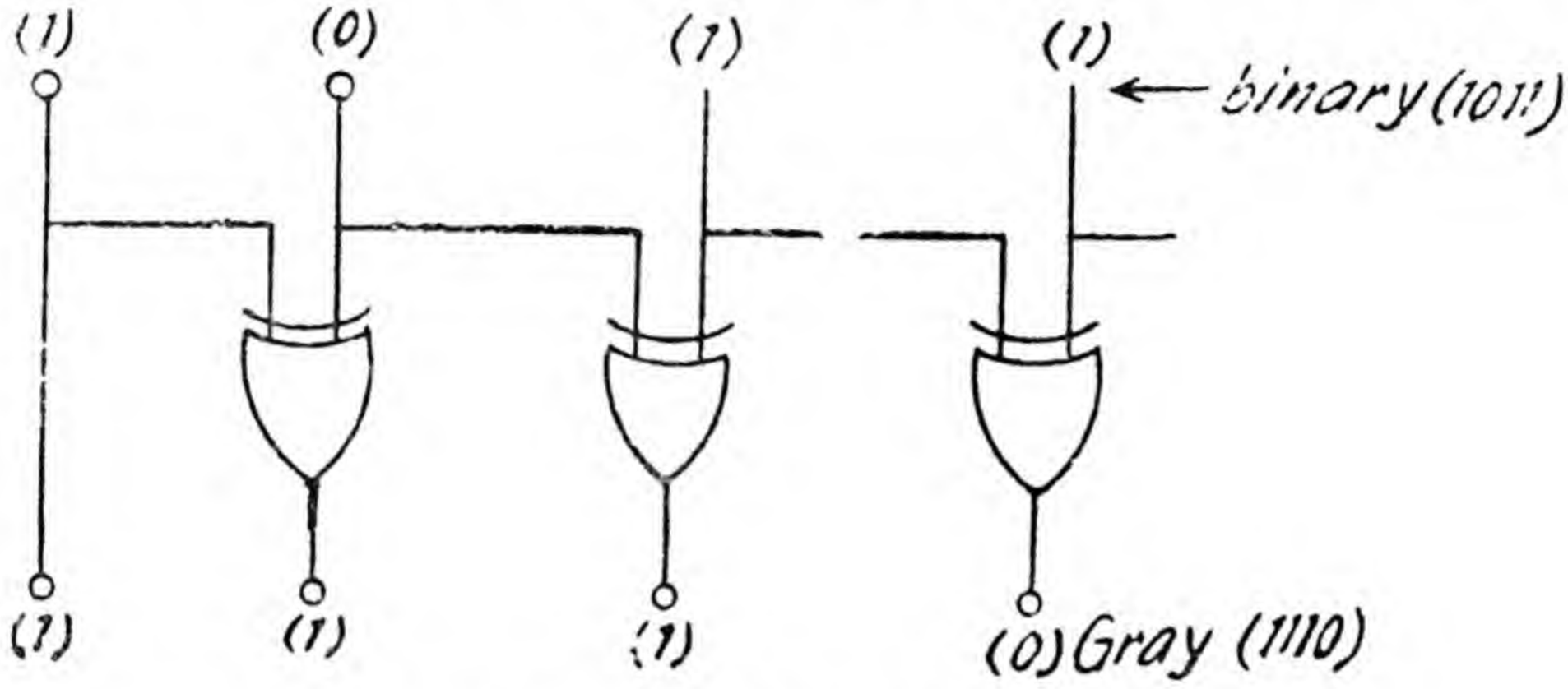


Fig 19. Binary to Gray Code converter.

(*ii*) *A parity checker* : During operation, a digital machine fetches each word (a group of bits) from its memory. Thus words [a group of bits, *e.g.*, machine is to add two *BCD* numbers, 0100 0000 (it is 40) with 0100 0101 (it is 45)] are moved around and it is very likely that 0 may change to a 1 or *viceversa*, *i.e.*, an error may creep in. In order to check it, a parity bit is attached to the word. Two kinds of parity bit exist : an even parity means an extra bit is attached to a group of bits to yield an even number of 1's, while an odd parity makes the number of 1's in a word odd. This new word can be moved and stored by the machine and we can check for even or odd parity at different points to ensure that no error has crept into the word. Thus we need a check whether a word has even parity or odd parity. For this we require a circuit that produces a 0 for even parity and a 1 for odd parity. It can be done if we employ a *mod–2–addition* on each bit of a group of bits (a word) as shown in figs. 20 (*a*) and 20 (*b*). The output $Y=0$ for even parity and $Y=1$ for odd parity.

The Half Adder : *It adds two binary digits at a time* and thus is a basic circuit for adders. When we add 1 and 1, we get

sum=0 and carry=1.

That is,
$$\begin{array}{l} 1 \leftarrow A \text{ input} \\ \underline{1 \leftarrow B \text{ input}} \\ 10 \end{array}$$

↑ ↑
| └ sum
└ – – – carry

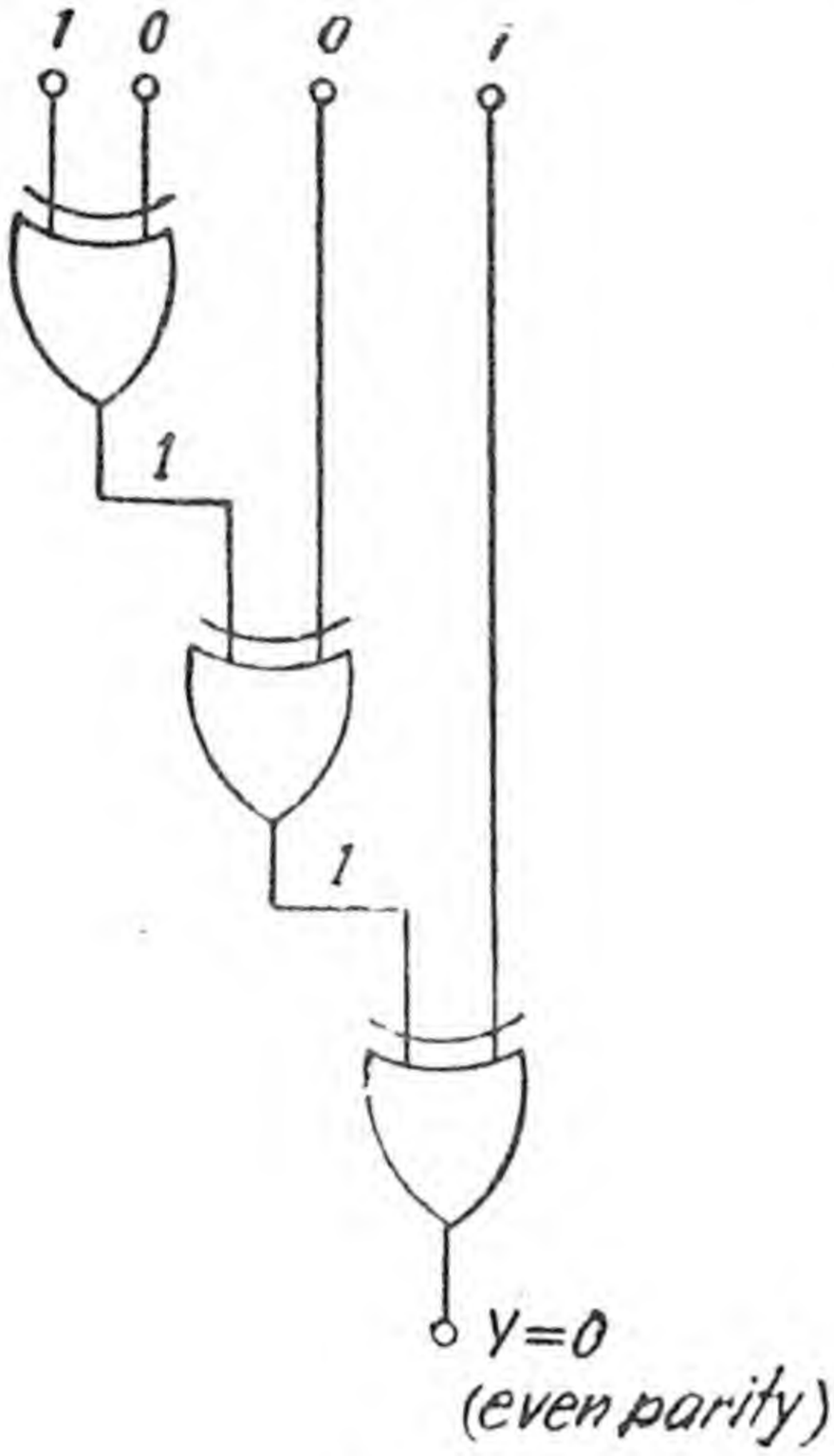


Fig. 20 (*a*)

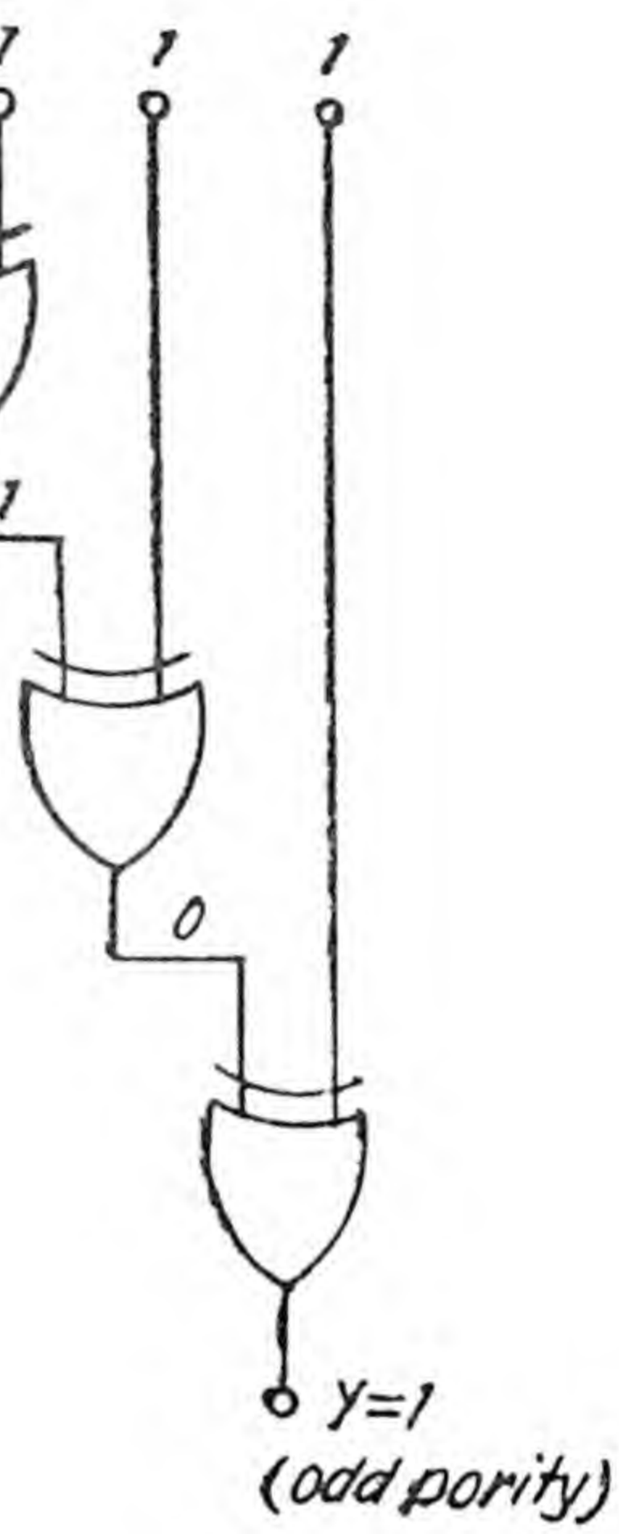


Fig. 20 (*b*)

Fig. 20 XOR gate as a parity checker.

Thus while adding two bits at a time, we get, besides the sum, a carry bit. For sum we use *XOR* operation and carry is obtained by ANDing the two bits. In fig. 21 (*a*) a half adder (including a *XOR* gate and an *AND* gate) is shown. Its symbol is shown in fig. 21 (*b*). The truth table is given below :

A	*B*	Sum	Carry
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

It is the simplest type of adder circuit since it does not have a carry input. Its usefulness is limited to the lowest order bit of a multiple stage adder circuit. It can not be used for higher order bits since it does not have carry input.

The Full-Adder : *It adds three binary digits at a time* and thus has a carry input. While adding 101 with 101, we get

$$\begin{array}{r} 10 \\ +101 \\ \hline 1010 \end{array}$$

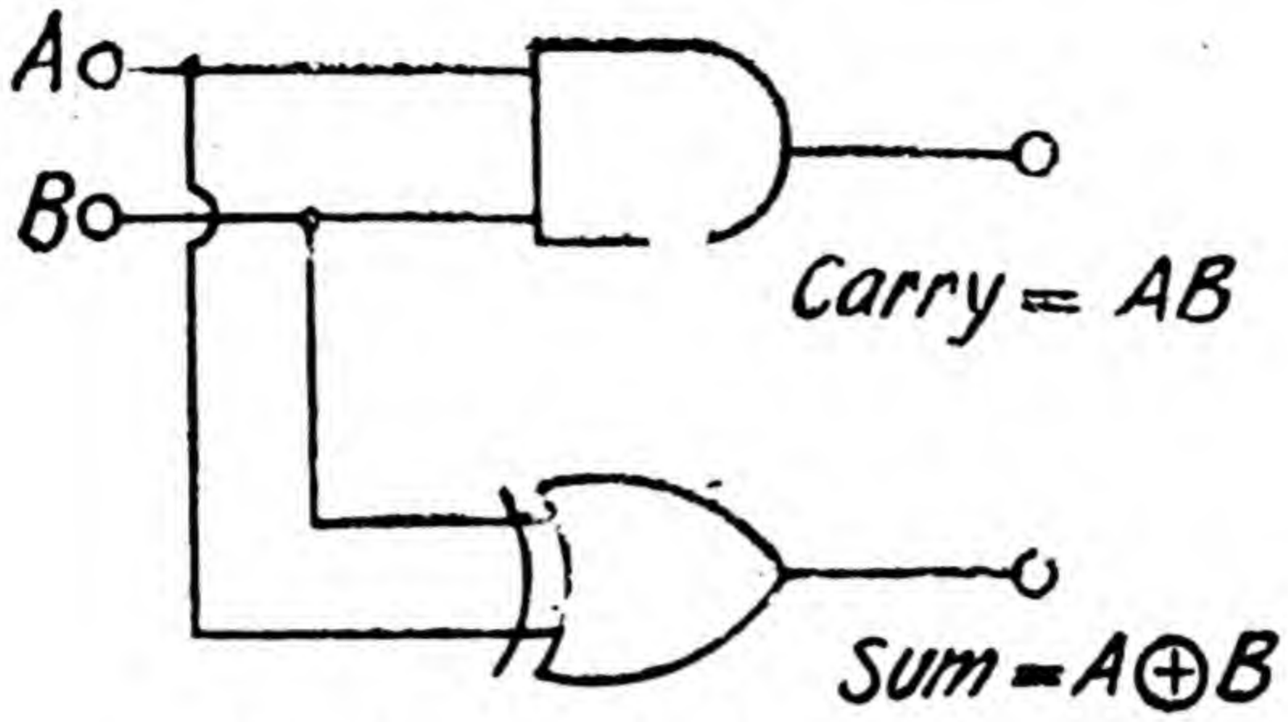


Fig. 21 (*a*) A half adder.



Fig. 21 (*b*) Symbol of half adder.

(*i*) so at the least significant column, 1+1=0 (Sum), and Carry=1

(*ii*) in next column, we have to add three digits because of Carry 0+0+1=1 (Sum), and Carry=0

(*iii*) in next column again we have to add three digits, 1+1+0=0 (Sum), and Carry=1

Thus a full adder needs electronic circuits that can handle three digits at a time. Two half adders and an *OR* gate arranged, as shown in fig. 22, give a full adder. Corresponding truth table is also given. We can write the above example as

```
  1 0 1   ←Carry
    1 0 1 ←A input
    1 0 1 ←B input
  -------
  1 0 1 0 ←Sum
```

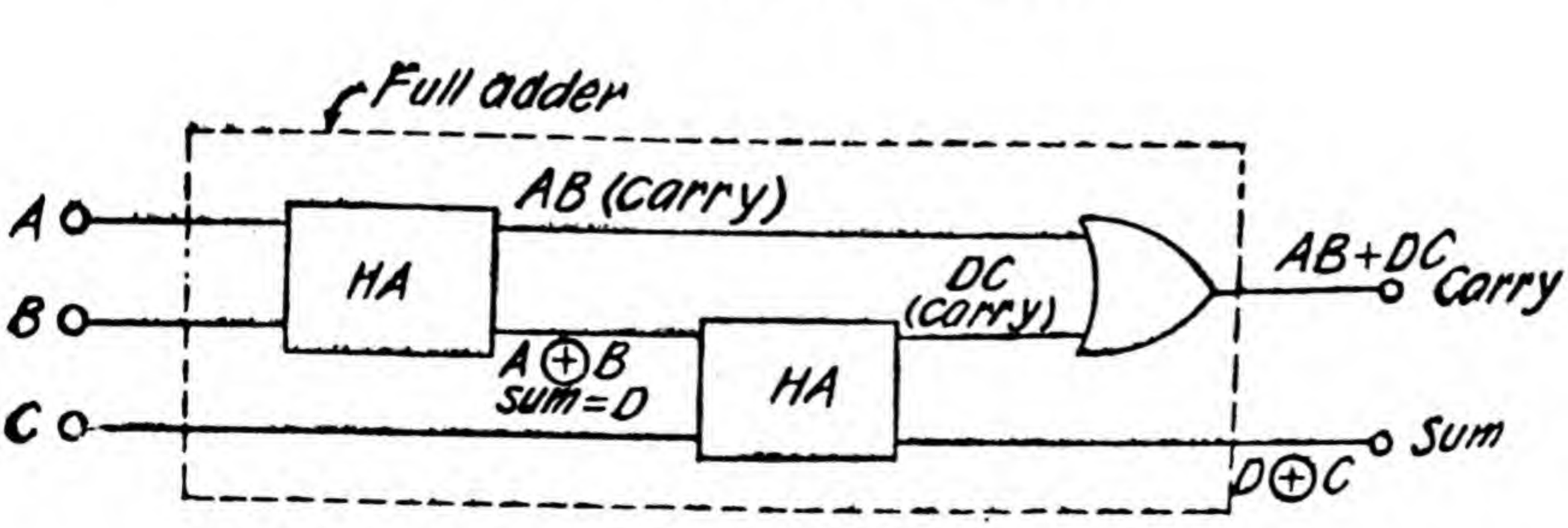


Fig. 22. A full adder using two half adder and an OR Gate.

Thus for third column next to least significant column, we have

0 ←Carry, C input of second adder
1 ←A input } of first adder
1 ←B input }

10
↑↑
| └———Sum of full adder
└————Carry of full adder

Now refer to fig. 22. With $A=1$, $B=1$, $C=0$, we have $D=A\oplus B=1\oplus 1=0$, so that

Carry of full adder $=AB+DC=1.1+0.0=1$
Sum of full adder $= D\oplus C =0\oplus 0=0$

Truth table of a full adder :

A	B	C	Sum	Carry
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

Parallel Adder : By cascading any number of full adders we can obtain an n–bit binary adder circuit, commonly called a parallel adder. The drawback of such a circuit is longest signal-delay time (time taken to give output after the instant the input is applied).

Suppose we are to add four bit binary numbers 1111 with 1011 then

111 ←Carry
1111 ←A input
+1011 ←B input

11010
↑
└————over flow carry if sum exceeds four bits.

or in general

$a_4\ a_3\ a_2\ a_1$
$+b_4\ b_3\ b_2\ b_1$

$S_4\ S_4\ S_3\ S_2\ S_1$
↑
└——over flow carry

The first column requires only one half adder as only two bits are to be added and there is no carry, but all subsequent columns need full adder because we are to add three bits (two digits and one carry). A parallel four bit binary adder is shown in fig. 23.

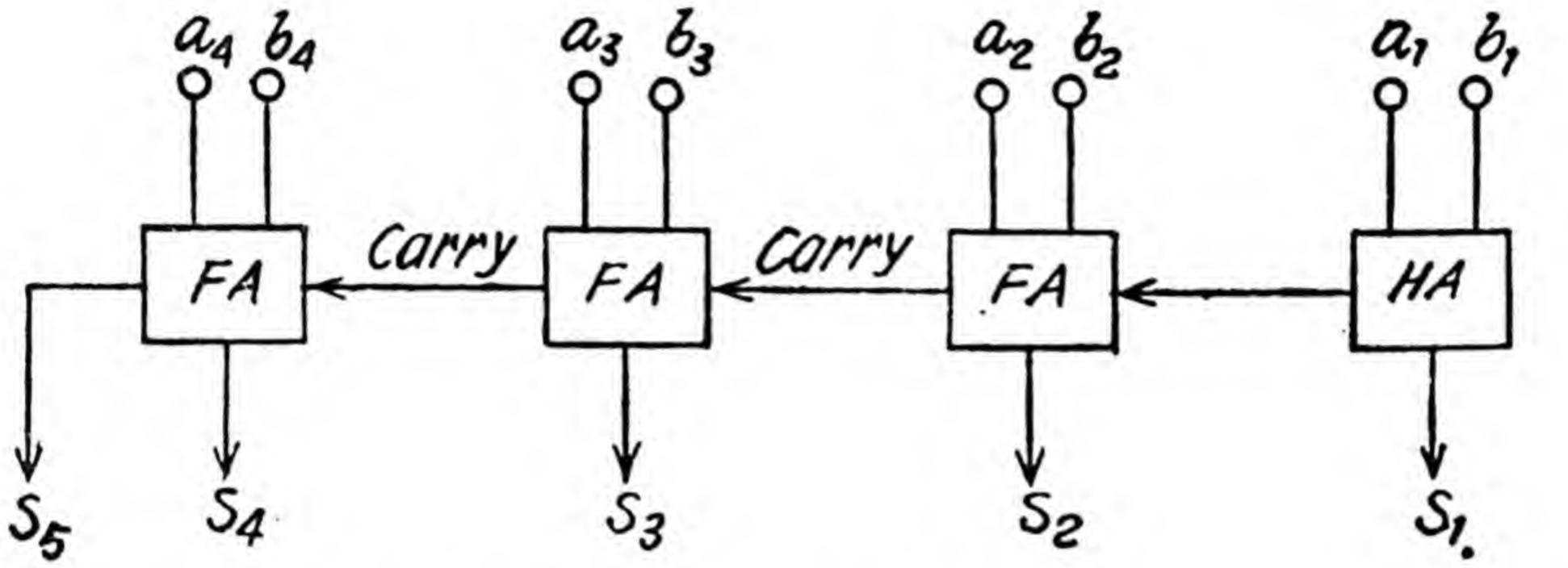


Fig. 23. A four bit binary adder.

Subtraction : In order to subtract two four-bit binar numbers *e.g.* subtraction of $b_4\, b_3\, b_2\, b_1$ from $a_4\, a_3\, a_2\, a_1$ we first invert each *b*-bit to get 1's complement of $b_4\, b_3\, b_2\, b_1$ *i.e.* $\bar{b}_4\, \bar{b}_3\, \bar{b}_2\, \bar{b}_1$ and then add to $a_4\, a_3\, a_2\, a_1$ in the following way :

We know that subtraction of a number *A* from *B* is in fact the addition of *A* and $-B$. Thus an adder *complements* the magnitude of *B* and then magnitude of *uncompensated* number (*A*) is added to the magnitude of the complemented, $\bar{B}$. If there is an overflow, called an *end-around carry* (*EAC*), it is added to the least significant bit (*lsb*) digit of the sum and the sum will be of the same sign as that of the number that was *not complemented* (means of the sign of *A*). Further if there is *no EAC*, the sum is complemented and the sign bit of the sum will be the same as that of the number that *was complemented*. A parallel four bit subtrator is shown in fig. 24. It uses four full adders and end around carry (*EAC*).

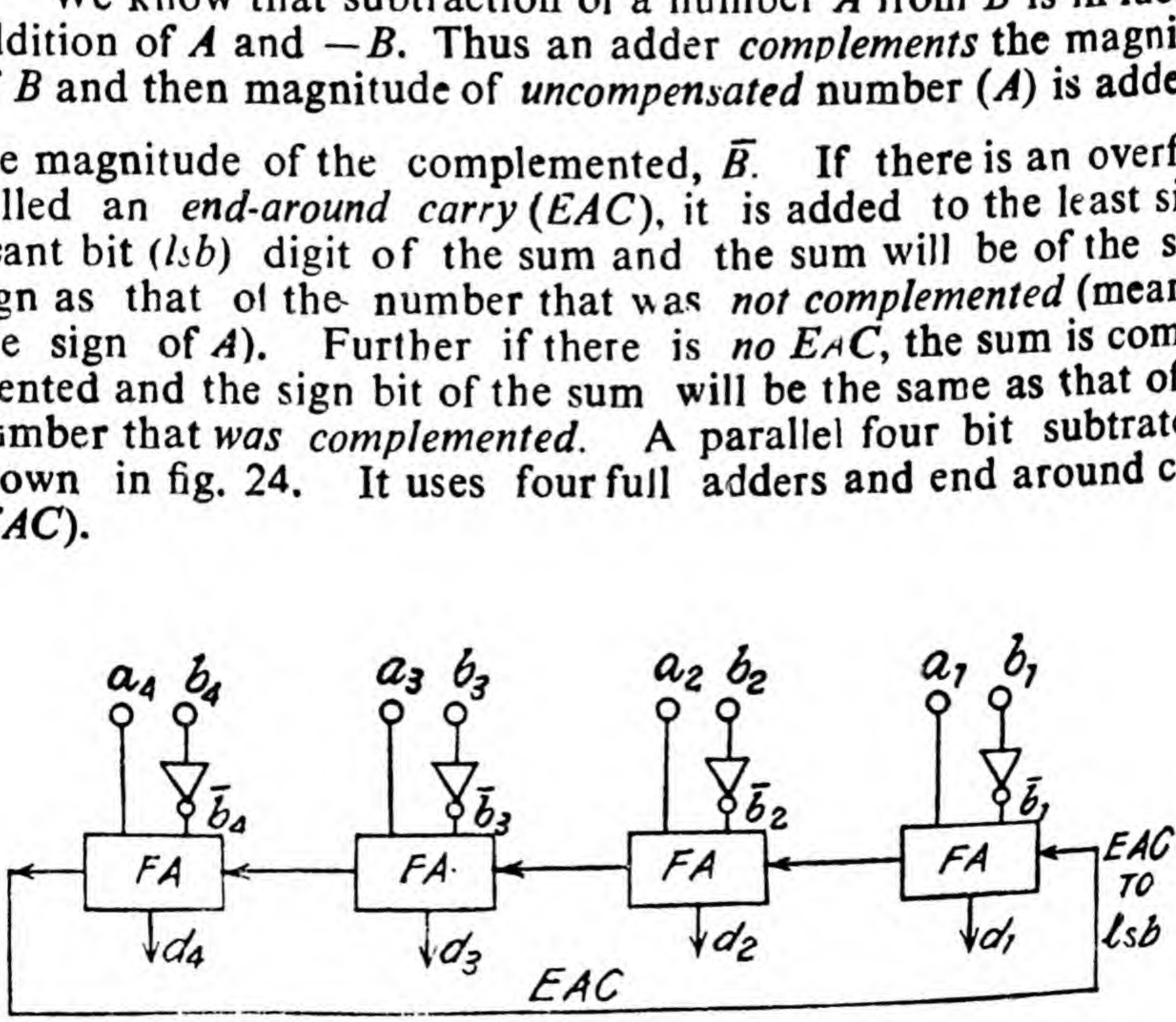


Fig. 24. A parallel four-bit subtractor.

25·10. SIMPLIFICATION USING KARNAUGH MAPS :

In order to reduce the cost of implementing boolean functions, boolean-algebra reduction techniques are employed. We discuss here a reduction short cut method, called Kernaugh-map reduction (pronounced kar no). It is an efficient way of reducing boolean functions to a minimum form so that minimum hardware is required to implement.

A truth table shows the output for each input condition. A Karnaugh map shows the *fundamental products* needed to produce the output 1's for corresponding input conditions.

Two Variable Karnaugh Maps : We shall construct a Karnaugh map from the truth table, taking as an example, of fig. 25 (*a*). The first step is to write $\bar{X}$ followed by X in vertical column and $\bar{Y}$

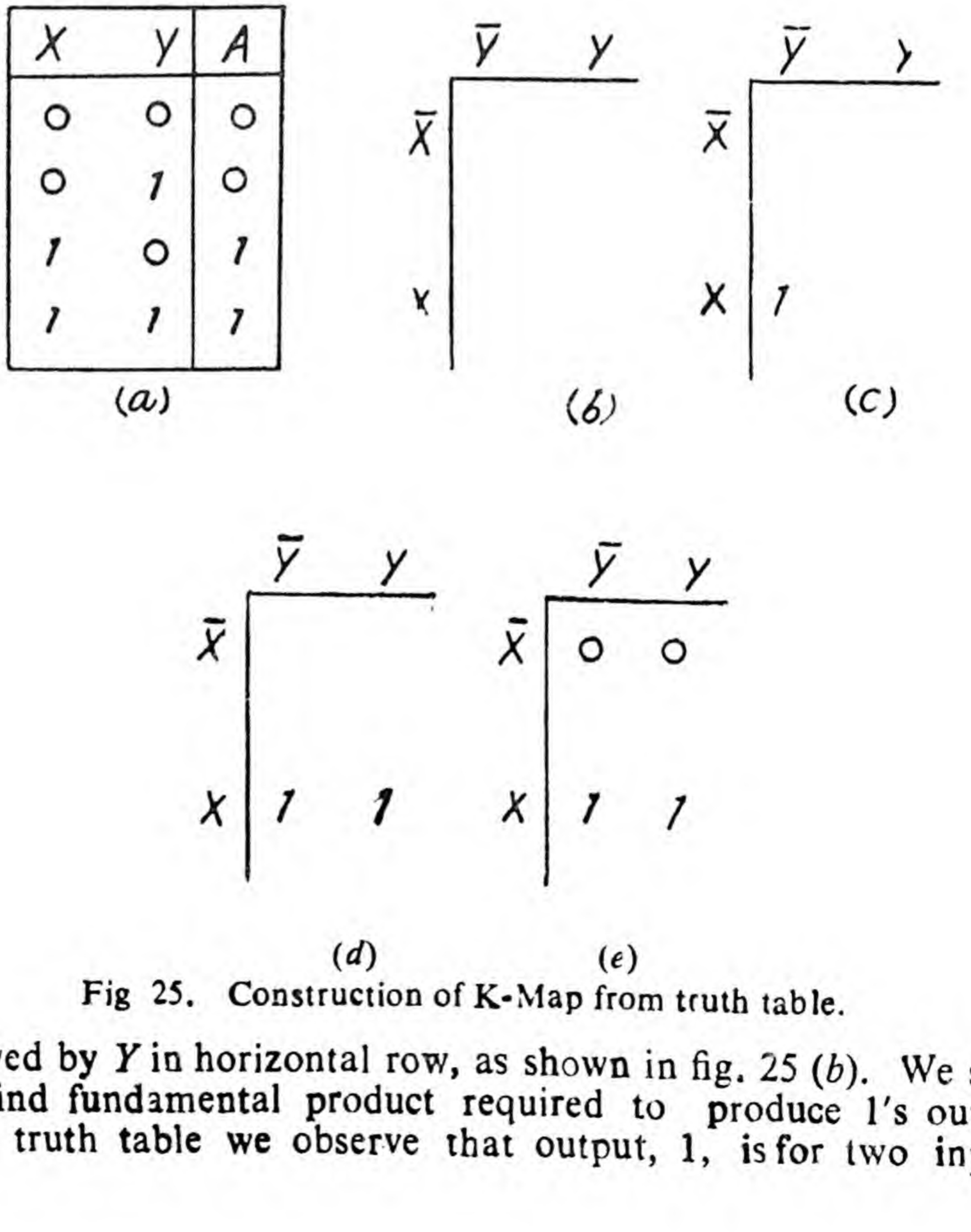


Fig 25. Construction of K-Map from truth table.

followed by Y in horizontal row, as shown in fig. 25 (*b*). We shall now find fundamental product required to produce 1's output. From truth table we observe that output, 1, is for two inputs.

	Output (A)	Input	Fundamental product for 1's output	Entry in Map
(i)	1	$XY=10$	$X\bar{Y}=1\bar{0}=11=1$	Enter 1 as shown in fig. 21 (c). It gives $X\bar{Y}$.
(ii)	1	$XY=11$	$XY=11=1$	Enter 1 more in the position shown in fig. 21 (d) so that it gives XY.

(ii) For remaining spaces in the map, enter 0's. These zero mean that fundamental products are not needed for the corresponding inputs $XY=00$ and $XY=01$.

Thus fig. 25 (e) shows the complete Karnaugh map for truth table of fig. 25 (a).

Multiple Variable Karnaugh Maps : We shall now construct a Karnaugh map for four input variables W, X, Y and Z. We have chosen four bit word because many digital computers use them. Truth table is shown in fig. 26 (a) and the corresponding K-map in fig. 26 (b).

W	X	Y	Z	A
0	0	0	0	0
0	0	0	1	1
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
0	1	1	0	1
0	1	1	1	1
1	0	0	0	0
1	0	0	1	0
1	0	1	0	0
1	0	1	1	0
1	1	0	0	0
1	1	0	1	0
1	1	1	0	1
1	1	1	1	0

(a) Truth table

	$\bar{Y}\bar{Z}$	$\bar{Y}Z$	YZ	$Y\bar{Z}$
$\bar{W}\bar{X}$	0	1	0	0
$\bar{W}X$	0	0	1	1
WX	0	0	0	1
$W\bar{X}$	0	0	0	0

(b) K-Map for truth table of fig. (a).

Fig. 26. Four variable Karnaugh map.

We observe from fig. 26 (*a*) that outpu' 1's appear for inputs : $WXYZ=0001$ So that fundamental products are : $\bar{W}\,\bar{X}\,\bar{Y}\,Z=1111=1$

$WXYZ=0110$ $\bar{W}\,X\,Y\,\bar{Z}=1111=1$

$WXYZ=0111$ $\bar{W}\,X\,Y\,Z=1111=1$

$WXYZ=1110$ $W\,X\,Y\,\bar{Z}=1111=1$

These fundamental products have been shown in K-map fig. 26 (*b*).

Pairs, Quads and Oc'ets :

Pair : Refer to fig. 27 (*a*). It contains a pair of 1's that are horizontally adjacent where

first 1 represents the fundamental product $W\,X\,Y\,Z$

second 1 " " " $W\,X\,Y\,\bar{Z}$

Thus as we move from first 1 to the second 1, only one variable goes from uncomplemented form to complemented form (Z to $\bar{Z}$) while other variables W, X, Y do not change. *Whenever this happens, we can eliminate the variable that changes form*, so that we can write the output as

$$A=WXYZ+W\,X\,Y\,Z$$
$$=WXY$$

Similarly for fig. 27 (*b*),

$$A=WXY\bar{Z}+W\bar{X}Y\bar{Z}$$
$$=WY\bar{Z}$$

For fig. 27 (*c*), we have to separate pairs for which

$$A=[\bar{W}X\bar{Y}Z+\bar{W}XYZ]+[WX\bar{Y}\,\bar{Z}+W\bar{X}\,\bar{Y}\,\bar{Z}]$$
$$=\bar{W}XZ+W\bar{Y}\,\bar{Z}$$

Pair is never formed with diagonally adjacent 1's as they do not lead to any simplification.

Quad : Refer to fig. 27 (*d*) and fig. 27 (*e*). A quad is a group of *four* 1's that are horizontal or vertically adjacent. Complements in products drop out of the Boolean equation. Thus for quad in fig. 27 (*d*), we have

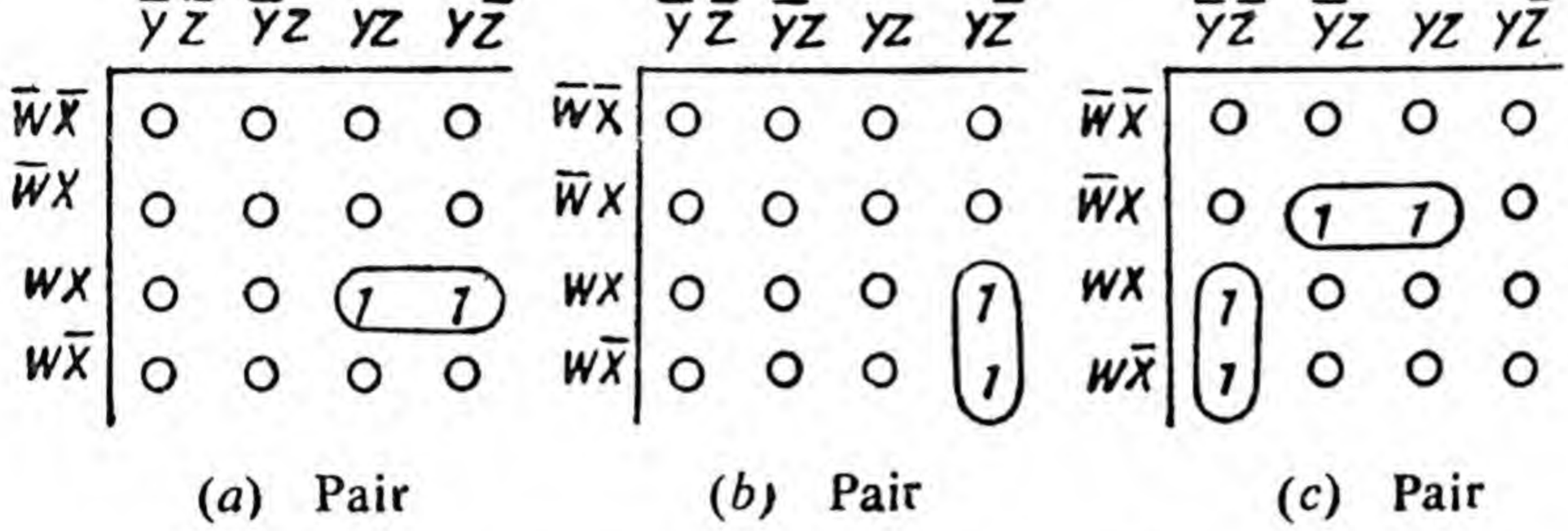
$$A=WX\bar{Y}\,\bar{Z}+WX\bar{Y}Z+WXYZ+WXY\bar{Z}$$
$$=WX$$

because Y and Z have been complimented and so drop out. We can see that how simplified becomes the expression for A after the reduction.

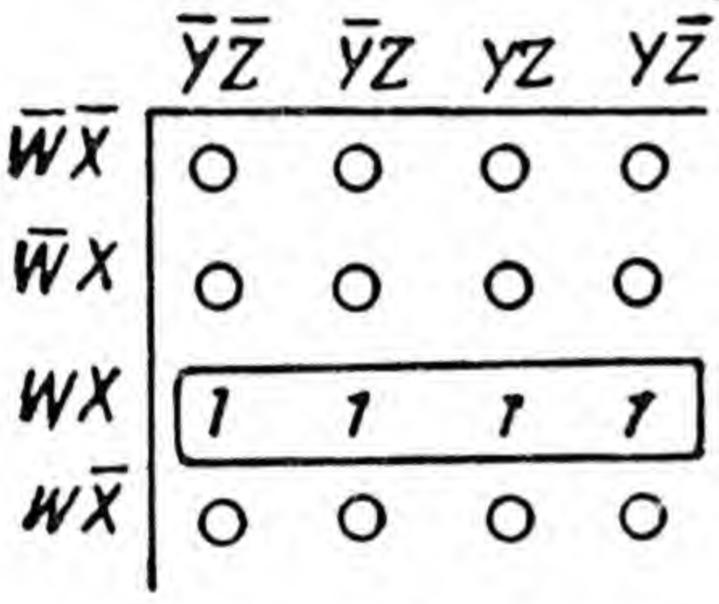
Octet : Refer to fig. 27 (f) and 27 (g). An octet is a group of eight adjacent 1's as shown in fig. 27 (f). The output A, for this octet is

$$A = WX\bar{Y}\,\bar{Z} + WX\bar{Y}Z + WXYZ + WXY\bar{Z} + W\bar{X}\,\bar{Y}\,\bar{Z} + W\bar{X}\,\bar{Y}\,Z + W\bar{X}\bar{Y}Z + W\bar{X}Y\,\bar{Z}$$
$$= W$$

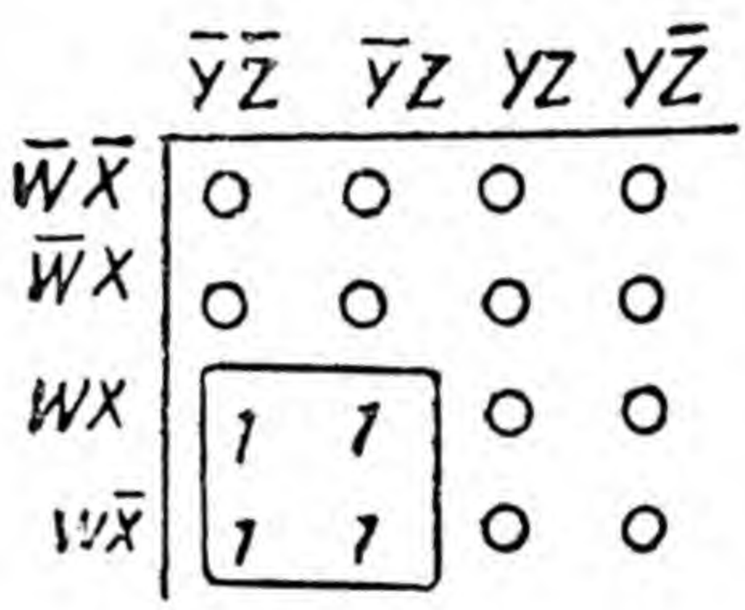
because each variable except W has been complemented and so drop out.



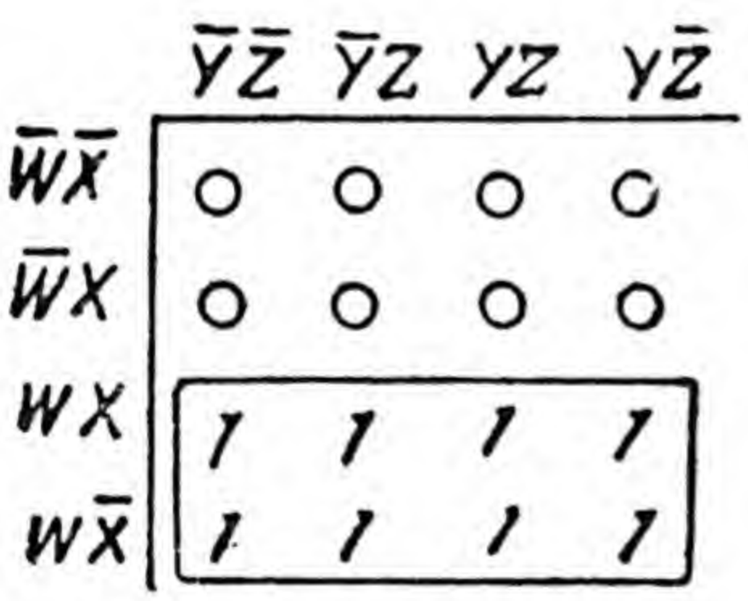
(*a*) Pair (*b*) Pair (*c*) Pair



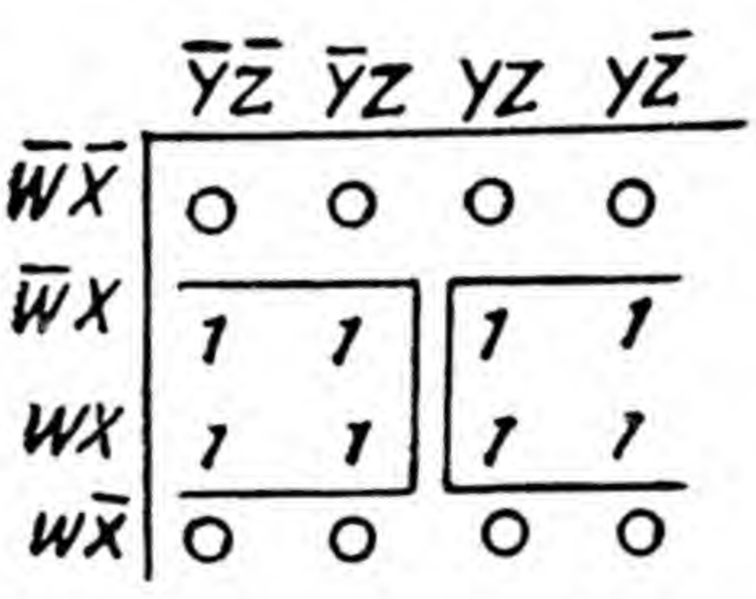
(*d*) Quad



(*e*) Quad



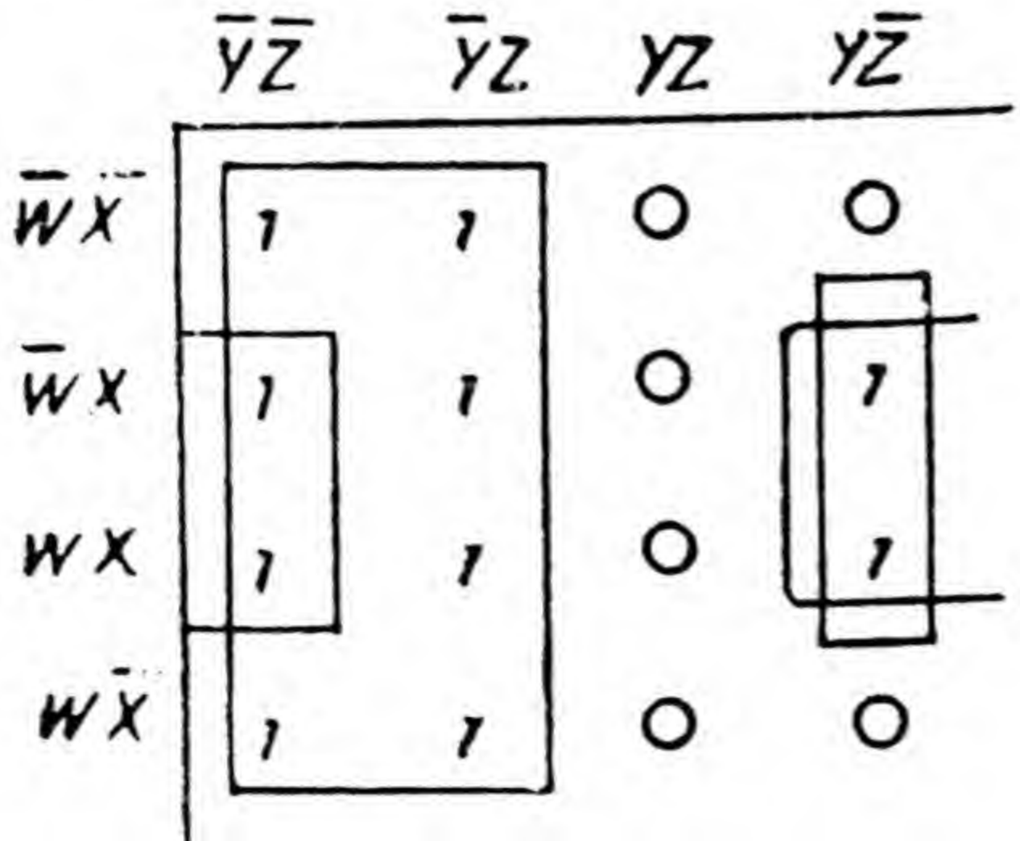
(*f*) Octet



(*g*) Octet, Rolling.

Fig. 27. Pair, quad and octet.

Karnaugh Rolling : Thus for simplified Boolean expression,



we look in K-map for octets, then for quads and finally for pairs and encircle them. If possible roll and overlap to get the largest groups. Fig. 28 shows an efficient way to encircle groups. If we take octet and pair in it then output

$$A=[\bar{W}X\bar{Y}Z+WXYZ]_{pair}$$
$$+[\bar{W}\,\bar{X}\,\bar{Y}\,\bar{Z}+\bar{W}\,\bar{X}\,\bar{Y}\,Z+\bar{W}X\bar{Y}\,\bar{Z}+\bar{W}X\bar{Y}Z$$
$$+WX\bar{Y}\,\bar{Z}+WXYZ+W\bar{X}\,\bar{Y}\,\bar{Z}+W\,\bar{X}\,\bar{Y}\,Z]_{octet}$$
$$=XYZ+\bar{Y}$$

But still more reduction can be achieved if rolling and overlapping is done. Then

$$A=\bar{Y}\text{ (for octet above)}+[\bar{W}X\bar{Y}\,\bar{Z}+WX\bar{Y}\,\bar{Z}+\bar{W}XYZ+WXY\bar{Z}]_{quad}$$
$$=\bar{Y}+X\bar{Z}$$

Simplification of Boolean equation and its realisation through gates : Let us take a K-map shown in fig. 29 (*a*). In it there are no octets but there are three quads using overlapping technique. The

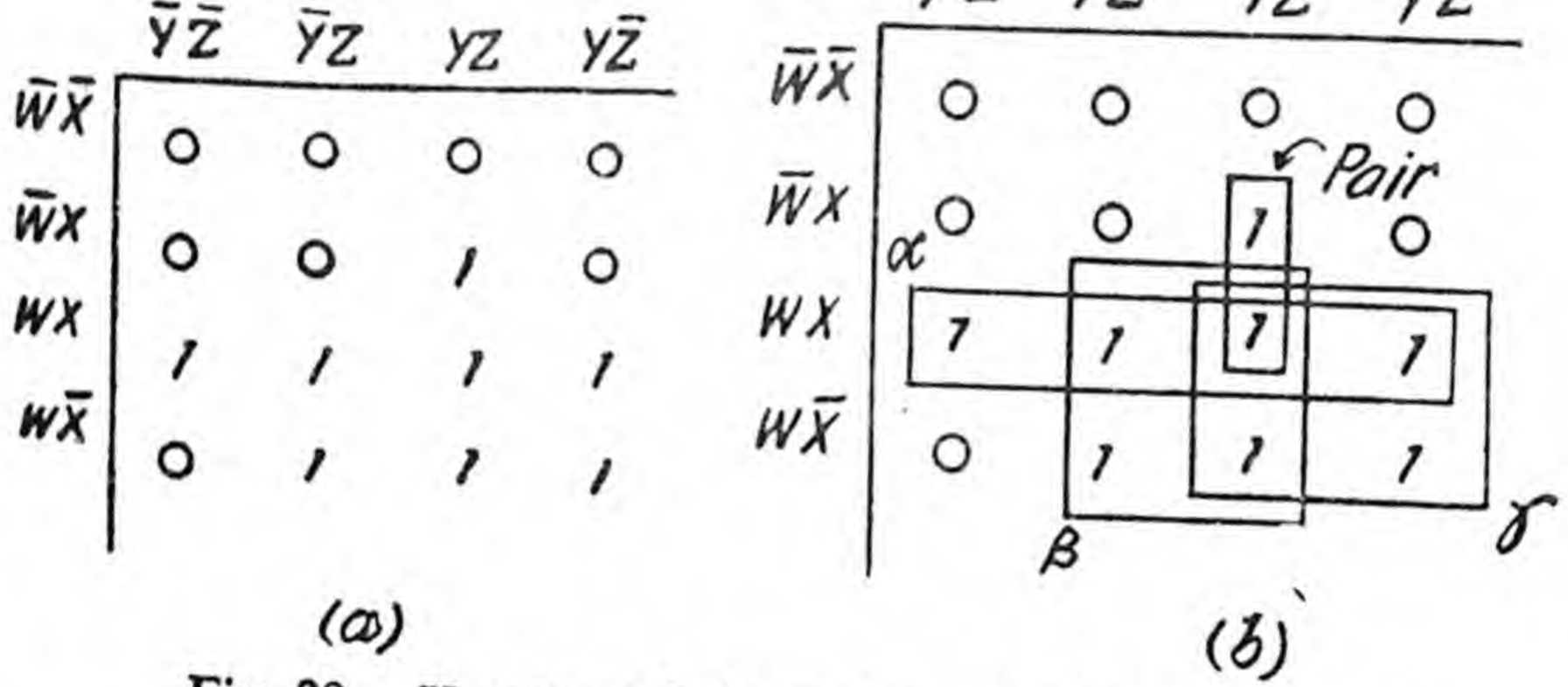


Fig. 29. K map and formation of quads and a pair.

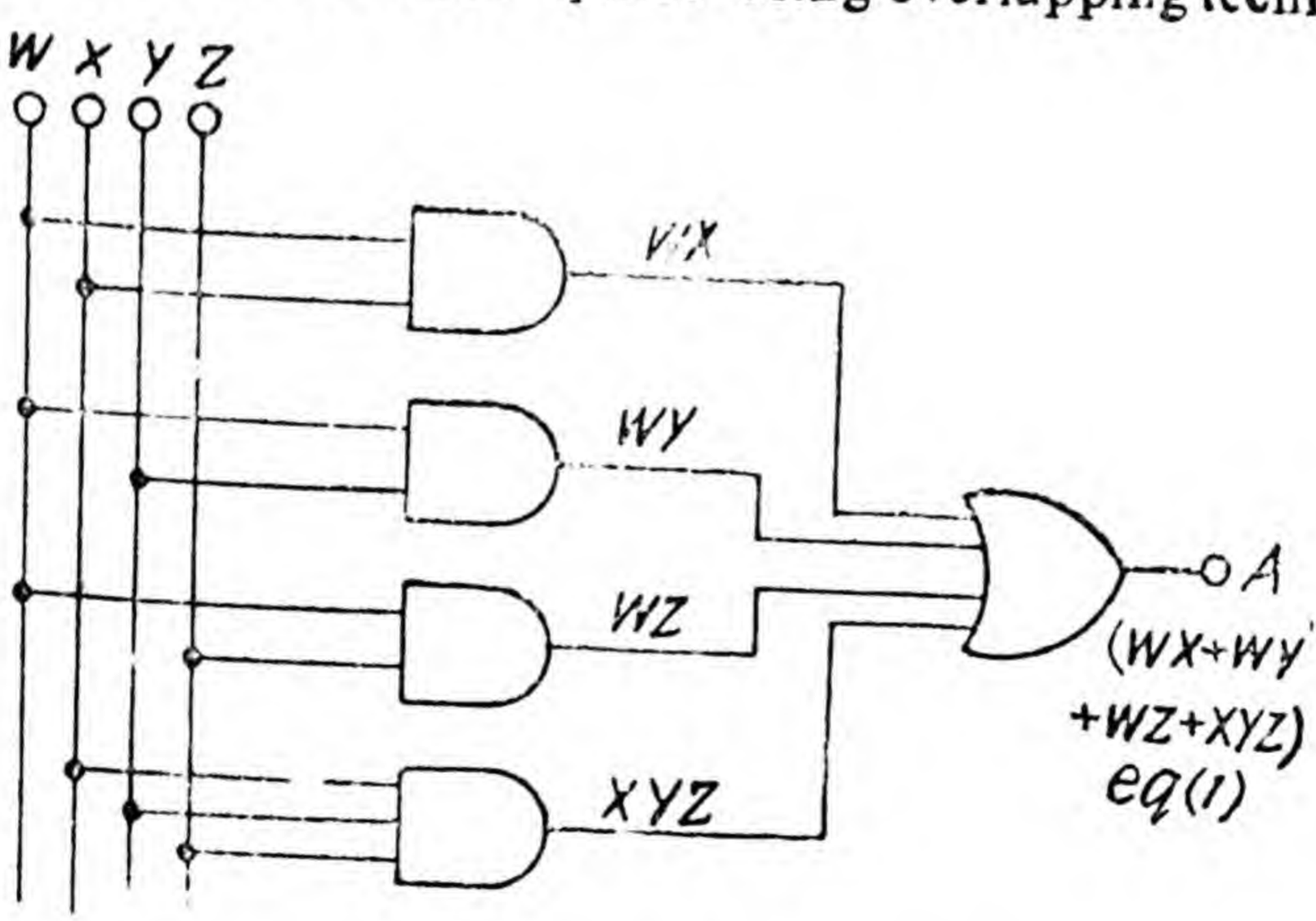


Fig. 30. (*a*) Realisation of Boolean eq. (1).

remaining one is made a part of overlapped pair fig. 29 (*b*). The horizontal quad denoted by α corresponds to the simplified product *WX*. the left square quad denoted by β corresponds to *WZ* and right square quad denoted by *Y* corresponds to *WY*. The simplified product for pair is *XYZ*. Thus simplified Boolean expression is

$$A = WX + WY + WZ + XYZ \quad \ldots(1)$$

or

$$A = W(X + Y + Z) + XYZ \quad \ldots(2)$$

Using AND-OR gate network eq. (1) [fig. 30 *a*] and eq. (2) [fig. 30 *b*] can be realised.

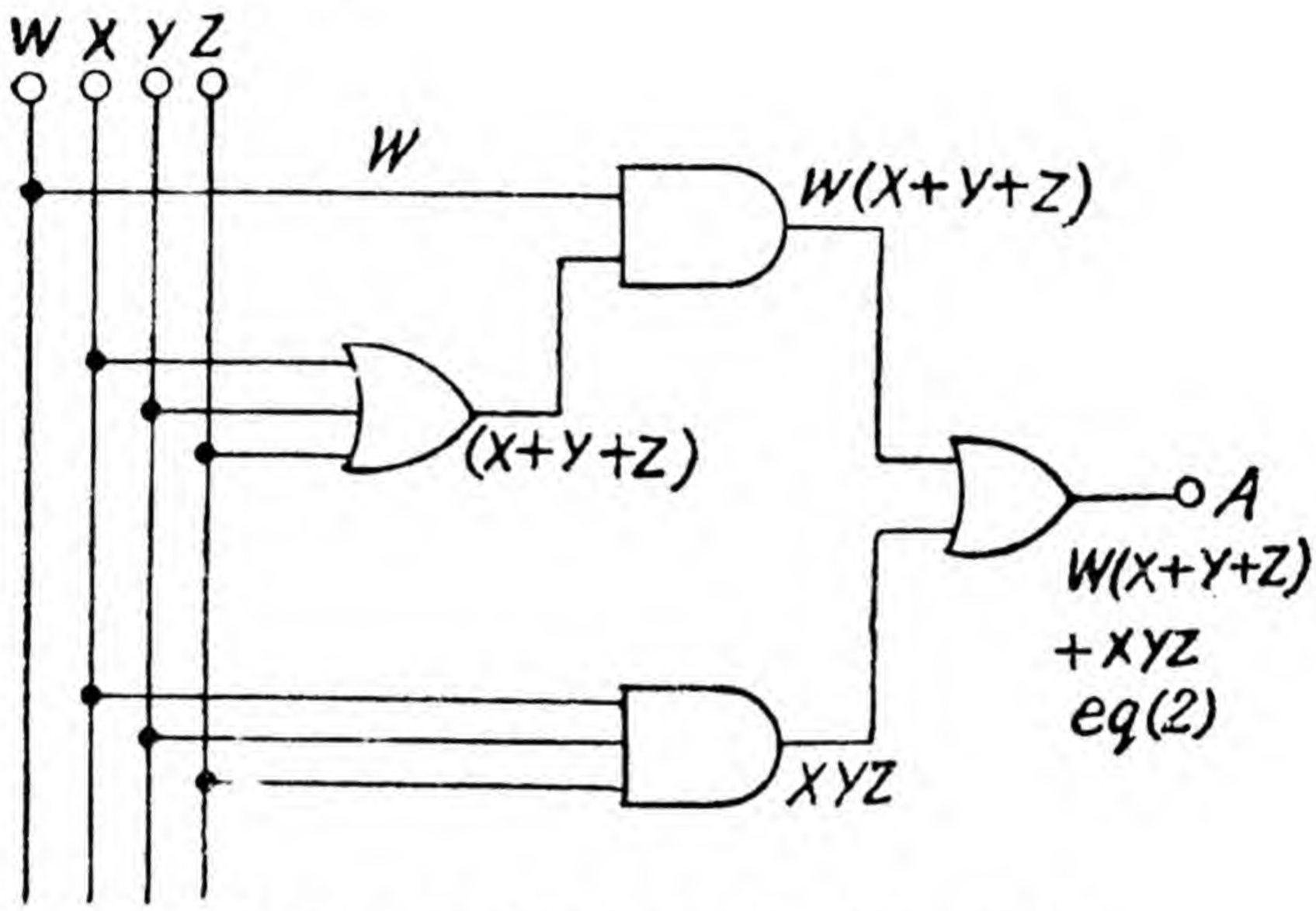


Fig. 30 (*b*). Realisation of Boolean eq. (2).

25·11. LOGIC GATE PARAMETERS :

There are a number of logic circuit families available from various manufacturers. Choice between them is based on the consideration of following parameters :

(i) **Fan-in and Fan-out :** The maximum number of inputs which can be applied to logic gate is known as 'Fan-in'. Thus *a* 3-input NAND gate has fan-in of 3.

The fan-out of a logic gate is the number of gates that can be driven by it. Thus, if a fanout of a typical NOR gate is '6', then it implies that this NOR gate can drive 6 such NOR gates.

(ii) **Noise Margin :** It is the property of logic circuit to withstand unwanted noise signals at input and is the maximum value of noise signals which a system can reject with performance unaffected.

(iii) **Noise Immunity :** It is the voltage which when applied as the input of a logic gate will cause the output to change the output state for a given supply voltage.

(iv) **Propagation Delay Times :** This determines how fast the logic system can operate. There is a definite time interval

between the appearance of the output and the application of input to the gate. This time interval is called propagation delay.

In other words, it can be defined as the time interval between a change in input state and the resulting change in output state. It is about 10 nano seconds for TTL gate. The delay for a single gate appears to be insignificant but becomes appreciable when gates are used in cascade.

(v) Power Dissipation : From the point of view thermal, packaging, reliability, economy, it is better to have low power dissipation per gate which is usually from few microwatts to 50 m.w. For TTL gate power dissipation is 8 mw. High operating speed requires higher power and therefore a compromise between operating speed and power dissipation is desirable.

25·12. LOGIC FAMILIES :

A number of logic families are available in the market. Resistor Transistor Logic (RTL) was the first family of logic circuits established. It offered high performance but low noise margins. The next family introduced was Diode Transistor Logic (DTL) which was slow but having better noise margins and larger fan-outs. The third to be introduced was Transistor–Transistor Logic (TTL). This provided greater operating speed than DTL. It is the most popularly used industrial family. Thus logic families are classified according to the various characteristics and for a particular application, proper choice may be made.

It is to remembered that logical 1 and 0 are represented, in most modern logic systems, by voltage levels. *Positive* logic (for active high levels) means that the most positive logic voltage level (also referred to as the *high* level) is defined to be the logical 1 state, and the most negative logic voltage level (as referred to as the *low* level) is defined to be the logical 0 state. *Negative* logic is just the opposite–high level is a 0 and low level is a 1. It is the preference of the designer to opt for positive or negative logic.

25·12-1. RESISTOR TRANSISTOR LOGIC (RTL) :

In this family logic is performed by resistors and transistors and we can realise inverters. NOR and NAND logic functions.

A typical NOR/NAND circuit is shown in fig. 31. The gating is performed by resistors, and, in this case, the gate itself (R_1, R_2, R_3 and R_4) is a negative OR a positive AND gate, depending upon the definition.

Operation : If all inputs are at zero (ground). then divider action of the resistors places a positive voltage on the base of *pnp* transistor and it is off. Thus the input is at $-V_{cc}$.

Now if any input goes negative, the transistor conducts and the output falls to ground. If more inputs go negative, the output remains at ground. Thus if *A* **OR** *B* **OR** *C* goes negative the output is at ground (NOR) Conversely. if *A* AND *B* AND *C* are at ground, the output is at $-V_{co}$ (NAND).

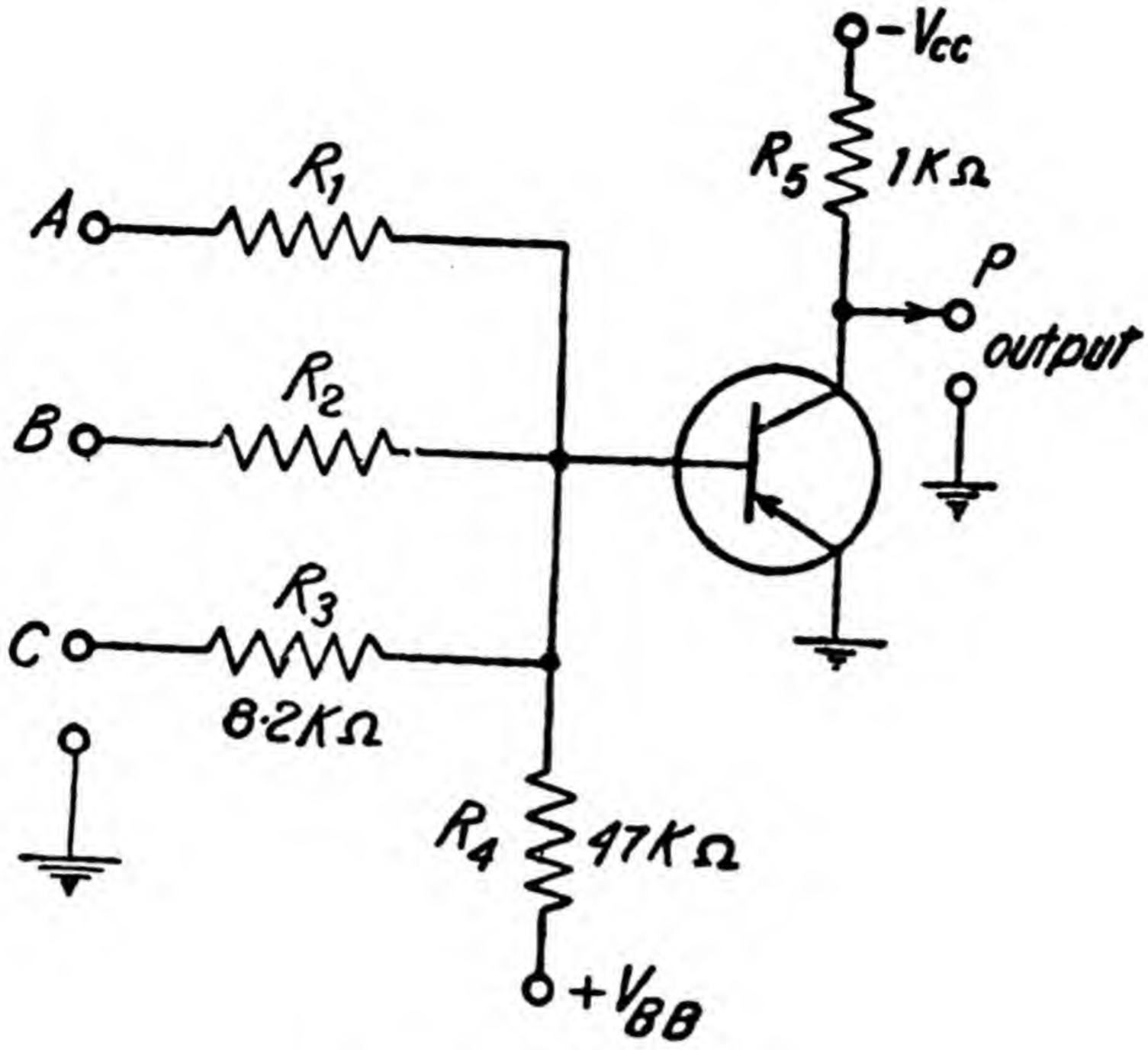


Fig 31. RTL circuit

Another RTL **NOR** gate is shown in fig. 32 using *npn* transistors. If input *A* is high, Q_1 will conduct and output at *P* will be grounded. Similarly if *B* is high, Q_2 will conduct and place the output at *P* to ground.

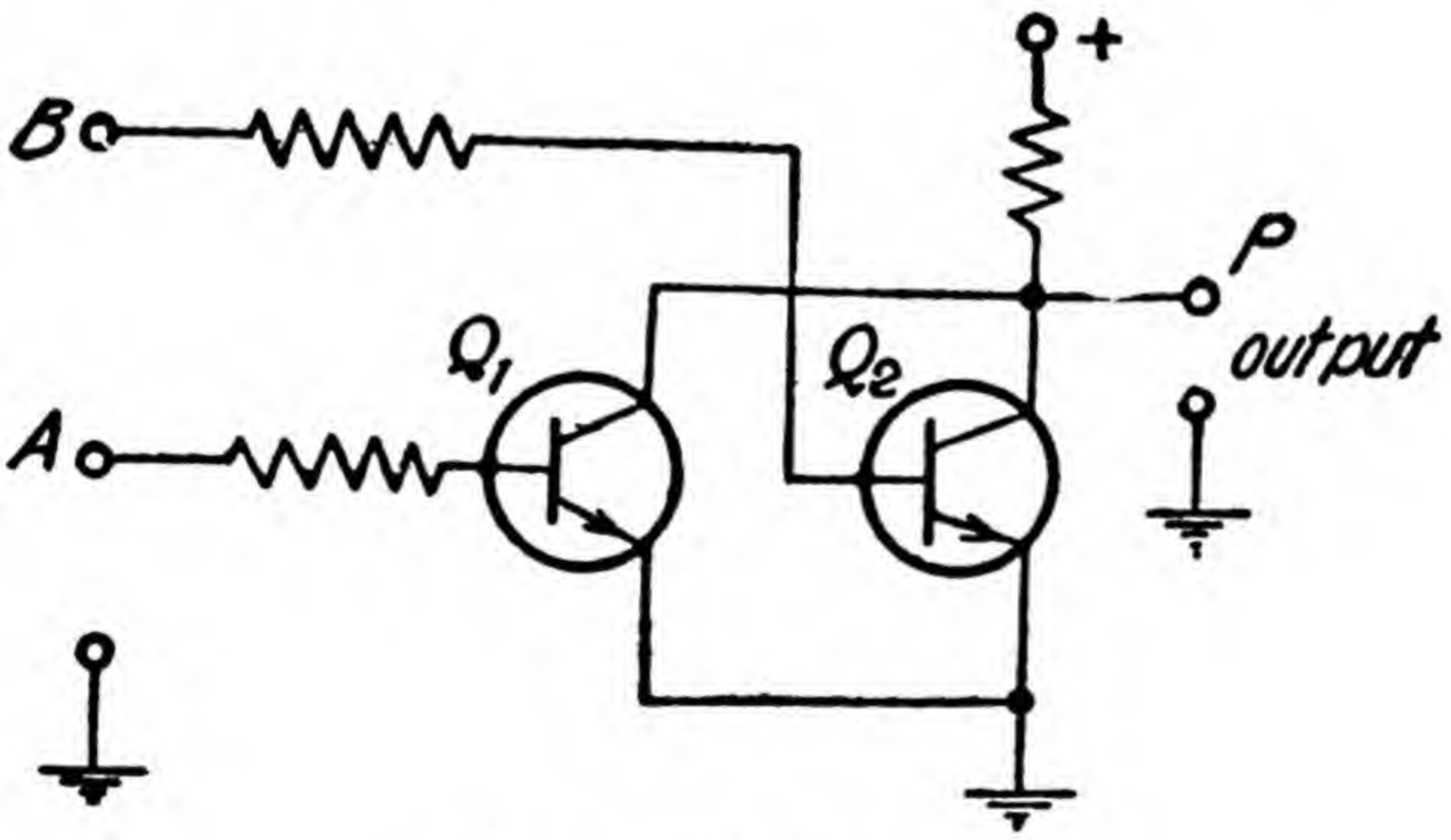


Fig. 32. RTL circuit.

Parameters : Referring to fig. 32, parameters of the logic circuit are :

Fan-out	4
Noise Margin	0·2 volts
Propagation Delay	12 *n* sec
Power Dissipation	30–100 mW
Power Supply voltage (V_{cc})	3·8 volts

Advantages and Disadvatages :

In RTL, resistors slow is the switching speed of the circuit. Thus RTL have low operating speed. Since there are only resistors to isolate one signal from another, there is a possibility of cross talk between inputs *i.e.* input *A* is fed a voltage change can be evident at input *B*. Consequently RTL has relatively poor noise immunity. The advantages of RTL is its low power dissipation.

25·12-2. DIODE-TRANSISTOR LOGIC (DTL) :

DTL logic circuits are shown below. DTL **AND** circuit is shown in fig 33, DIL **OR** circuit is shown in fig. 34 and DTL, **NAND circuit** is shown in fig. 35.

We assume positive logic inputs of $+5V$ for the logic state 1, and 0 V or ground for logic state 0.

DTL **AND Circuit** : Circuit is shown in fig. 33. If *any input*

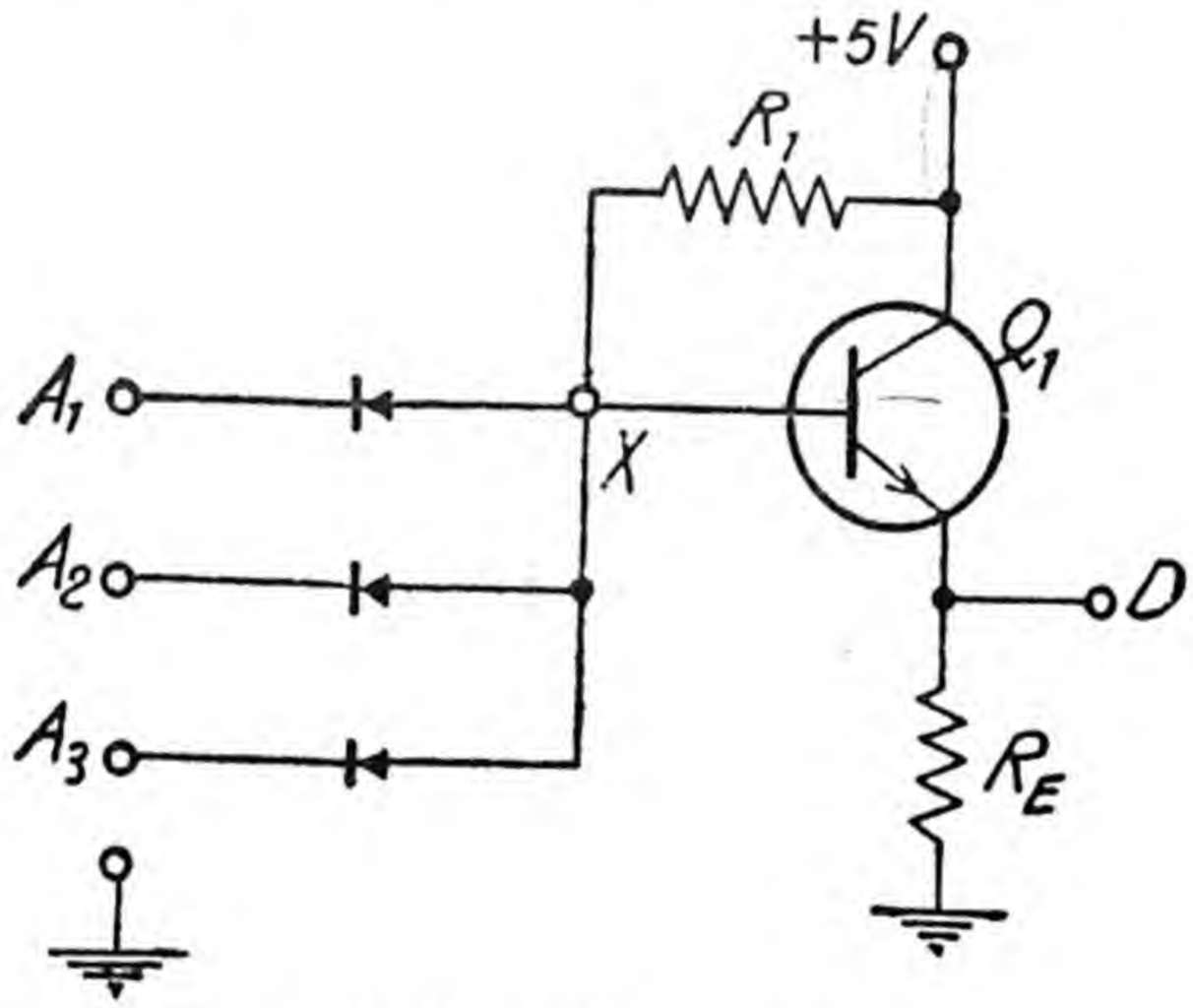


Fig. 33. DTL AND circuit.

is at ground or logic 0, its diode is forward biased through R_1. Consequently voltage at X is clamped to ground and the transistor is cut off. The output is low (at logic 0).

If *all inputs* are at logic 1 ($+5V$), diodes do not conduct and voltage at X is $+5V$ (logic 1 level). The transistor conducts heavily to saturation so that output is high (at logic 1).

DTL **OR Circuit** : Circuit is shown in fig. 34.

If *any input* is at logic 1 ($+5V$), its diode is conducting and $+5V$ appears at X. The transistor is in saturation and the output is high (at logic 1). If *all inputs* are at ground (logic 0), all diodes do not conduct and the voltage at X is zero. Consequently, the transistor is off and the output is low (at logic 0).

DTL **NAND** circuit :

Circuit is shown in fig 35. If any input is at ground (logic 0) its diode will be forward biased, voltage at X is clamped to ground and the transistor Q_1 is cut off. This prevents any base current being provided to Q_2. Thus Q_2 will cut off, leaving point D high (logic state).

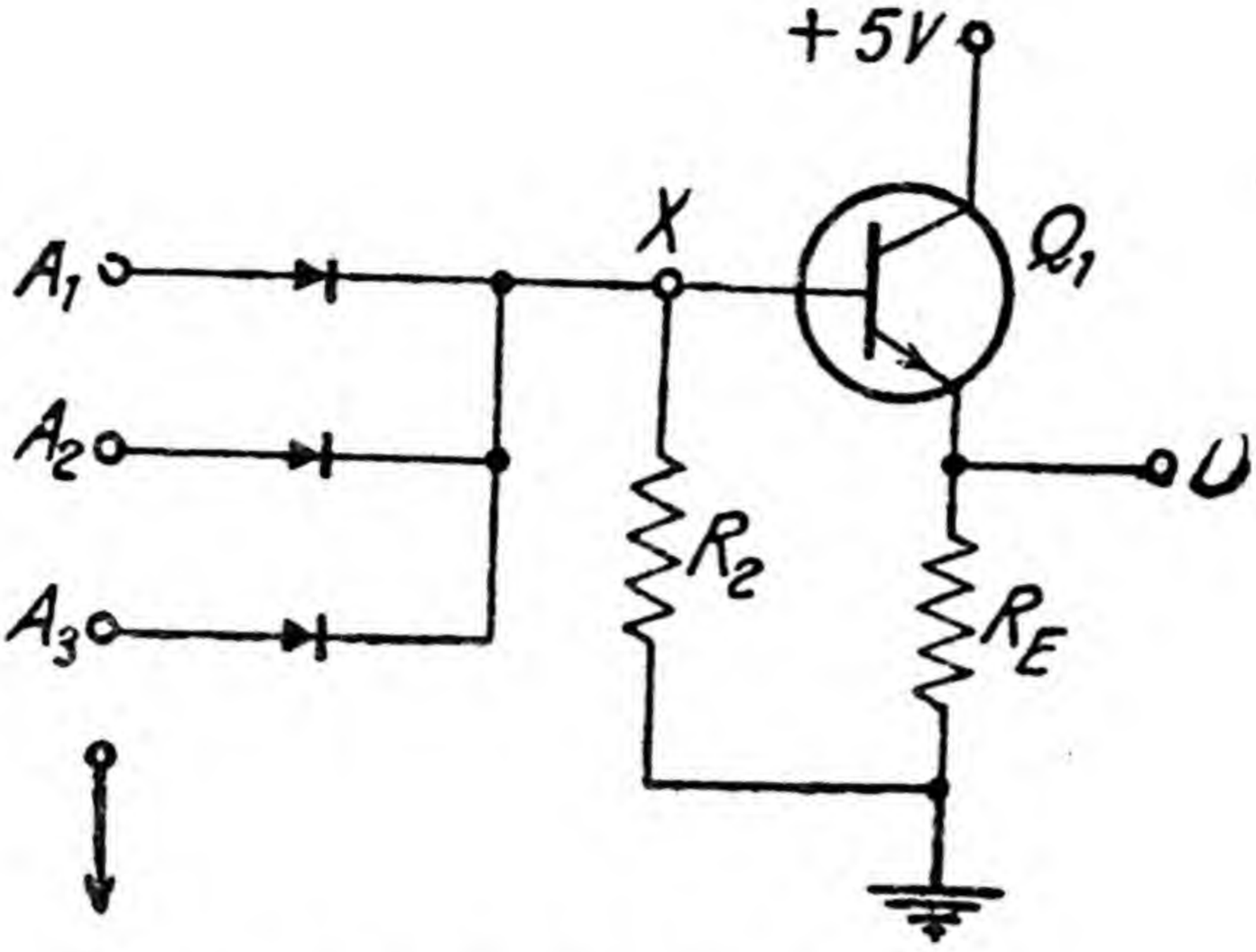


Fig. 34. DTL OR circuit.

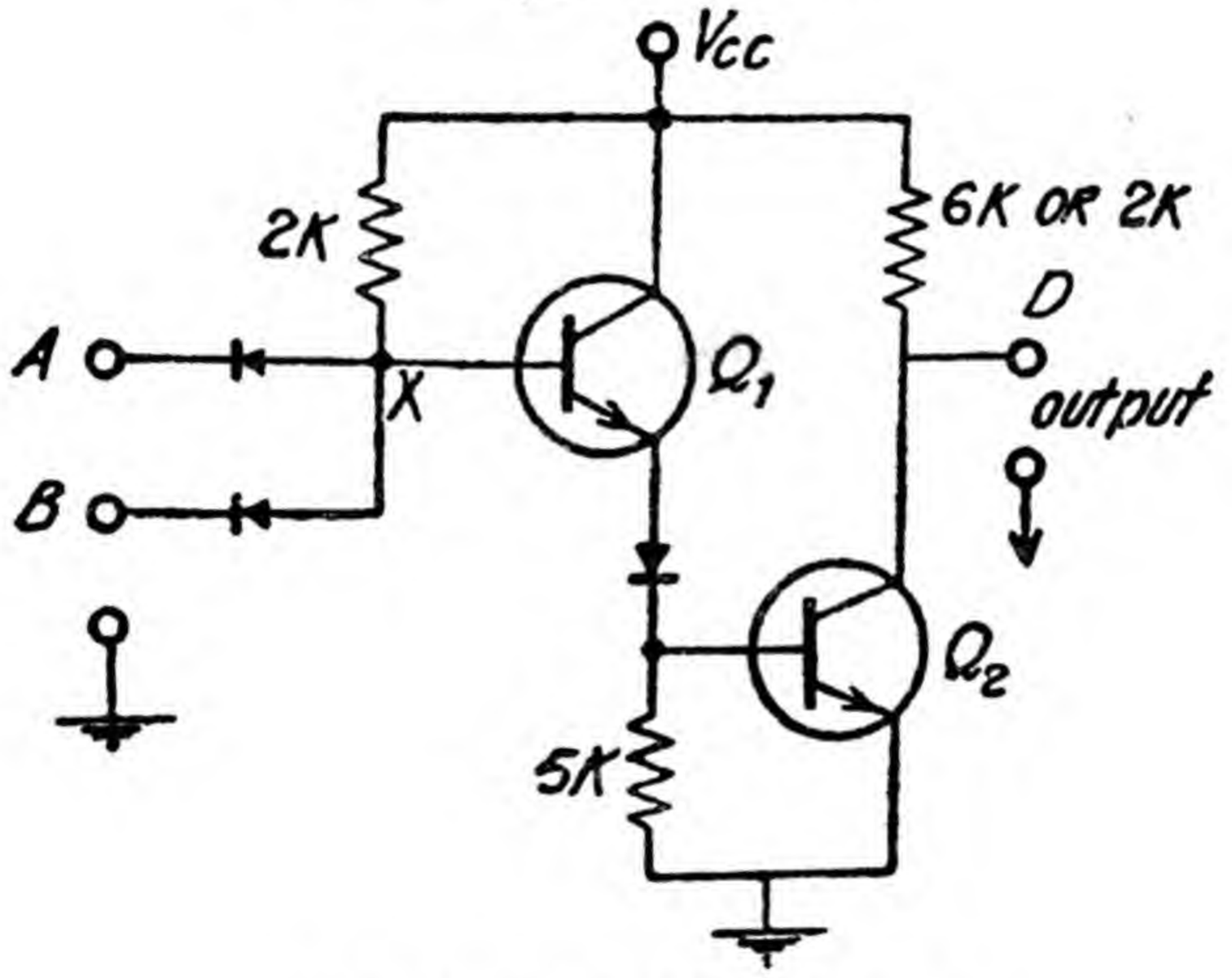


Fig. 35. DTL NAND circuit.

Parameters :

Fan out	8
Noise Margin	0·7 V
Propagation Delay	30 *n* sec.
Power Dissipation	60 *m* W
Power Supply Voltage (V_{cc})	+5 V

Advantages and Disadvantages :

The DTL circuit switches faster than the RTL circuit because the signal passes through the low forward resistance of the diodes to the transistor. A fan out as high as 8 is possible because of the high input impedance of the subsequent gates in the 1 state. Use of diodes rather than resistors and capacitors makes the DTL circuit more economical in integrated circuit form. It has low power dissipation.

Disadvantages of this circuit is its poor noise margin and low operating speed.

25·12-3. TRANSISTOR TRANSISTOR LOGIC (TTL) :

A TTL NAND gate is shown in fig. 36. It uses a multiple-emitter configuration for an input transistor inplace of diodes of DTL, giving greater speed of switching.

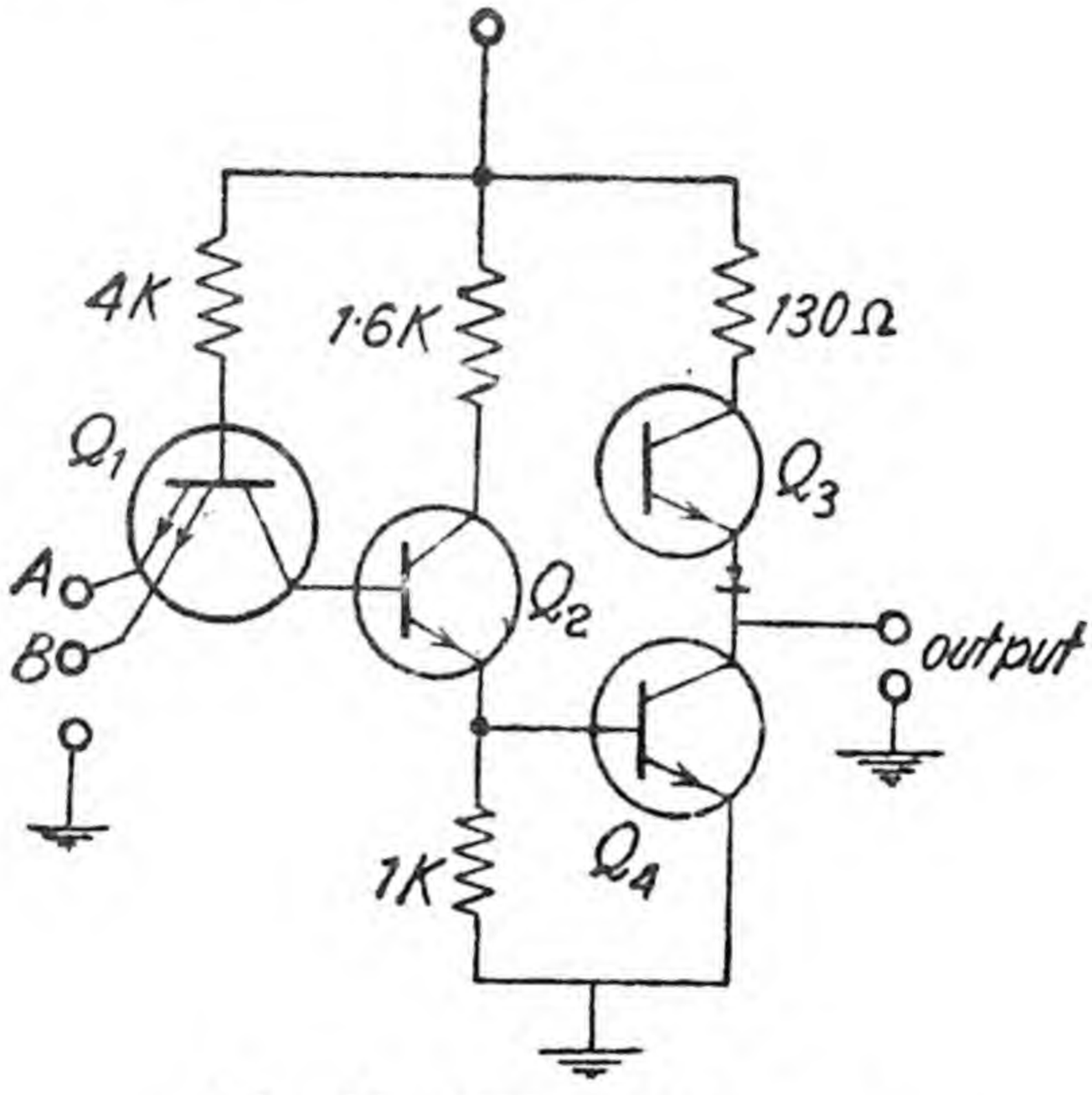


Fig. 36. TTL NAND circuit.

Operation :

If *both A* and *B* are *high* (logical 1) then Q_1 has no emitter current, but its base-collector junction is forward biased and therefore it will supply a base current to Q_2 which, in turn, feeds base current to Q_4, causing it to conduct. We get an output. Note that collector of Q_2 goes low and cuts off Q_3.

If *either A* or *B* goes *low* (logical 0), then Q_1 will have base-emitter current placing Q_2 to ground and thus Q_2 is cut off. Consequently, collector of Q_2 will go high and cause Q_3 to conduct and Q_4 to cut off. Thus we again get output.

Thus we get an output in both the cases.

Parameters :

Fan out	10
Noise Margin	0·4 V
Propagation Delay	7·10 *n* sec.
Power Dissipation	100 *m* W
Power Supply Voltage	+5 V

Advantages :

The multiple-emitter transistor replaces combinations of

diodes, resistors, and transistors found in other logic circuits. Thus its geometrical size is small. Smaller size yields lower costs or more functions or both per given IC chip. Other than low cost advantage, it has high operating speed, and high fan out.

25 12-4. EMITTER COUPLED LOGIC (ECL) :

This logic is designed for fast switching speed. It is the only family that does not operate fully saturated or cut off. Output levels are $-0{\cdot}8$ V for logical 1 and $-1{\cdot}8$ V for logical 0. A basic emitter-coupled logic gate is shown in fig. 37.

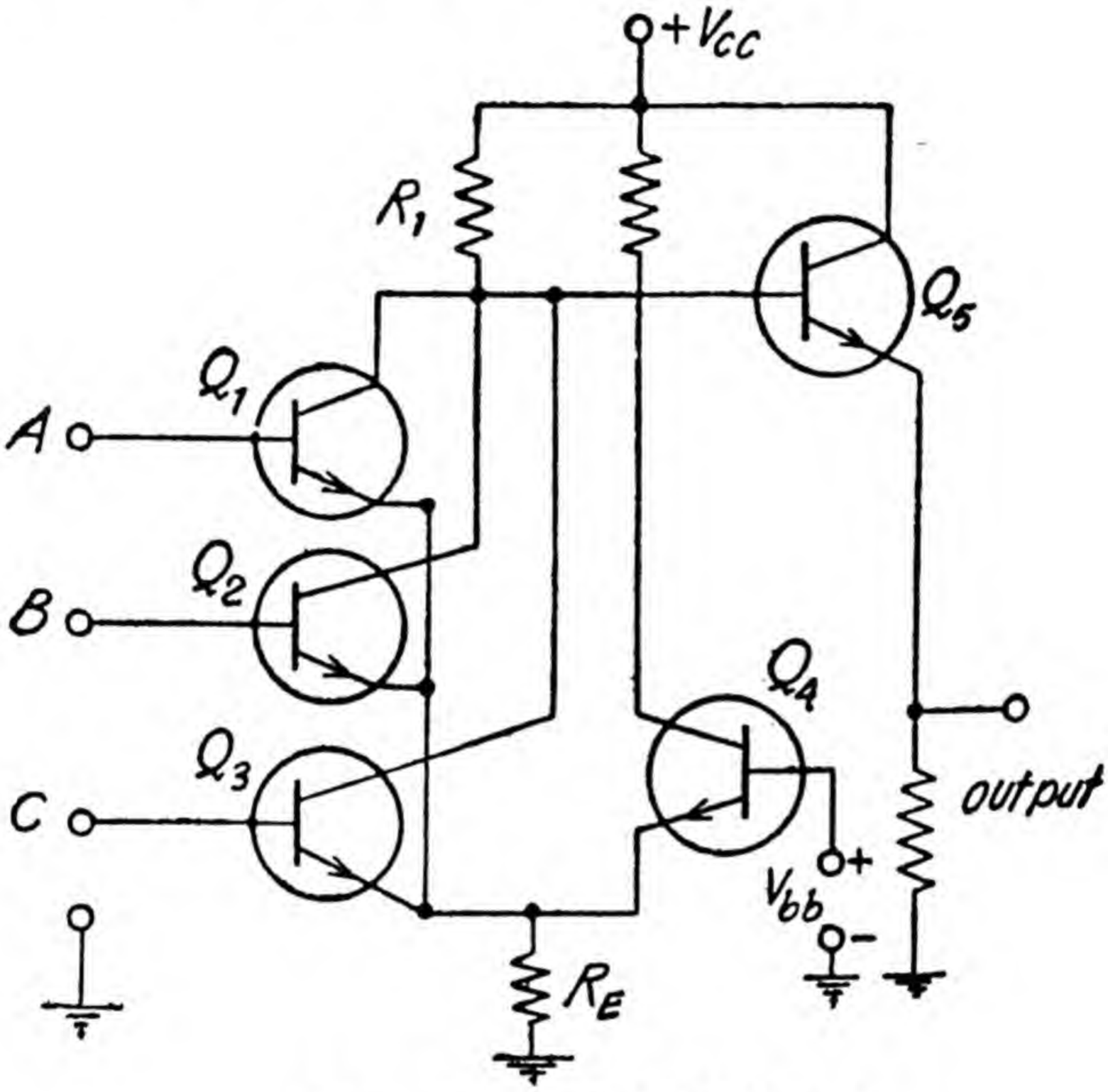


Fig. 37. Basic ECL Circuit.

The emitters of logic transistor Q_1, Q_2 and Q_3 are coupled to the emitter of a reference transistor Q_4. The common-emitter resistor R_E is high and acts as a constant-current source. The base of Q_4 is connected to a reference voltage V_{bb}.

Operation :

If all the inputs are near ground potential (logical 0), Q_1, Q_2 and Q_3 are cut off. No current flows through R_1 and the common collector potential rises toward V_{cc}. This drives Q_5 into conduction. Consequently the output from the emitter of Q_5 goes positive to give a logical 1 output.

If one of the inputs is made positive and higher than the reference voltage V_{bb} (logical 1) then current will flow through the associated transistor of that input. The collector potential will then fall, causing a fall in output from the emitter of Q_5. This gives a logical 0 output.

Advantages :

Note that emitter coupling prevents the transistors from going into saturation because, with R_E acting as constant current source, as current through logic transisto (*Q*1, *Q*2 or *Q*3) increases, current through reference transistor, *Q*4, decreases. This prevention from going into saturation imparts a very fast switching speed to the circuit, of the order of only a few nanoseconds. Power dissipation is relatively high, typically 50 *mW*. The use of an emitter follower output circuit gives a very low output impedance, allowing a very high fan out of upto 25.

Such a logic family is preferred only in the largest computers where many disadvantages can be suffered for the sake of high operating speed.

25·13. SEQUENTIAL CIRCUITS :

The five basic logic gates AND, OR, INVERTER, NAND, NOR) are classified as *combinational logic circuits*. A combinational logic circuit is a logic circuit without feedback. The output of such a circuit depends *only on its external inputs*. A *sequential logic circuit* is a logic circuit with feedback. Its output depends on the external inputs *as well as the present state of its outputs*, which are fedback.

Combinational logic circuits have no memory and can operate as fast as the devices of which they are made whereas sequential logic circuits have memory and are generally made to operate at the speed of a *master clock* which provides *square wave timing pulses*. There are two types of sequential logic circuits : synchronous (timed or clocked) and asynchronous (untimed). In the forthcoming articles, we shall describe sequential circuits.

25·14. FLIP-FLOP (BISTABLE Or MEMORY CIRCUIT) : S-R FLIP-FLOP :

A bistable multivibrator or flip-flop has two stable states and can stay in one of the two possible states after an input has been applied. Since the state does not change following the removal of the input, the flip flop is a 1-*bit memory* or storage device.

Let us describe a *SR* bistable multivibrator circuit. Its operation that follows is asynchronous. The way in which it is labelled is shown in fig. 38. This is most basic and useful sequential logic circuit, consisting two cross-coupled NOR gates (can also have two cross-coupled NAND gates). It has two inputs : the *S* (set) and *R* (reset) inputs and two outputs, which are complementary to each other, denoted by Q and $\bar{Q}$.

Since sequential logic circuits have feedback, the output signals (on logic lines Q and $\bar{Q}$) not only depend on the external input

signals (on logic lines S and R) but also on the *initial output*

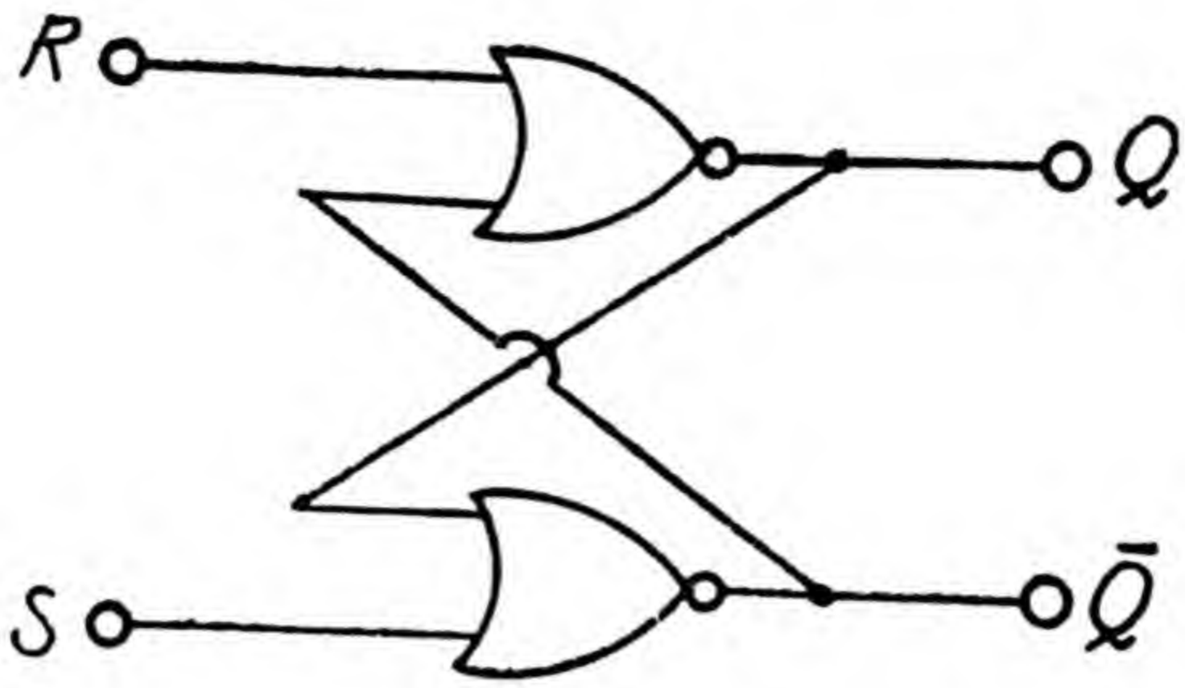


Fig. 38. S–R latch logic circuit using two cross coupled NOR gates. R–reset, S–set inputs. Q and $\bar{Q}$ outputs.

signals (on logic lines Q and $\bar{Q}$ themselves). The initial signals on logic lines Q and $\bar{Q}$ are referred to as the *present state signals* or simply the *present state*. After S OR R or *both* S AND R change, the signals that appear on logic lines Q and $\bar{Q}$ are called the *next state signals* or simply the *next state*. We shall use Q_n and Q_{n+1} to denote the present and next state of the flip flop. Since it may store one bit of information (either $Q=0$ or $Q=1$) it is called as 1-bit memory, and since this information is locked or latched, this flip flop is also called S–R latch. An S–R latch is called an asynchronous circuit since its inputs are not timed or synchronised.

S–R latch circuit (fig. 38) has four different input possibilities (but only two states, high and low for output logic line Q or $\bar{Q}$). The outputs for these four possible inputs are being described below :

(i) As long as both input lines S and R carry 0 signals, the flip flop remains in the same state, *i.e.*, for $S=0$ and $R=0$, value of Q is unchanged and $Q_{n+1}=Q_n$.

(ii) A signal on the R line and *a* 0 signal on the S line cause the flip flop to *reset* to the 0 state, *i.e.*, for $R=1$, and $S=0$, regardless of the values of Q_n, Q_{n+1} will be equal to 0.

(iii) For inputs $S=1$ and $R=0$, regardless of the values of Q_n, Q_{n+1} will be equal to 1.

(iv) Inputs $S=1$ and $R=1$ at the same time is forbidden. If this occurs flip flop can go to either state because input is telling the flip flop to go to both *set* and *reset* at the same time. Thus for $S=1$, $R=1$, $Q_{n+1}=\bar{Q}_{n+1}=0$ or 1 which is not allowed. We can put above output/input correspondence in a truth table (for fig. 38) as follows :

S	R	Q	$\bar{Q}$	
0	0	Q	$\bar{Q}$	← state is the same as before
0	1	0	1	
1	0	1	0	
1	1	ND	ND	← Not defined

Thus we conclude that :

(i) S–R latch is a two state sequential circuit. For $R=0$ and $S=0$, the next state is equal to the present state of the circuit. Thus the next state of this circuit is a function of its present inputs and its present state.

(ii) S–R latch state time depends solely upon the delays of the two NOR gates and is thus asynchronous.

(iii) S–R latch is a memory element. To store a binary digit 0 we apply $S=0$, and $R=1$, and to store a binary digit 1 we apply $S=1$ and $R=0$. Thus information remains in the latch as long as the inputs R and S are kept at 0 and the power supply to the device is kept on.

The operation of SR latch is more easily followed by the circuit representation of fig. 38. This is shown in fig. 39.

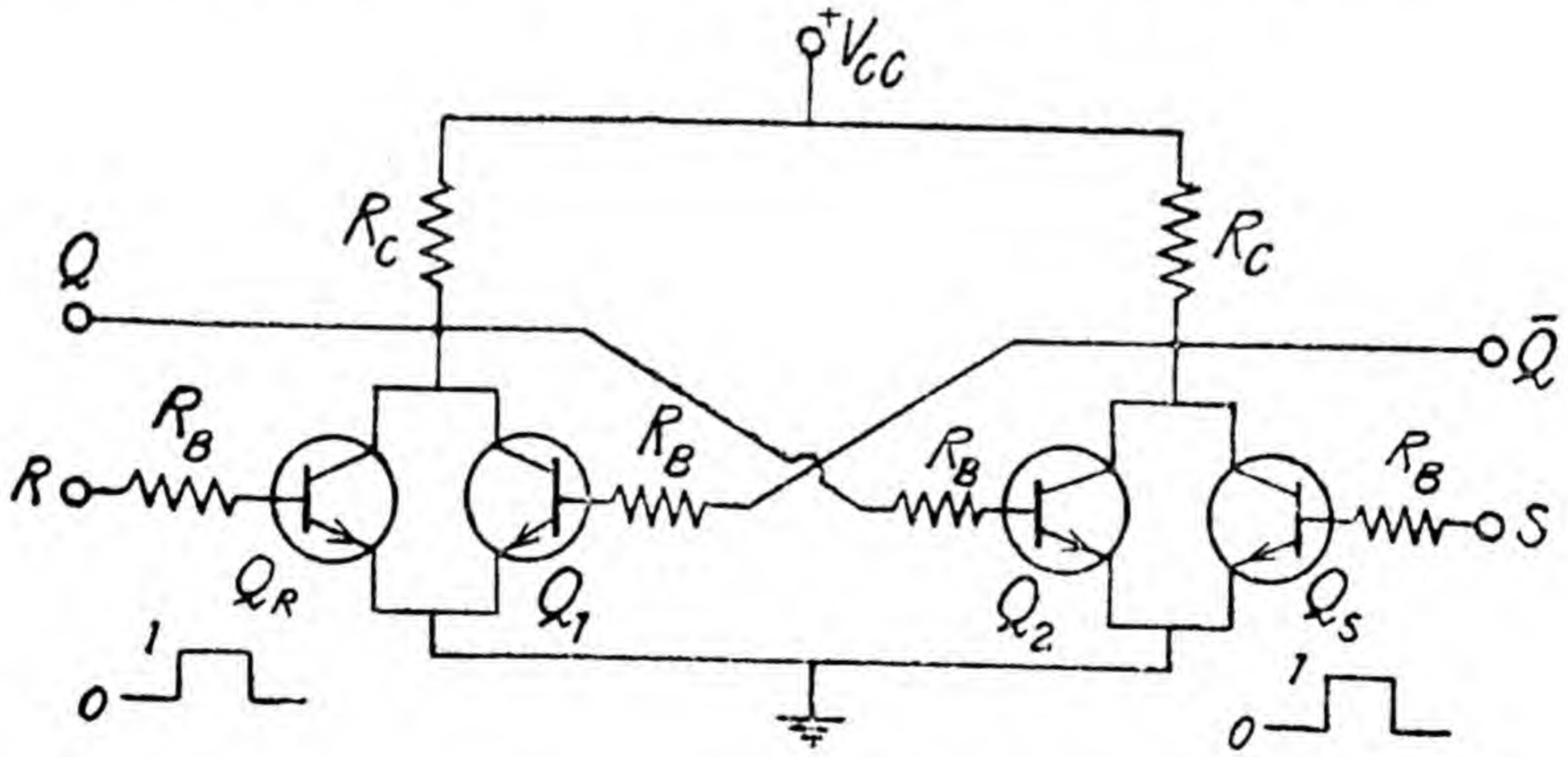


Fig. 39. Circuit representation of fig. 38.

To produce a logic 1 at the output terminal Q and a logic 0 at the output terminal $\bar{Q}$, it is necessary that Q_2 be conducting strongly and that Q_1 be cut off. To achieve this state, we can apply

a logic 1 (or positive) voltage to the terminal marked *S*. It will cause Q_s to go strongly into conduction and thereby bring the voltage at terminal $\bar{Q}$ essentially to zero. With a zero voltage at $\bar{Q}$, the base of Q_1 will not be receiving any current so that Q_1 will be cut off and its collector voltage will be essentially equal to $+V_{cc}$. This high voltage at Q will drive Q_2 into conduction so that state of Q_2 is same as that of Q_s.

To change the states, *i.e.*, $\bar{Q}$ at logic 1 and Q at logic 0, a positive voltage is applied at *R*. This will cause Q_R and Q_1 to conduct strongly while Q_s and Q_2 are cut off.

Assume Q at logic, 1 which implies that Q_R and Q_1 are off. But as Q is connected to the base of Q_2, $\bar{Q}$ is at logic, 0. If we now apply a logic, 1 pulse to *R*, Q_R saturates and Q changes to logic, 0. Due to cross coupling $\bar{Q}$ changes to logic, 1. (Thus when $R=1$, $S=0$, $Q=0$, $\bar{Q}=1$, see truth table). Now when *S* is set at a logic, 1, the output $\bar{Q}$ becomes logic, 0 and Q goes to logic, 1 (Thus when $S=1$, $R=0$, $Q=1$, $\bar{Q}=0$ see truth table).

Thus, a set input, $S=1$, yields Q at logic 1

a reset input, $R=1$, yields Q at logic 0.

Clocked Flip Flop : *S–R* latch discussed above is asynchronous circuit and acts as a short-term memory, remembering whether a set or reset pulse was last applied. Such circuits are not preferred in industry because they are generally harder to analyse and design. Sometimes we wish to control the operation of the set reset operations with a clock or timing pulse. The operation of a set (or reset) puise is delayed until a clock pulse appears. Such an operation is timed and is called synchronous switching. Readers will find any book on digital concepts us ful for detailed study in adcition to art. 25·16.

25·15. CLOCK PULSES (CP) :

In many applications, it is desired that, throughout the digital system, signals may move in synchronism. This is achieved by controlling the states of flip flop by clocked pulses or trigger pulses produced by a *masier clock* which is an accurately controlled oscillator producing a train of pulses similar to that shown in fig. 40.

Some features of clock pulse are :

(i) pulse duration is much smaller than the period of clock rate (333 *k c/s* for pulse in fig. 40),

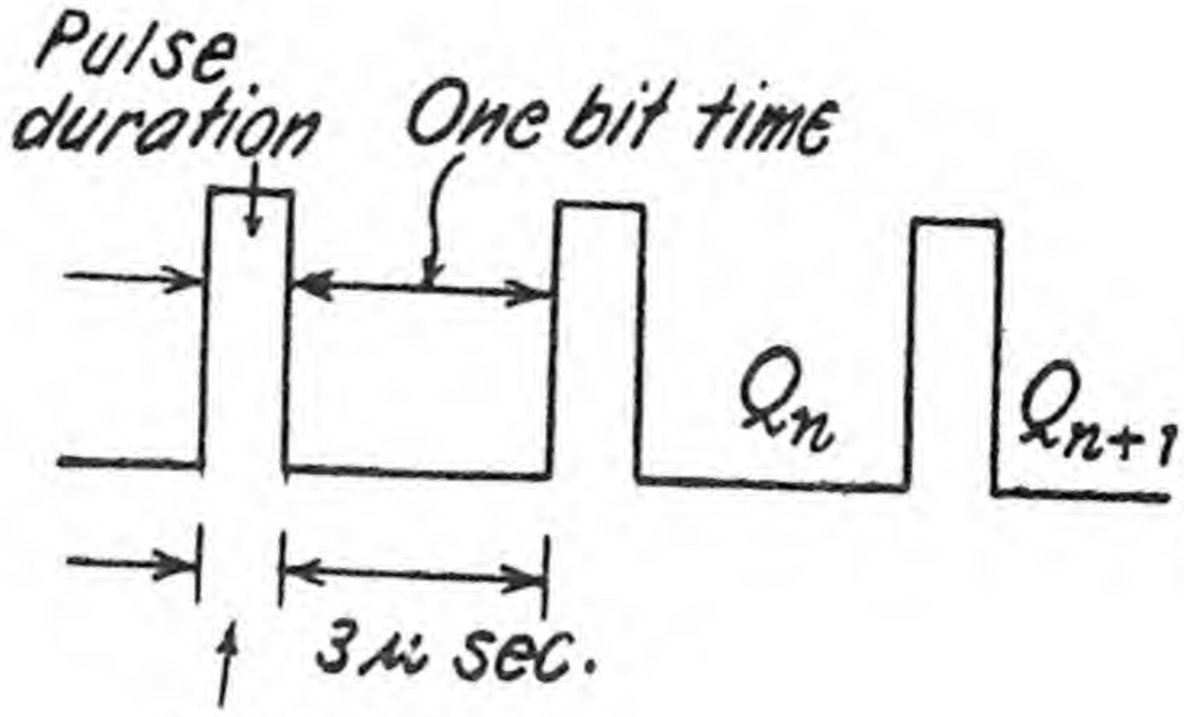


Fig. 40. A typical clock pulse (*CP*) waveform.

(ii) time between pulses is called *one bit time* and may vary from less than 1 μ sec. to over 15 μ sec. in computer involving many operations. This provides the time for certain operations to take place within the computer.

(iii) in clocked flip flop, the change of state on applying an input can take place only *during the application of a clock pulse* or change of state only occurrs on the rising and falling edge of the clock pulse. If, following one clock pulse, a flip flop output is designated by Q_n, the state following the next clock pulse is designated as Q_{n+1}.

25·16. RS/T FLIP FLOP Or CLOCKED FLIP FLOP :

In such a circuit, action of flip flop is controlled by pulses generated by a controlling oscillator elsewhere in the system called 'Clock.' These are called clock pulses or trigger pulses (CP).

Circuit is shown in fig. 41 (*a*). Suppose point S is at $+V_{CC}$, the supply voltage and R is at zero voltage (grounded) *i.e.* R is at logic 0 and S is at logic 1. When the clock pulse is applied then C_1R_1 circuit can be drawn as in fig. 41 (*b*). It acts as a differentiating circuit. At positive peak, the diode D_1 will not conduct because positive voltage is being applied to the cathode of the diode. But at negative peak diode will conduct and the negative pulse will reach the base of Q_1 so that Q_1 is *brought into cut off* while Q_2 is *turned on.*

Note that effect of clock pulse on S is different. At leading edge of the clock pulse voltage rises to $+Vcc$ and it simply counterbalances the $+Vcc$ voltage of point S and so no current flows in C_2R_2 circuit. At lagging edge of clock pulse, voltage falls to zero and negative pulse developed in C_2R_2 circuit simply brings the voltage at the junction of C_2R_2 close to zero but this does not produce sufficient voltage to turn D_2 on. Thus Q_2 is not affected by the incoming clock pulse.

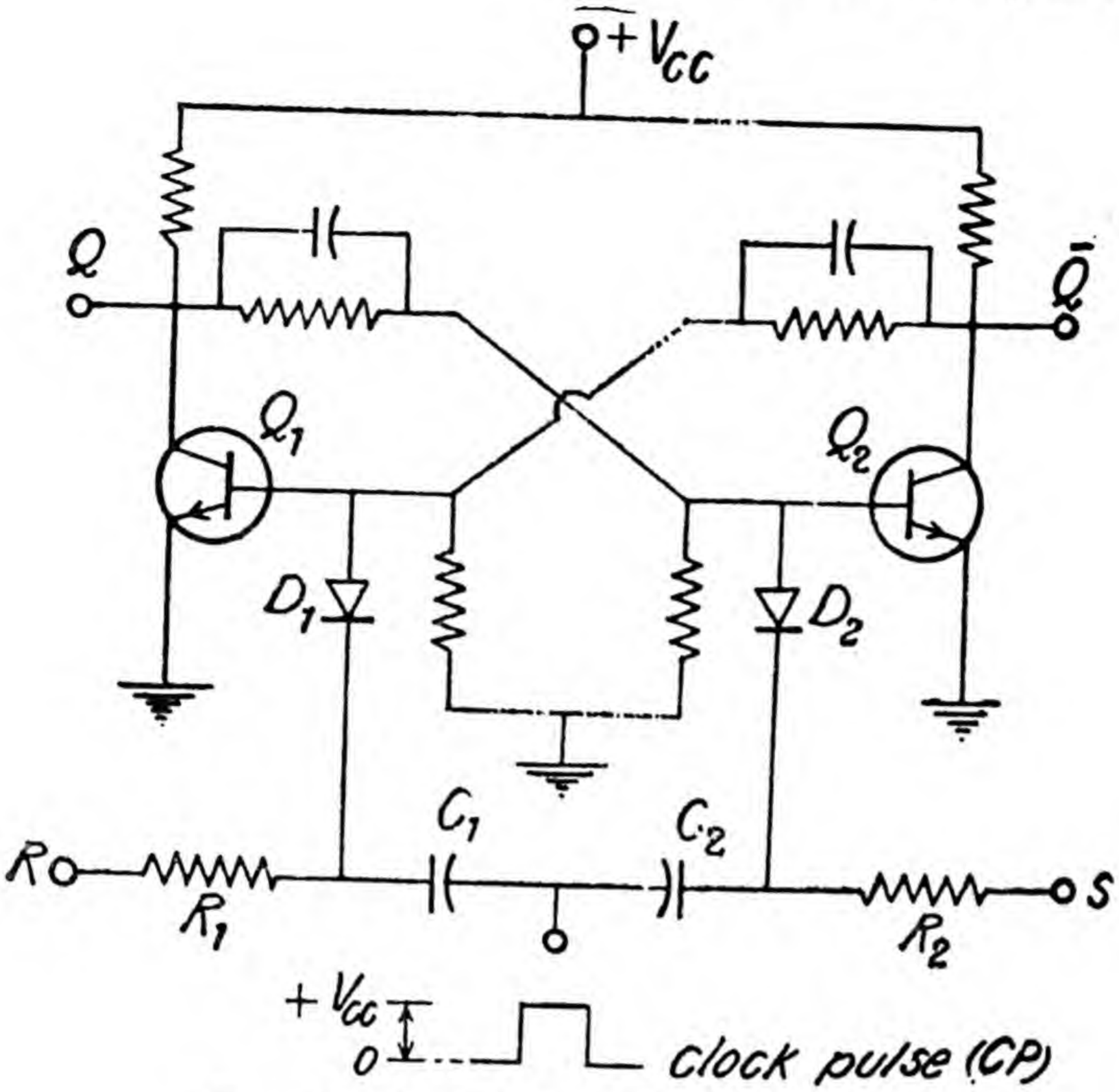


Fig. 41 (*a*) Clocked *RS* flip flop.

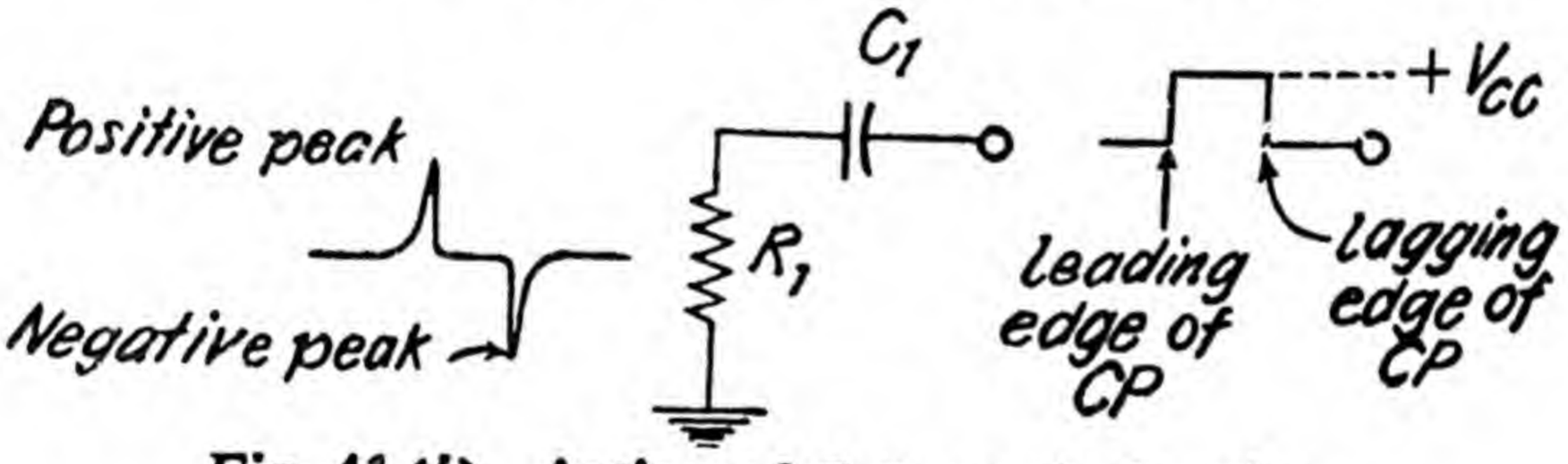
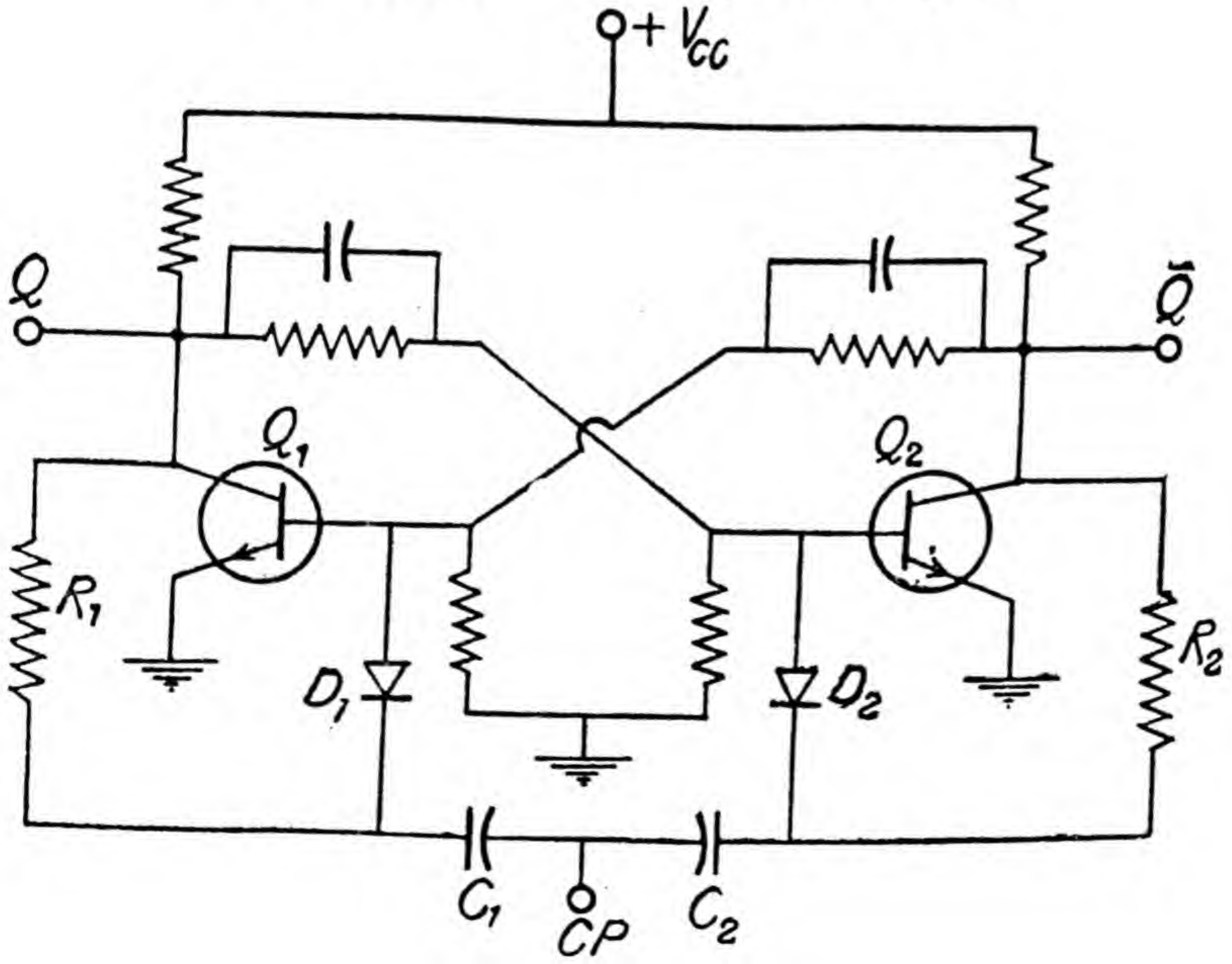


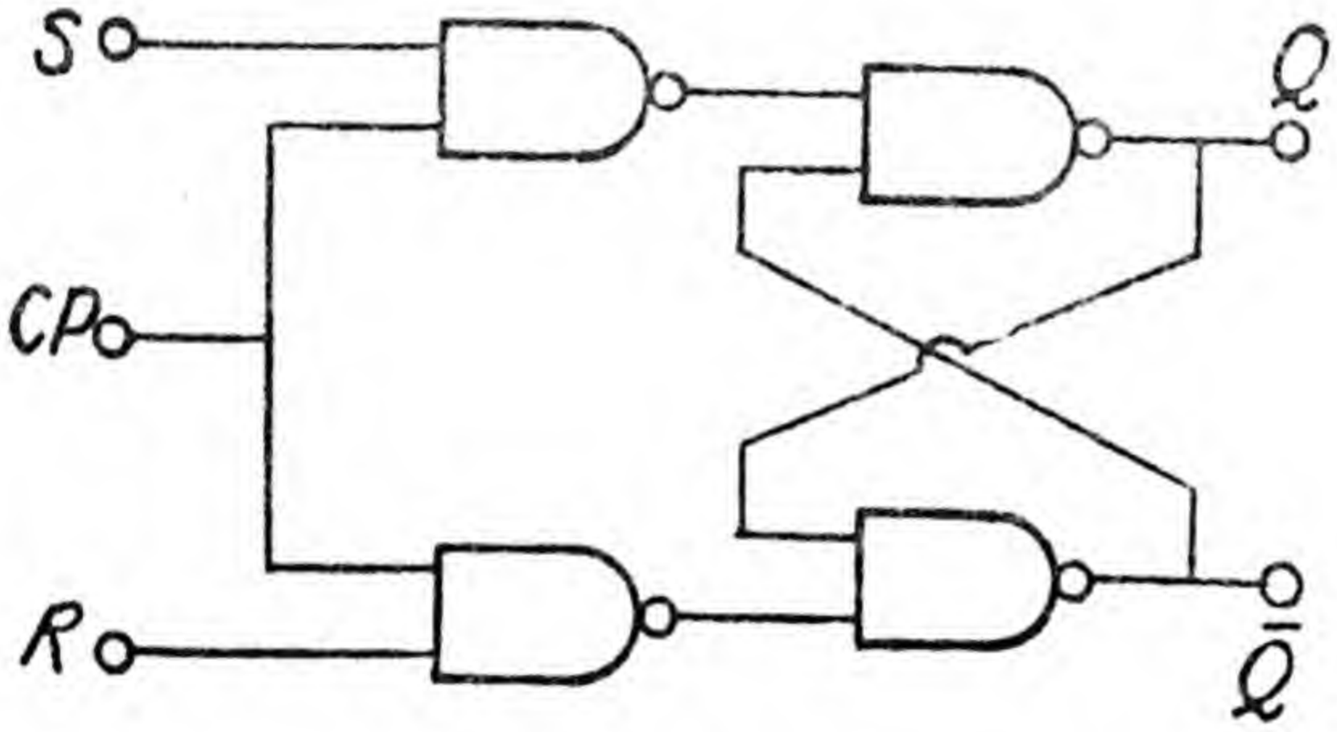
Fig. 41 (*b*) Action of C_1R_1 on clock pulse.



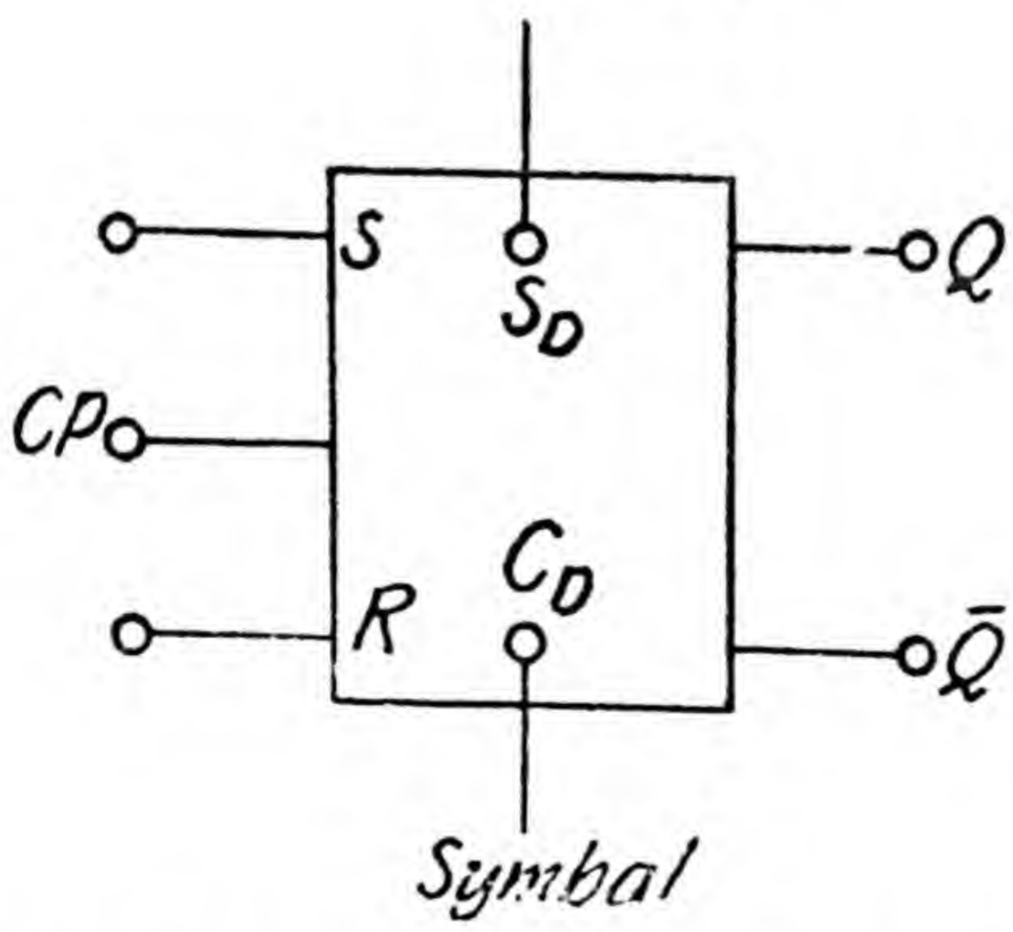
Suppose now voltages at R and S are reversed *i.e.* R is at logic 1 and S is at logic 0, then due to clocked pulse, Q_2 will be turned off while Q_1 will be brought into conduction.

To apply logic 0 to R and logic-1 to S or *vice versa*, we simply connect them to output terminals Q and $\bar{Q}$ because such voltages are already available at these terminals. Then circuit appears as shown in fig. 41 (*c*). Hence if we connect terminal S to $\bar{Q}$ and R to Q, the voltages at R and S terminals would be correct and all incoming clock pulses from clock generator will be directed to the proper transistor, determining which transistor will conduct or cut off. *Thus condition of output voltage at Q and $\bar{Q}$ is directly determined by the incoming clock pulses, and every time such a pulse is applied, there is a change in the voltage condition at these output terminals. Thus clock pulse controls the action of flip flop.*

A Gate Circuit : A functional block diagram of a clocked *R–S* flip flop is shown in fig. 41 (*d*), along with its symbol. The truth table is given in fig. 41 (*e*). Its circuit consists of a NAND



Circuit



Symbal

Fig. 41 (*d*) Clocked *R–S* flip flop with symbol.

R	S	Q_{n+1}
0	0	Q_n
0	1	1
1	0	0
1	1	?

Fig. 41 (*e*) Truth table of clocked *R-S* flip flop.

gate *R-S* flip flop whose input are gated (steered) by the clock pulses fed into two additional NAND gates. Due to inversion at input gates, Q now appears on the 'same line' as S while $\bar{Q}$ on the 'same line' as R (in contrast to *R-S* flip flop. See fig. 38).

In its symbol, two additional inputs (not shown in the circuit) called direct set (S_D) and direct clear (C_D) are shown. They permit direct setting and clearing of the flip flop *without regard to clock puls input*. If it is desired to clear the latch *i.e.*, the output be $Q=0$ when $CP=0$, then operation is performed by programming the clear input (C_D) to 0 and preset input (S_D) to 1 (that is, $C_D=0$. $S_D=1$ and $CP=0$). Similarly, if it is desired to preset the latch into 1 state then operation is performed by programming $S_D=0$, $C_D=1$ and $CP=0$. The S_D and C_D inputs are called direct or *asynchronou inputs* as they are not in synchronism with the clock but may be applied at any time in between clock pulses. Once the state of flip flop is established (asynchronously) the direct input can be maintained at $S_D=1$ and $C_D=1$ before the next pulse arrives inorder to 'enable' the flip flop.

The transitions of the states of the clocked *R-S* flip flop can be described by its truth table. The condition $R=S=1$ results in undeterminate state. The reason is as foilows. With inputs $R=S=1$, the outputs of the first two NAND gates will be 0 when clock pulse is applied. Consequently, both outputs from the second two NAND gates will be 1, indicating cut off in the output transistors. Now after the passage of clock pulse, both of the output transistors will try to go into conduction and since due to circuit imbalance one transistor will go into saturation faster than the other, the resultant state will remain indeterminate. Thus a condition $R=S=1$ must be avoided in *R-S* flip flops. A method to avoid this ambiguity of output for 1-1 input is found in *J-K* flip flop.

25·17. J-K FLIP FLOP :

Fig. 42 (*a*) shows the block diagram, symbol and truth table of a *J-K* flip flop. It is simply an *R-S* flip flop with two AND gates at the inputs.

We note that there are cross connections from output to *J-K* inputs. $\bar{Q}$ is connected to J input while Q is connected to K input. Its effect is that it provides a means of steering 1-1 input to the necessary S or R terminal to provide a change of state at the output. For example, suppose before CP is applied

$$Q=1, \quad \text{and} \quad \bar{Q}=0$$

and let 1 – 1 input at J and K

then, as they are AND gates, at the input of

J gate is 1, 0

K gate is 1, 1

so that at outputs of these gates,

S will 0 (low)

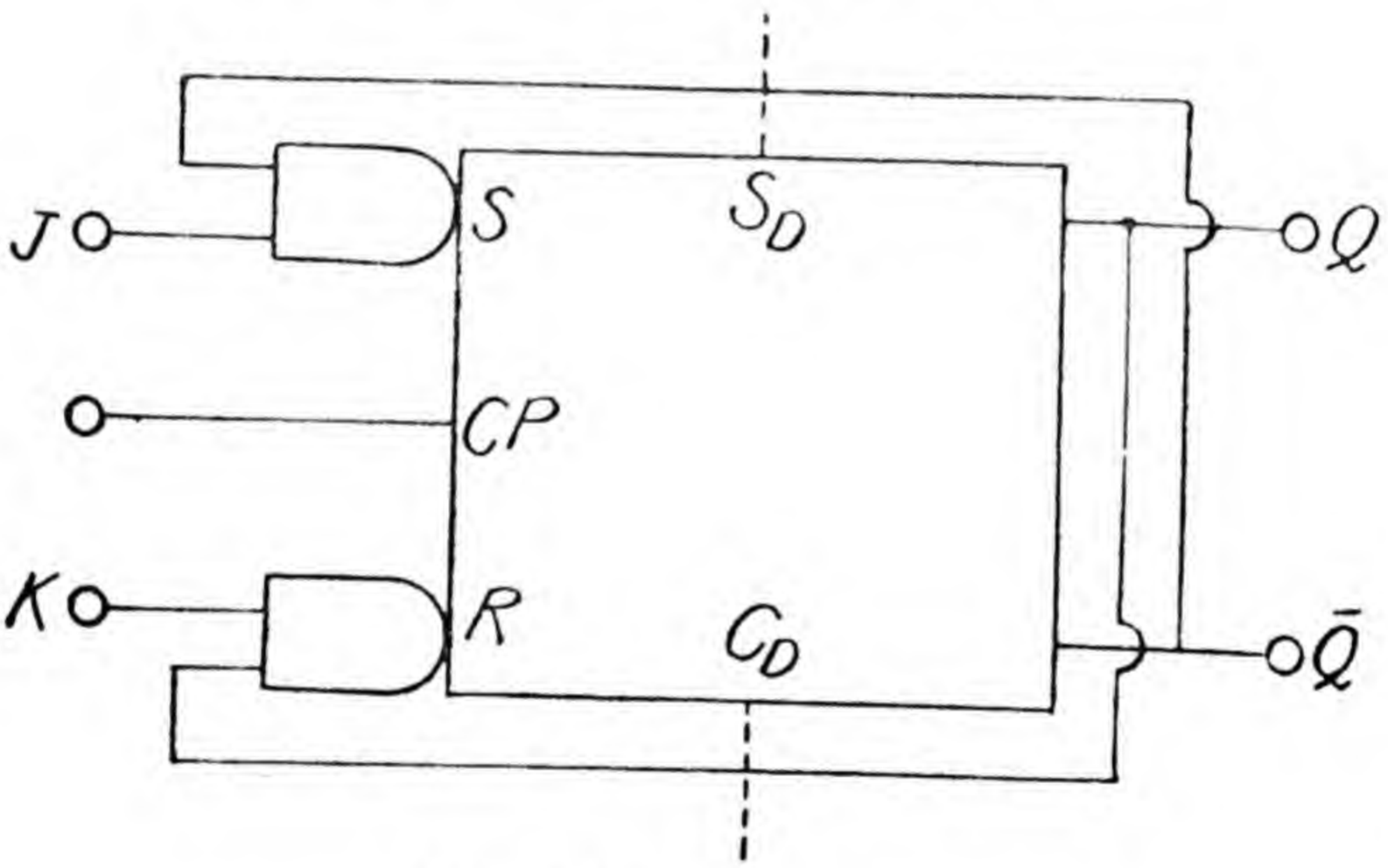
R will 1 (high)

and when the next clock pulse arrives, this will make the outputs of flip flop

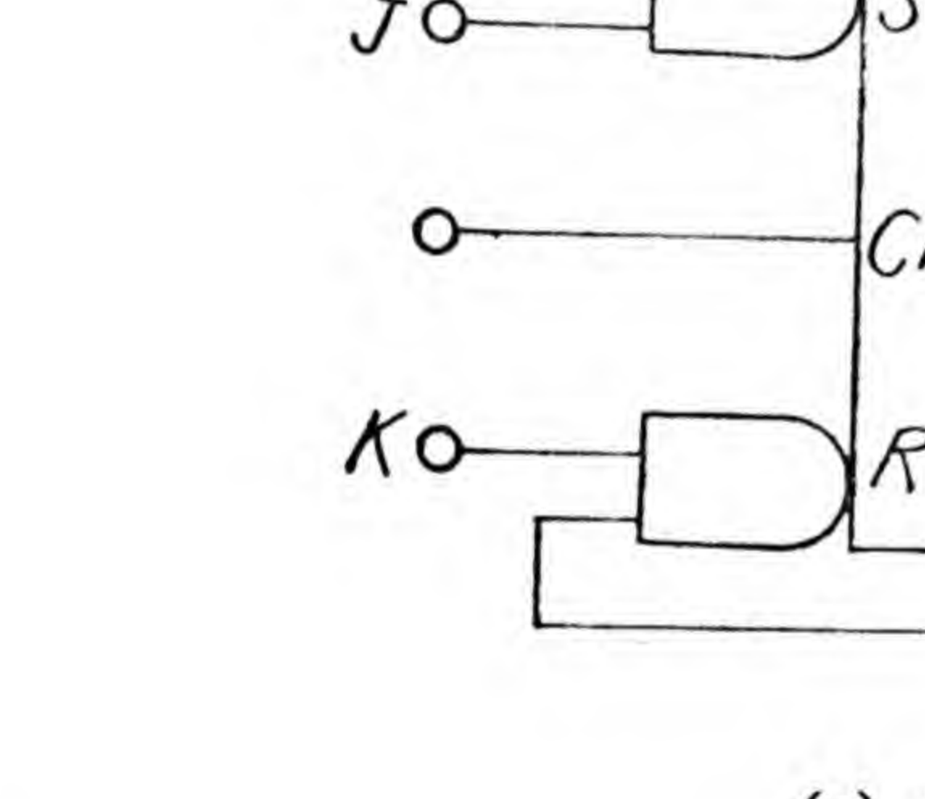
$$Q=0$$

and $$\bar{Q}=1$$

Thus a change of state has occurred. It means state will change each time CP appears if $J=K=1$, *i.e.* flip flop *toggles* from one state to the other at each CP. This property makes it very useful in counting applications. Note that in the symbol, fig. 42 (*b*), we have used T in place of CP to show this property.



(*a*) Block diagram.



(*b*) Symbol.

J	K	Q_{n+1}
0	0	Q_n
0	1	0
1	0	1
1	1	$\bar{Q}_n$

(*c*) Truth table.

Fig. 42. *J–K* Flip flop.

25·18. MASTER SLAVE J-K FLIP FLOP :

In some clocked *J-K* flip flops, there may still be some ambiguity in the output when $J=K=1$ input is applied. This arises due to a *race around* condition in which the change in state of signals Q $\bar{Q}$ fed back from output to input of AND gates on applying *CP*, may cause an oscillatory condition for the *duration* of the clock pulse (*CP*). Therefore when *CP* ends, *outputs may be quite arbitrary, depending upon the propagation delay through the flip-flop and the duration of CP.*

A circuit to aviod this race around condition is called a *J-K* master slave flip flop. Its simplified diagram is shown in fig. 43, while a bit descriptive in fig. 44.

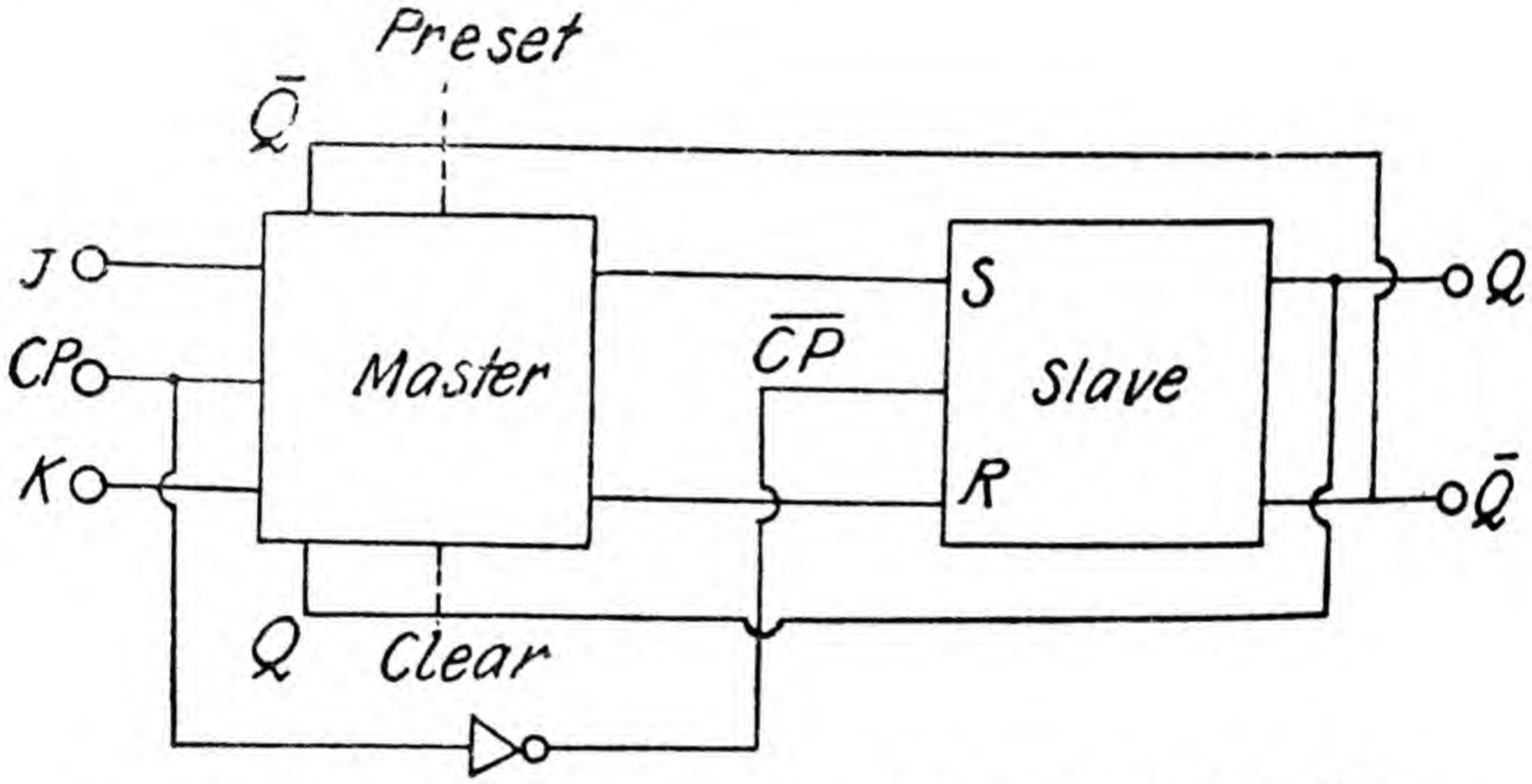


Fig. 43. Simplified block diagram of *J-K* master slave flip-flop.

As is obvious from the figure 44, at the input of the slave, trigger (*CP*) appears after inversion. As $T=1$, $\bar{T}$ will be 0. Therefore slave *R-S* flip flop cannot change its state (inhibited) and Q_n is *invariant for pulse duration.* This is the situation when master is enabled. But when trigger pulse passes, $T=0$ so that master is inhibited and as $\bar{T}=1$, slave will be enabled. Since slave is a $R\,S$ flip flop, its operation follows *R·S* flip flop truth table.

$$\left.\begin{array}{l}\text{If} \quad S=Q_m=1 \text{ and } R=\bar{Q}_m=0\\ \text{then} \quad Q=1 \quad \text{and} \quad \bar{Q}=0\end{array}\right\}$$

$$\left.\begin{array}{l}\text{If} \quad S=Q_m=0 \text{ and } R=\bar{Q}_m=1\\ \text{then} \quad Q=0 \quad \text{and} \quad \bar{Q}=1\end{array}\right\}$$

Thus in the interval between clock pulses, the value of Q_m transferred to the output Q (or $\bar{Q}_m$ is transferred as $\bar{Q}$). Erratic beha-

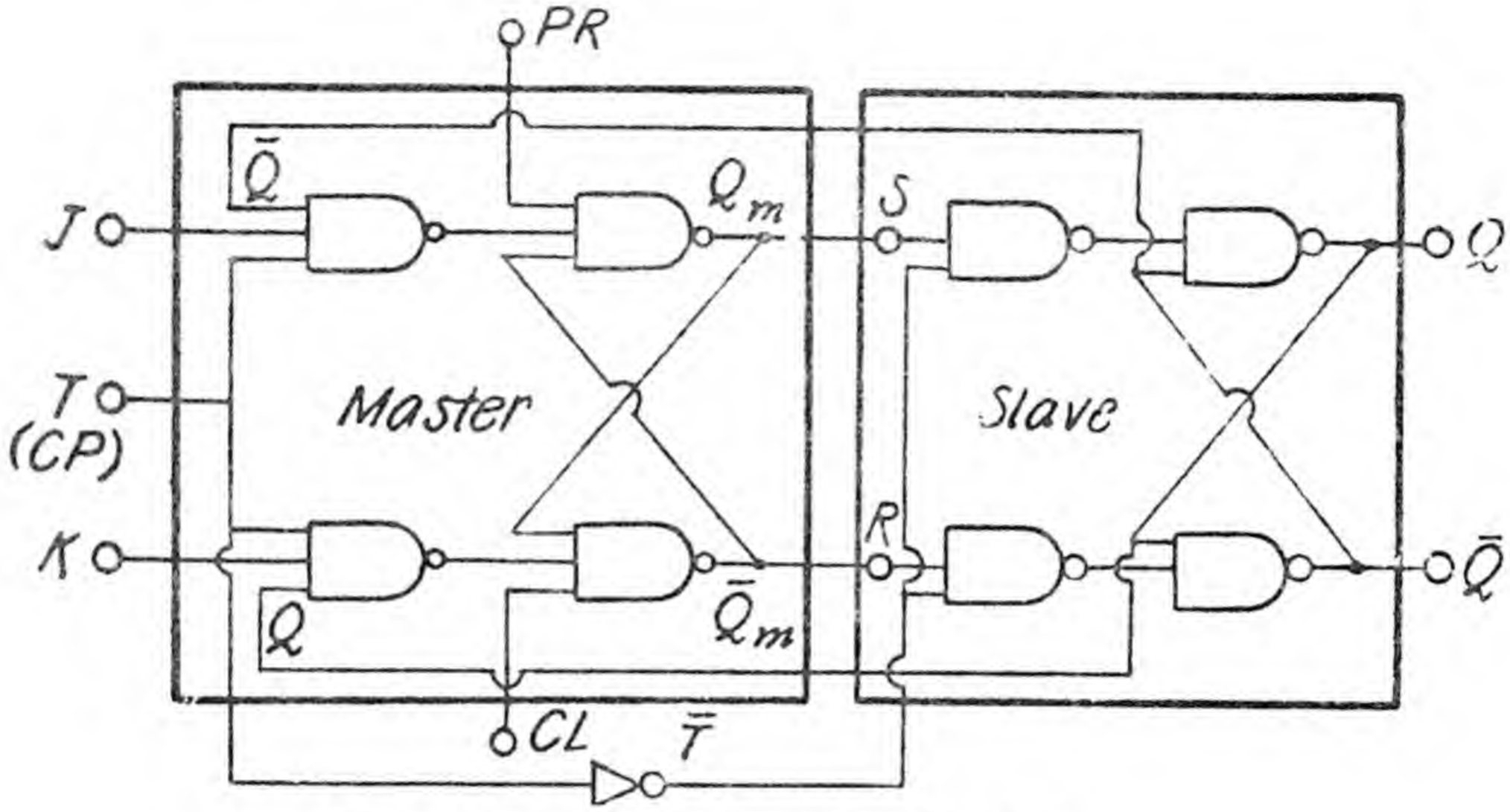


Fig. 44. *J-K* Master slave flip flop.

viour does not occur during the clock pulse, because the cross connected feedback is from the output of the slave to the input of the master. *Since the outputs* Q and $\bar{Q}$ *do not change during the clock-pulse, no oscillations can take place.* Thus possibility of arbitrary outputs when *CP* ends are removed.

The D type Flip-Flop :In this type of flip-flop, there is no possibilty of ambiguous state. This is achieved by providing *J-K* flip flop with an inverter on one of the input lines. Thus only one data input (*D*) to the flip flop is necessary. Its symbol is shown in fig. 45.

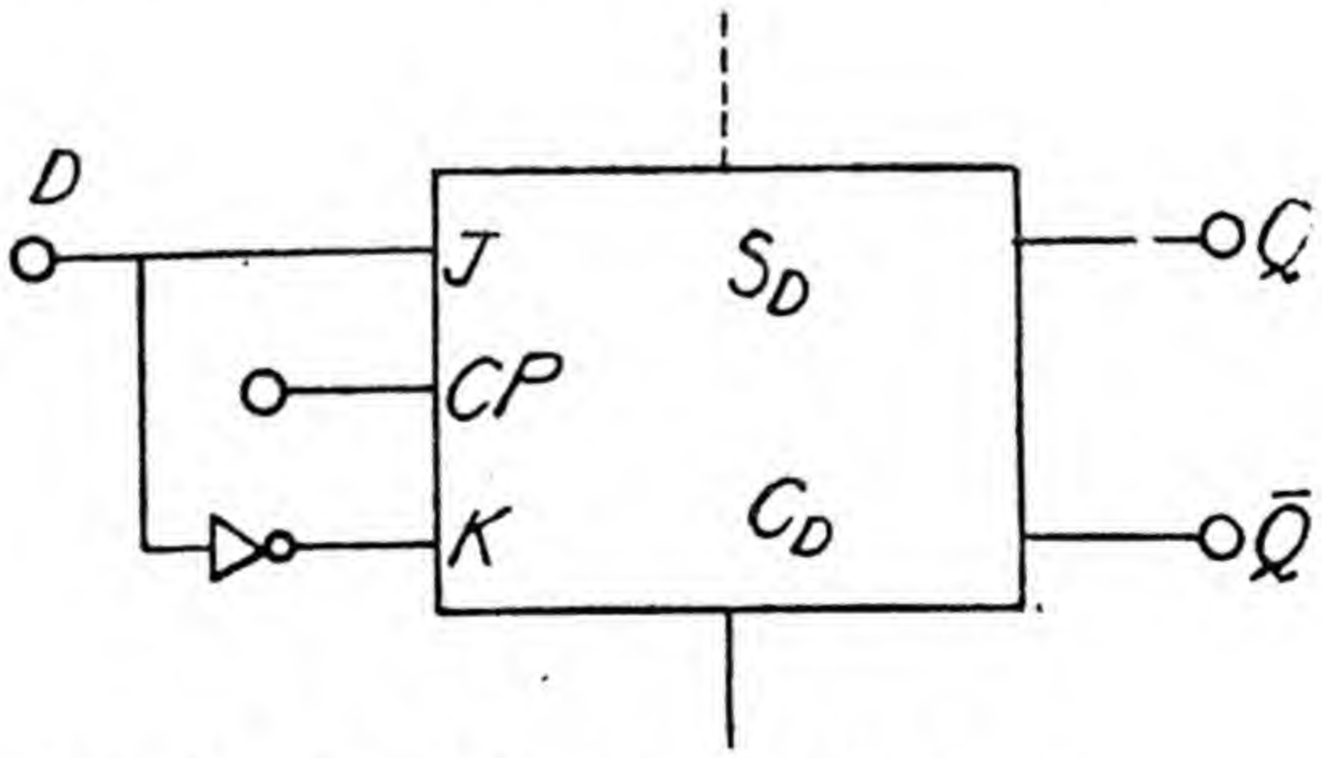


Fig. 45. Symbol of *D*-type flip flop.

If $D=1$ (and $Q=0$) then Q does not become 1 until $CP=1$. That is, *the output after CP, equals the input at D before CP.* Therefore it is thought of as a *delay* (*D*) type of flip flop. The bit on *D* line is transferred to the output at the next clock pulse and hence this unit functions as a 1-bit delay device and is used as a temporary storage device. It is known as a 1-bit bistable latch.

25·19 COUNTERS :

A counter can be described as a tallying device that tallies, or counts, some number of events. An electronic counter needs that

the data be converted into electric pulses, each of which represent one bit of information, or one *happening*. These pulses are introduced at the input and the counter tallies each and every one keeping a cumulative total.

When several flip flops are connected together in a certain way, they form a counting chain that is capable of actually counting or tallying events. Consider four interconnected flip flops as shown in fig. 46. (*a*). The pulses to be counted are applied to the first flip flop. The negative-going pulses have been considered.

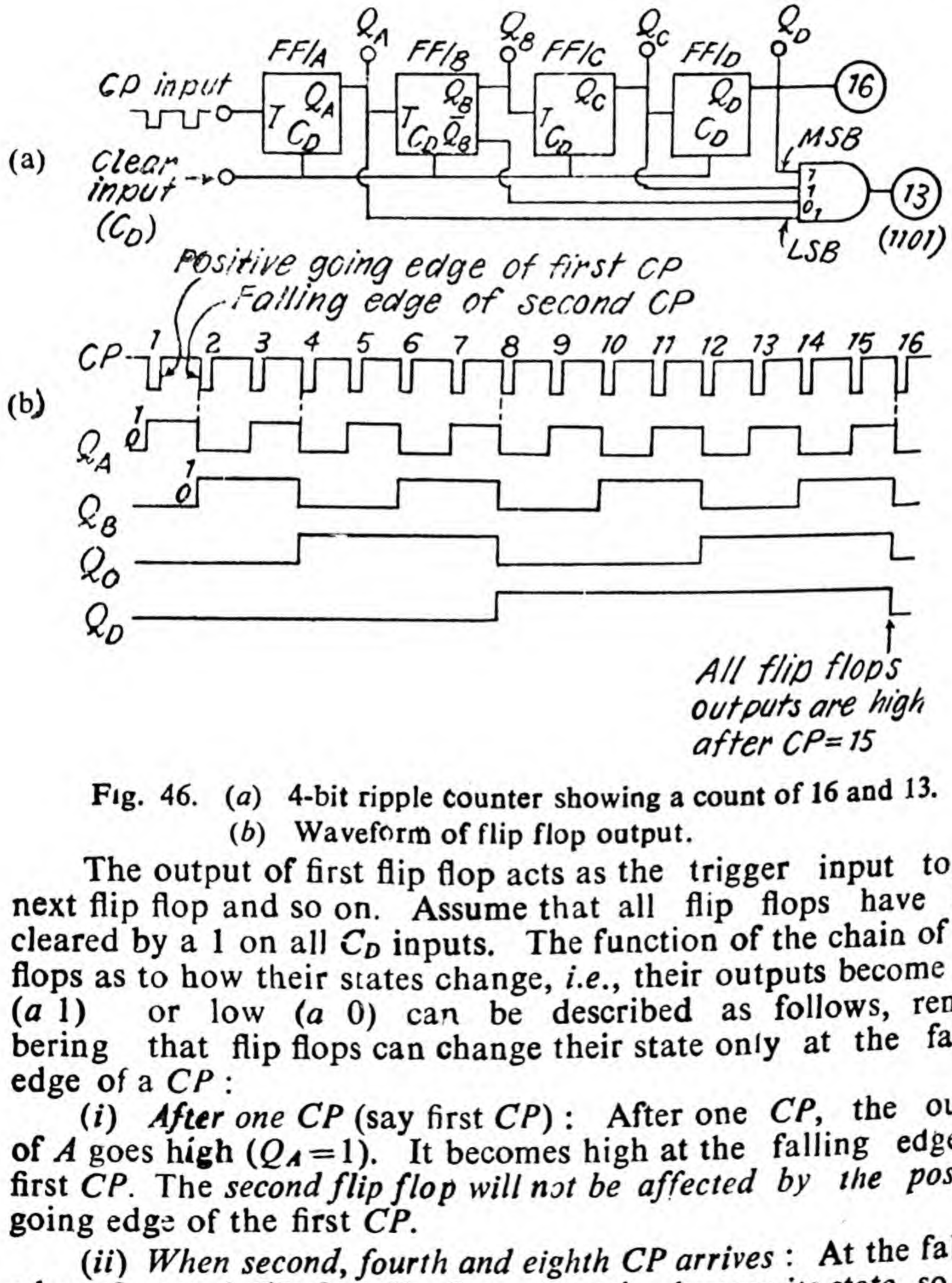


Fig. 46. (*a*) 4-bit ripple counter showing a count of 16 and 13.
(*b*) Waveform of flip flop output.

The output of first flip flop acts as the trigger input to the next flip flop and so on. Assume that all flip flops have been cleared by a 1 on all C_D inputs. The function of the chain of flip flops as to how their states change, *i.e.*, their outputs become high (*a* 1) or low (*a* 0) can be described as follows, remembering that flip flops can change their state only at the falling edge of a *CP* :

(*i*) *After one CP* (say first *CP*) : After one *CP*, the output of *A* goes high ($Q_A=1$). It becomes high at the falling edge of first *CP*. The *second flip flop will not be affected by the positive* going edge of the first *CP*.

(*ii*) *When second, fourth and eighth CP arrives* : At the falling edge of second *CP*, first flip-flop, *A*, again changes its state so that $Q_A=0$ (a low). Thus we get the falling edge of Q_A at which the second flip flop *B*, will be triggered and then its output, Q_B, goes high (*a* 1). It will remain in this state until two more *CP* (num-

bered 3rd and 4th in fig. 46*b*) occur. This is because falling edge of Q_A will occur only at the arrival of fourth *CP*. At the falling edge of Q_A (the input of flip-flop *B*) second flip-flop changes its state so that $Q_B=0$ (fig. 46 *b*).

But with $Q_B=0$ we get the falling edge of Q_B at which the third flip-flop, *C*, will be triggered so that its output $Q_C=1$. It will remain in this state until the arrival of eighth *CP* because falling edge of Q_B will occur only at the arrival of eighth *CP*. At the falling edge of Q_B, third flip-flp changes its state so that $Q_C=0$ (fig. 46 *b*).

But with $Q_C=0$, we get the falling edge of Q_C at which the fourth flip-flop *D* will be triggered so that its output goes high, *i.e.*, $Q_D=1$. It will remain in this state until the arrival of sixteenth *CP* because falling edge of Q_C will occur only at the arrival of sixteenth *CP* (fig. 46 *b*).

(*iii*) *Arrival of sixteenth CP* : After 15 clock pulses, all the flip-flop *Q* outputs are high (*a* 1). Thus when the sixteenth *CP* arrives to reset *A*, all other flip-flop are reset *i.e.* outputs Q_A, Q_B, Q_C and Q_D all go low (*a* 0). The circuit then starts to count once more. This is called *a scale of* 16 *counter*. Fig. 46 (c) shows the state of the flip-flops at each count.

Input Pulse	Flip flop outputs			
	Q_D	Q_C	Q_B	Q_A
0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
5	0	1	0	1
6	0	1	1	0
7	0	1	1	1
8	1	0	0	0
9	1	0	0	1
10	1	0	1	0
11	1	0	1	1
12	1	1	0	0
13	1	1	0	1
14	1	1	1	0
15	1	1	1	1
16	0	0	0	0

For CP=13, $Q_A=1$, $Q_B=0$, $Q_C=1$, $Q_D=1$

All flip flop outputs go high after CP=15

When CP=16 arrives to reset A, all the others are also reset because of falling outputs.

Fig. 46. (*c*) **Table showing state of flips flops at each count.**

Other Uses of this Circuit :

(*i*) *To decode a count of* 13 : Refer to the table fig 46 (*c*). At count 13, $Q_A=1$, $Q_B=0$, $Q_C=1$, $Q_D=1$, Therefore to get a count of 13, we use a four input AND gate with Q_A, $\bar{Q}_B$, Q_C and Q_D as input. All inputs are *a* 1, *i.e.*, $Q_A=1$, $Q_B=1$, $Q_C=1$, $Q_D=1$ so that $\bar{Q}_B=0$. Thus we apply 1101 at the input of AND gate and get a count of 13. We note that most significant bit (MSB) is *D* while least significant bit (LSB) is *A* (fig. 46 *a*).

(*ii*) *As frequency divider* : From fig. 46 (*b*), it is clear that the frequency of output from *A* is one half that of *CP* and the frequency of output from *B* is one-quarter that of *CP*. It leads to the way of designing 'divide by—' counters, *e.g.*, 'divide by 60' etc.

25·20. BCD DECADE COUNTER :

A binary coded decimal (*BCD*) counter is shown in fig. 47. The counter will reset itself to zero after a count of 10. Forced feedback is used to return the counter to a reset state after nine natural binary counts. On count 10, all the outputs are low, *i.e.* outputs of all flipflops are 0. Truth table for this counter is thus same as for a scale of 16 counter upto and including count 9.

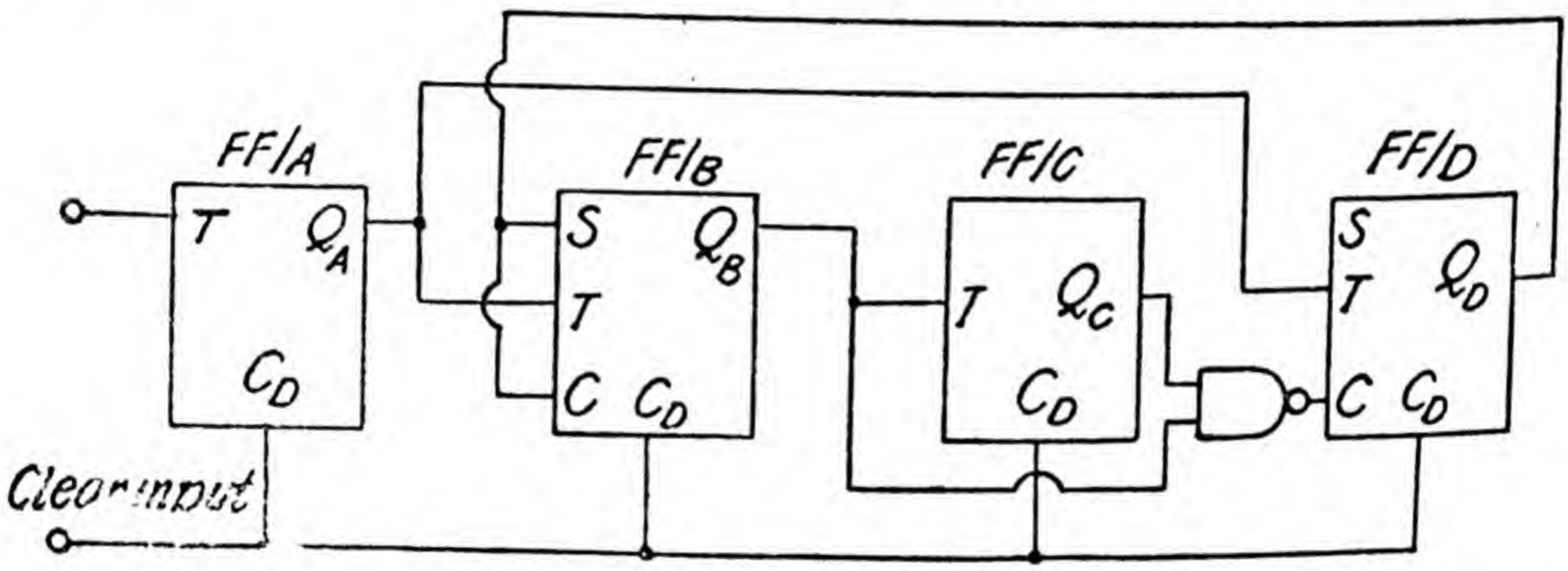


Fig. 47. A BCD decade counter.

The operation can be briefed with the help of truth table (fig. 46*b*) as follows :

(*i*) At the falling edge of every *CP*, flip flop A changes its state *i.e* from a 1 to a 0, and *vice versa.*

(ii) Flip flop *C* changes its state every time *B* changes from a 1 to a 0, *i.e.*, at the falling edge of Q_B (fig. 46*b*). Thus triggering of flip flop *C* depends upon the output of *B*.

(iii) Flip flop *B* changes its state when Q_A goes from a 1 to a 0 (*i.e.* at the falling edge of Q_A, fig, 46*b*) , *except* on tenth pulse From fig. (46*b*), we note that though at falling edge of tenth pulse Q_A goes from a 1 to a 0 but Q_B will not rise from a 0 to a 1. The reason is that because now Q_D has also become a 1 and as it is fedback to the input of *B*, the input of *B* remains high (*a* 1) and thus its output Q_B can not change its state, *i.e.*, it continues to be a low (a 0) In otherwords *B is inhibited from changing back to a* 1 *at this time.*

Thus at the arrival of tenth pulse, Q_A goes from a 1 to a 0, Q_B is forced to 0, and $Q_C=0$ (already low). Since $Q_B=0$, $Q_C=0$ there will a 1 at the clear input, C, of the flip flop D so that Q_D will also be a 0. Thus Q outputs of all flip flops go to a 0 *i.e.* counter resets itself to zero, to be ready for next count.

It remained to explain when Q_D goes 1. We observe from the truth table that flip flop D never changes its state during counts 1 through 7, This is because at its clear input, C, a *sustained* 1 is always present on account of either Q_B or Q_C (or both) being 0 when Q_A drops from 1 to 0. Thus input of D remains high and $Q_D=0$. But after count 7, this high clear input is removed and Q_D goes high (all when Q_A goes from 1 to 0 This happens at tenth pulse. The result is that FF_B is inhibited and counter resets itself. It may be remembered that such a reset condition is achieved only after the application of 16th pulse in a scale of 16 counter.

EXERCISES AND PROBLEMS

1. Express the decimal number 27 in binary.
2. Express the binary number 101101 in decimal.
3. Express the decimal number 537 as a binary number.
4. Add the following binary numbers. Prove yours answers :

   (*a*) 11010011 + 10010111 ——— (*b*) 001101 + 101111 ———

5. Subtract the following binary numbers and prove yours answers :

   (*a*) 1000010 − 0100011 ——— (*b*) 1110001 − 0010111 ———

6. Express the number 275 in *BCD*.
7. Exdress 1011 in decimal form using Gray code.
8. Draw the symbol for a three variable :

   (*a*) AND gate (*b*) OR gate (c) NAND gate (*d*) NOR gate.

9. In the equation $R=\overline{AB}+\overline{C}$ indicate whether R is a 0 or a 1 under the following conditions :

   (*a*) A, B, and $C=0$

   (*b*) $A=1$, $B=0$, $C=0$

   (*c*) A, B, and $C=1$.

10. Reduce the following Boolean functions :

    (*a*) $A+\bar{A}B+AB$

    (*b*) $A\bar{B}+\bar{A}B+AB+\bar{A}\bar{B}$

    (*c*) $A\bar{B}D+\bar{A}\bar{B}D+\bar{B}\bar{D}$

    (*d*) $A\bar{B}C+\bar{A}\bar{B}C$

    (*e*) $(1+B)(ABC)$.

11. Without reducing convert the following expressions to AND/OR/ invert logic :
    (*a*) $A+B+AB$
    (*b*) $\overline{ABC}+A\bar{B}$.
12. Discuss, with merits and demerits, the various logic families.
13. Describe an exclusive OR gate and its applications.
14. Describe the operation of half and full adder. What is a parallel adder ?
15. Discribe how Karnaugh maps help to simplify a Boolean equation.
16. What do you understand by sequential and combinational logic circuits. Describe a *S–R* latch operation.
17. Desciribe a clock R–S flip flop, explaining clearly the way in which trigger pulse controls the change of state of flip flops.
18. Describe J–K flip flop and master slave J–K flip flop. Describe its the merits over clocked flip flop.
19. Describe the basic circuit involved in counters. Explain the working of a *BCD* decade counter.

26 TRANSMISSION LINES

Radio communication requires the transference of energy between the transmitter and receiver. During this transmission three main intermediate phenomena are involved as stated below :

(*i*) The intelligence in the form of radio frequency power is available at the transmitter output. For the release of this energy to space, it is necessary to feed it to some radiating conductor. Because of disturbance, this process of radiation cannot be performed near to transmitter location. Thus first step is to transfer this r.f. power from transmitter to the radiating element. A device with minimum energy losses, used for the transference of r.f. energy to some radiator is called a 'TRANSMISSION LINE'.

(*ii*) The second step is the release of this r.f. energy into space or the communication of intelligence. For this energy release, transmission line is terminated by a radiating element ; transferred r.f. energy affected by the intelligence, is applied to this radiator, which sets up high frequency currents along its length The electromagnetic waves are produced into the space due to the escape of electrical energy from this radiating element. which then travel with the velocity of light. This radiating element is usually known as'ANTE NNA'.

(*iii*) The third step is the mechanism of the 'PROPAGATION OF ELECTROMAGNETIC WAVES'. Generally these are propagated as ground waves, sky waves or space waves.

Transmission line is a device to transmit radio frequency power from one place to another. In order to avoid the disturbance which would occur if the transmitter and antenna are located at the same place, a transmission line with minimum loss of energy is applied between the transmitter and antenna. The most usual form of transmission lines are :

(*a*) a pair of conducting lines in the forms of parallel wires,

(*b*) a pair of concentric tubular conductors, *i.e.*, coaxial cable.

The length of the line system is usually great enough so that, at any given instant of time, the current distribution in it as a result of applied e.m.f. is not uniform. The line consists of resistance, inductance and capacicance which are distributed along the length of the line.

As the purpose of transmission line is to transmit power from generator to load, therefore its study should include the following :

(*i*) in order to have an idea of power at different points on the line, it is essential to know the magnitudes, and phase relationships of voltage and current at different line lengths. This requires the derivation of a relation which connects voltage or current with line length.

(*ii*) it is very important to study that how the termination of the line affects the power transmission by the line on account of changes in the magnitude and phase of voltages and currents on the line.

(*iii*) in order that maximum of energy received by the line be transferred to the load, it is essential to obtain the conditions under which losses and distortion on the line be minimum.

DISTRIBUTED PARAMETERS :

Transmission of energy is affected at high frequencies at which the inductance of short length of conductors, and capacitances between short conductors and their surroundings become *appreciable* and their effects *intermingle* at each point. Thus instead of lumped parameters, we owe transmission line concepts to distributed parameters. We first account for the distributed parameters of a transmission line.

Distributed resistance : The wires of a transmission line formed of some conducting material, will have resistance. This resistance, because of uniform size wires, will be distributed uniformly along their entire length.

Furthermore, when radio frequency waves travel along the transmission line, it is also possible that some part of their energy is lost due to radiation. To account for this consumption (loss) of radiated energy, we can regard a resistance called 'radiation resistance' in addition to line resistance. In the present analysis of lines, we shall ignore this resistance.

Distributed inductance : When radio frequency current flows in the wires, a magnetic field is developed about them. This field offers a sort of resistance to any rise or fall of the line current. Since this field links the wire current, inductance is present and is distributed uniformly along the entire length of the wire. As the inductance impedes the flow of current, it should be assumed in series with the line resistance.

Distributed conductance and capacitance : The fact that input and output currents of the line differ, leads one to realise that there is admittance between the wires even though there is no apparent connection between them. This shunt admittance consists of two components. First, capacitance is present because the two conductors of the parallel line are separated by a dielectric and thus forms

a condenser. This capacitance shunts the line and is, therefore, called shunting capacitance. Second, the dielectric between the wires is not perfect, a conduction current will flow between the wires. This conduction or leakage path may be presented by a conductance between the wires.

Thus an approximate equivalent circuit of a transmission line can be drawn as shown in fig. 1 (*a*).

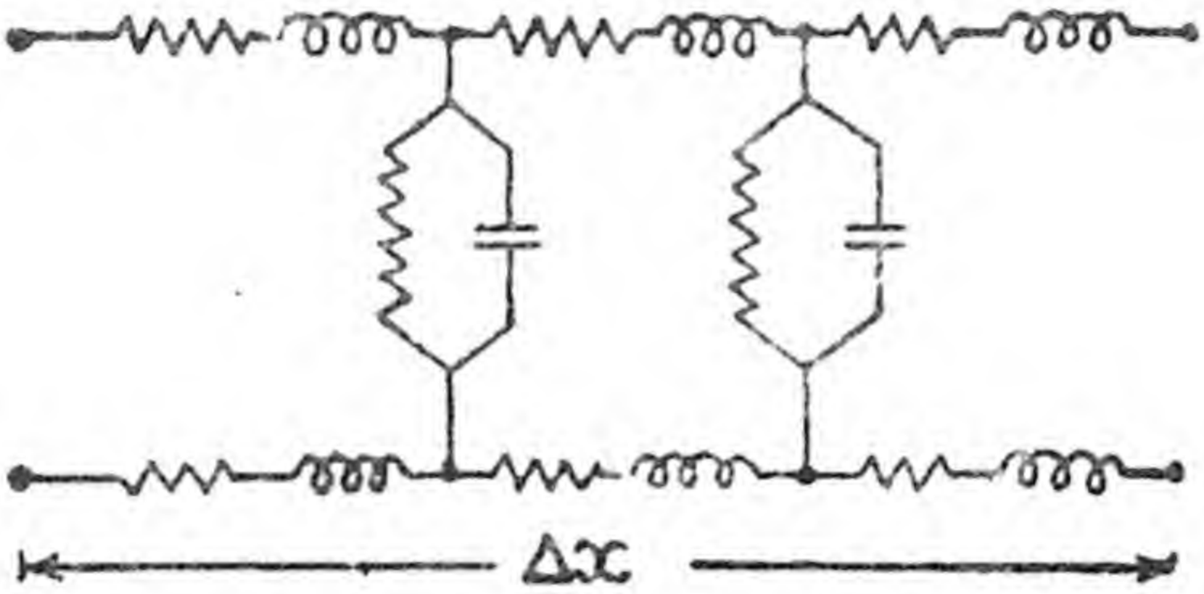


Fig 1. (a) Equivalent circuit of length $\triangle x$ of a transmission line.

TYPES OF TRANSMISSION LINES :

As mentioned earlier there are two types of transmission lines :

1. **Parallel wire line :** This is usually of two types, namely, two wire parallel line and four wire parallel line. Horizontal two wire line and four wire line are balanced to ground devices, *i.e.*. if at a point of one conductor there is some positive voltage with respect to ground, then it will be balanced by an equal negative voltage at the corresponding point of the other conductor. A two wire line in vertical configuration is unbalanced structure because of which it gives rise to unbalanced currents. The unbalanced currents are undesirable for they increase energy losses. Parallel wire lines are used at frequencies of the order 100 mc/s. Above 200 mc/s., power loss due to radiation and induction become excessive. The characteristic impedance of such lines ranges from 100 ohms to 300 ohms.

2. **Coaxial line :** This is best suited at high frequencies. It consists of two concentric conductors separated by an insulating material. The main advantages of such a line are (*a*) losses due to induction and radiation are reduced to minimum because electro-magnetic field associated with a coaxial line is confined to the space between inner and outer conductors, (*b*) characteristic impedance is low, of the order of 50 to 70 ohm.. and (*c*) they can be used at 3000 mc/s; and all lower frequencies because their attenuation even at very high frequencies is very low.

A coaxial line is unbalanced to ground but since the outer conductor acts as a shield, it does not give rise to unbalanced currents.

A coaxial line is also of two forms :

(*a*) *Flexible cables* : In such a cable the whole inner conductor is surrounded by a flexible plastic material and a braid of conducting material woven over the dielectric serves as the outer conductor. To make the cable mechanically strong, entire cable is covered with a steel braid or armour. Such cables are not preferred at 3000 mc/s., since losses in the dielectric medium become appreciable.

(*b*) *Rigid cables* : To avoid energy losses in dielectric medium at high frequencies, rigid coaxial lines with air as dielectric are employed. To support inner conductor within a large outer conductor, beads or quarter wave stubs are employed.

26 0. CALCULATION OF LINE PARAMETERS :

As mentioned above, there are four line parameters, R, L, C and G distributed per unit length of the line. We shall now find expressions for inductance and capacitance of a line of two parallel round conductors and coaxial line separately.

INDUCTANCE OF A LINE OF TWO PARALLEL ROUND CONDUCTORS :

A section of a long round wire with a hypothetical cylinder of length one meter and wall thickness dr is shown in fig. 1 (*b*). Permeability μ is defined as

$$\mu = \mu_r \mu_v,$$

where μ_v is the magnetic permeability of space and has the value $4\pi \times 10^{-7}$ in *mks* units, and μ_r is the relative permeability of the particular material. In a material or region where μ_r is independent of flux density (*e.g.*, air and most dielectrics), inductance can be defined as

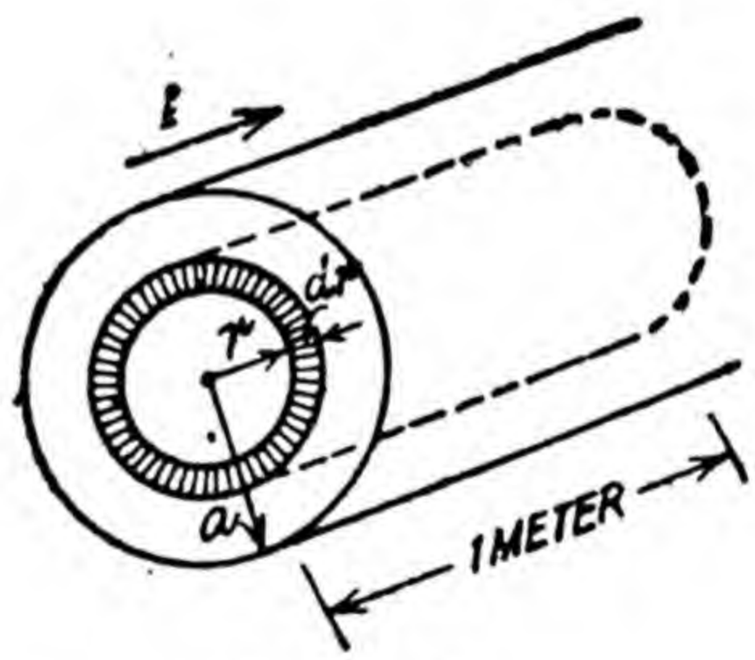


Fig 1 (b)

$$L = \frac{\text{flux linkage of the current}}{\text{total conductor current}} = \frac{\lambda}{I}.$$

Therefore, the problem now reduces to the calculation of flux linkage of the current. In a long, straight, round wire, the flux paths surrounding the current, are concentric circles. The magnetic field intensity at any distance r from the centre of the conductor is, by definition,

$$H_r = \frac{NI_r}{l},$$

where I_r is the current enclosed by the flux path around which H is measured. For only one turn or current path,

$$H_r = \frac{I_r}{l}.$$

The length, l, of the flux path at any distance r is $2\pi r$. Therefore

$$H_r = \frac{I_r}{2\pi r}.$$

If total conductor current, I, is supposed to be distributed uniformly over the cross section of the cylinder then

$$I_r = \frac{\pi r^2}{\pi a^2} I$$

so that

$$H_r = \frac{rI}{2\pi a^2}$$

$$B_r = \mu H_r = \frac{\mu r I}{2\pi a^2},$$

which is the flux dersity at any point *inside* the conductor. Therefore flux in the wall of hollow cylinder of area dA will be

$$B_r dA.$$

For the cylinder of thickness dr and length 1 meter, it will be

$$d\phi = B_r dr = \frac{\mu r I}{2\pi a^2} dr.$$

This flux linkage corresponds to current I_r. In order to correspond it to total current I, we should multiply the above expression by $\frac{r^2}{a^2}$. Then the flux present in the wall of the cylinder of radius r, width dr, and one meter in length, will be

$$= \frac{\mu r I}{2\pi a^2} dr . \frac{r^2}{a^2}$$

Therefore total internal linkages, λ_{int}, due to all the flux inside the conductor will be

$$\lambda_{int} = \frac{\mu I}{2\pi a^4} \int_0^a r^3 \, dr$$

$$= \frac{\mu I}{8\pi} \text{ linkages/meter,}$$

where μ is the permeability of the conductor material.

We shall now determine external flux linkage. Refer to fig. 1 (*c*). The flux density in the space *outside* the conductor is

$$B_r = \mu_v H_r \frac{\mu_v I}{2\pi r},$$

where μ_v is the permeability of the space between the conductors. The flux present in the dashed hollow cylinder of radius r.

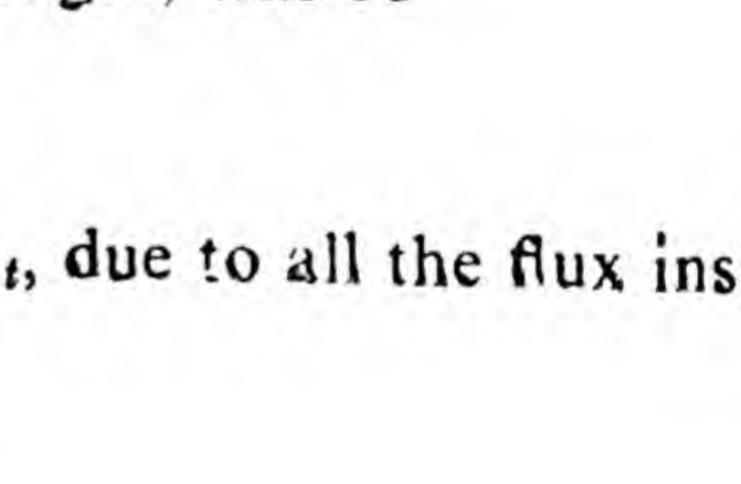


Fig. 1 (*c*)

width dr and one meter length will be

$$d\phi=\frac{\mu_v I}{2\pi r}dr.$$

Consequently flux passing in space between the two conductors and arround the current in X will be

$$\lambda^1_{ext}=\frac{\mu_v I}{2\pi}\int_a^{d-a}\frac{dr}{r}.$$

We shall now calculate the linkage contributed by the flux cutting wire Y. Flux due to the current in X, that passes through wire Y, links values of current varying between

$+I$ for flux element with r just greater than $(d-a)$.

and 0 for flux element with $r=(d+a)$

An average value may be obtained by integrating to the centre of wire Y. accounting as full linkages that flux passing through the left half of Y and as zero linkages the flux passing through right half of Y. Then the contribution by the flux cutting wire Y is

$$\lambda^2_{ext}=\frac{\mu_v I}{2\pi}\int_{d-a}^{d}\frac{dr}{r}.$$

Flux surrounding the conductor Y $[r>(d+a)]$, links both $+I$ and $-I$ currents *i e.* zero net current to X and therefore contributes zero flux linkages.

We now write for the *total external flux* linkages as

$$\lambda_{ext.}=\lambda^1_{ext.}+\lambda^2_{ext.}$$

$$=\frac{\mu_v I}{2\pi}\left[\int_a^{d-a}\frac{dr}{r}+\int_{d-a}^{d}\frac{dr}{r}\right]$$

$$=\frac{\mu_v I}{2\pi}\int_a^d\frac{dr}{r}=\frac{\mu_v I}{2\pi}\log_e\frac{d}{a}\text{ linkges/meter}$$

Therefore total inductance of one wire will be

$$\frac{\lambda}{I}=\frac{\lambda_{int.}+\lambda_{ext}}{I}=\frac{\mu}{8\pi}+\frac{\mu_v}{2\pi}\log_e\frac{d}{a},$$

and for foth the wires or the line, it will be

$$L=2\left(\frac{\mu}{8\pi}+\frac{\mu_v}{2\pi}\log_e\frac{d}{a}\right)$$

$$=\frac{\mu_v}{4\pi}\left(\frac{\mu}{\mu_v}+4\log_e\frac{d}{a}\right)$$

$$=\frac{4\pi\times10^{-7}}{4\pi}\left(\mu_r+9{\cdot}210\log_{10}\frac{d}{a}\right)$$

$$=10^{-7}\left(\mu_r+9{\cdot}210\log_{10}\frac{d}{a}\right)\text{ henry/meter}$$

$$=\left(0{\cdot}1609\,\mu_r+1{\cdot}482\,\log_{10}\frac{d}{a}\right)\text{ mh/mile}$$

Since first term arises due to internal flux linkages, it is called internal inductance of the line; whereas the second term arises due to the linkages of flux external to the wires, it is called external inductance At radio frequencies the current flows almost in the surface of the conductor so that internal induc ance becomes negligible. It should be born in mind that relation has been derived on the assumption that the current distribution over the conductors is uniform which can be true only if spacing between wires is quite large compared to their radius ($d \gg a$).

INDUCTANCE OF COAXIAL LINE :

A coaxial line crosssection is shown in fig. 1 (*d*) The radius of inner conductor is *a*, internal radius of outer conductor is *b*, and external radius is *c*. We assume that the frequency is low so that current is distributed uniformly over the cross section of the conductor. Inner conductor carries current $+I$ amperes and outer conductor $-I$ amperes. It is thus obvious that any flux path surrounding the outer conductor will enclose $(+I-I=0)$ zero current and consequently there is no flux external to outer conductor.

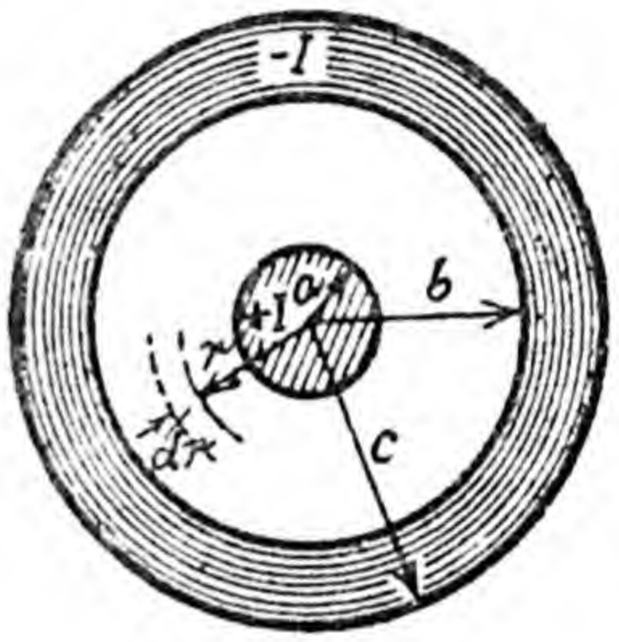


Fig. 1 (d)

In this case, the total flux linkages will be made up of following three contributions :

(*i*) Flux linkages internal to the central conductor : This we have calculated in previous case and can be shown to be given by

$$\lambda_1 = \frac{\mu_1 I}{8\pi} \text{ link ıges/meter,}$$

where μ_1 is the permeability of the material of inner conductor.

(*ii*) Flux linkages of the central conductor current $+I$, due to flux in the dielectric between the conductors. This we have also calculated previously and is given by

$$\lambda_2 = \frac{\mu_v I}{2\pi}\int_a^b \frac{dr}{r} = \frac{\mu_v I}{2\pi}\log_e \frac{b}{a} \text{ linkages/meter.}$$

where μ_v is the permeability of the dielectric.

(*iit*) Flux linkages due to the flux passing through the outer conductor, between *b* and *c*.

For this we shall assume a cylindrical shell of width *dr* and radius *r* which is greater than *b* but less than *c*, that is inside the outer conductor. The current that produces the flux surrounding the inner conductor passing within the outer conductor is $+I$ at $r=b$ and varies to zero upto $r=c$. Therefore current enclosed by flux at radius *r* is

$$=I+\frac{\pi\,(r^2-b^2)}{\pi\,(c^2-b^2)}\,(-I)$$

$$=I\frac{(c^2-r^2)}{(c^2-b^2)}$$

so that corresponding flux density B_r at radius r will be

$$B_r=\frac{\mu_2 I}{2\pi r}\,\frac{(c^2-r^2)}{(c^2-b^2)},$$

where μ_2 is the permeability of the material of the outer conductor.

Here it is important to note that a flux line at $r=b$ makes a complete linkage of current I but on the otherhand. a flux line at $r=c$ links zero current. That is, contribution of linkage varies from unity to zero. Then fractional part of a linkage contributed by a flux line at radius r between b and c is

$$1-\frac{\pi\,(r^2-b^2)}{\pi\,(c^2-b^2)}=\frac{c^2-r^2}{c^2-b^2}.$$

Finally, we now write linkages due to the flux passing through outer conductor between b and c. That is

$$\lambda_3=\frac{\mu}{2\pi}I\int_b^c\frac{c^2-r^2}{c^2-b^2}\cdot\frac{c^2-r^2}{c^2-b^2}\cdot\frac{dr}{r}$$

$$=\frac{\mu_2 I}{2\pi\,(c^2-b^2,^2}\left[c^4\log_e\frac{c}{b}-c^2(c^2-b^2)+\frac{1}{4}\,(c^4-b^4)\right]$$

linkages/meter

The inductance of the coaxial line will be

$$L=\frac{\lambda_1+\lambda_2+\lambda_3}{I}.$$

Taking $\mu_1=\mu_v=\mu_2$ (because ordinarily the material of a line are non-magnetic), we arrive at

$$L=\frac{\mu_v}{8\pi}\left[1+4\log_e\frac{b}{a}+\frac{4}{(c^2-b^2)^2}\left\{c^4\log_e\frac{c}{b}\right.\right.$$

$$\left.\left.-c^2\,(c^2-b^2)+\tfrac{1}{4}\,(c^4-b^4)\right\}\right]$$

But $\mu_v=4\pi.10^{-7}$, we write

$$L=10^{-7}\left[2\log_e\frac{b}{a}+\frac{2c^4\log_e c/b}{(c^2-b^2)^2}-\frac{c^2}{c^2-b^2}\right]$$

henrys/meter

At radio frequencies, the current crowds to the outer surface of the inner conductor and inner surface of the outer conductor so that λ_1 and λ_3 are negligible. In that case

$$L=2\times10^{-7}\log_e\frac{b}{a}\text{ henrys/meter}$$

$$= 0{\cdot}741 \log_{10} \frac{b}{a} \text{ mh/mile}$$

We have assumed here that there is no flux outside the outer conductor but if outer conductor is braided from small wires (in order to have a flexible cable) then some flux may leak out and appear external to the cable.

CAPACITANCE OF TWO PARALLEL ROUND CONDUCTORS:

Let us consider a section of a long round conductor of radius a. Let us enclose this cylinder by another concentric hypothetical cylinder of radius r and length one meter. Then electric flux density D at P, a point on the enclosing cylinder is

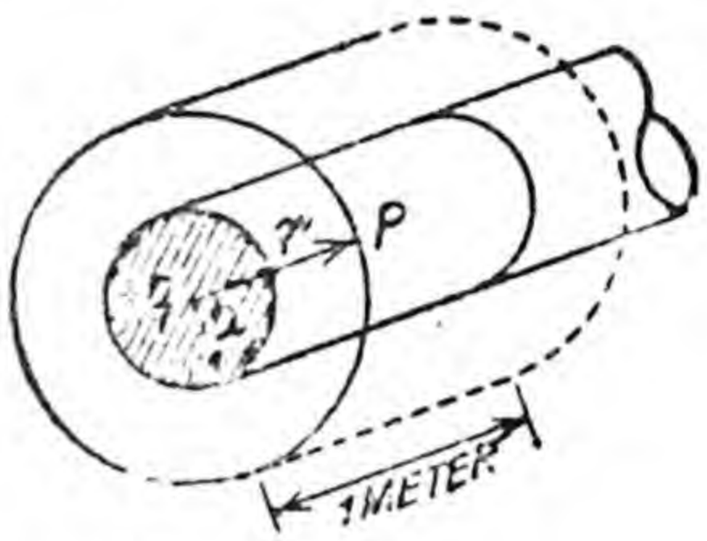


Fig. 1 (e)

$$D = \frac{q}{2\pi r} \text{ coulombs/meter}^2$$

so that electric field intensity at radius r is

$$E = \frac{D}{\epsilon} = \frac{q}{2\pi r \epsilon} \text{ volt/meter}$$

where ϵ is permittivity defined as

$$\epsilon = \epsilon_r \, \epsilon_v$$

ϵ_r is the dielectric constant or relative permittivity and ϵ_v is the permittivity of space. The potential at P will be, then,

$$V_r = -\int_a^r \frac{q}{2\pi r \epsilon} \, dr = \frac{-q}{2\pi\epsilon} \log_e \frac{r}{a} \text{ volts}$$

A parallel wire transmission line is ordinarily constructed so that spacing d between wire is large with respect to the radius a of a conductor. The charge can then be considered as uniformly distributed around the periphery of each conductor.

The potential difference between the two conductors, shown in fig. 1 (f) when left wire has a charge of $+q$ and right wire a charge of $-q$ coulombs per unit length, taking as a reference the negatively charged wire is

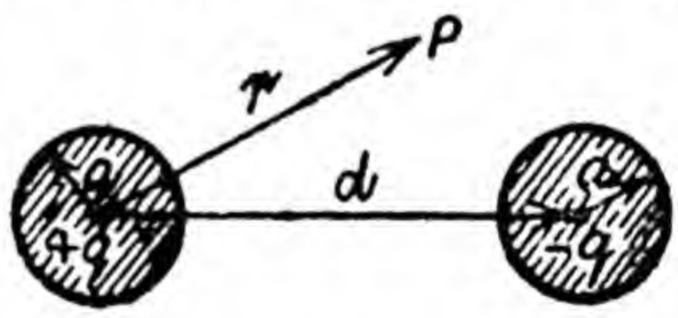


Fig. 1 (f)

$$V = \frac{q}{\pi\epsilon} \log_e \frac{d-a}{a} - \frac{(-q)}{\pi\epsilon} \log_e \frac{d-a}{a}$$

$$= \frac{q}{\pi\epsilon} \log_e \frac{(d-a)}{a} \text{ volts.}$$

Therefore capacitance is

$$C = \frac{q}{V} = \frac{q}{\frac{q}{\pi\epsilon} \log_e \frac{d-a}{a}} = \frac{\pi\epsilon}{\log_e \frac{d-a}{a}} \text{ farads/}m.$$

Since $d \gg a$, we can write d inplace of $(d-a)$. Therefore

$$C = \frac{\pi\epsilon}{\log_e \frac{d}{a}} \text{ farads/meter}$$

$$= \frac{12\,07\ \epsilon_r}{\log_e \frac{d}{a}}\ \mu\mu f/m$$

$$= \frac{1{\cdot}943 \times 10^4}{\log_e \frac{d}{a}}\ \mu\mu f/\text{mile.}$$

CAPACITANCE OF THE COAXIAL LINE :

Cross section of a coaxial transmission line is shown in figu e 1 (g). The inner conductor is given a charge $+q$ and outer conductor is given a charge $-q$, which are distributed uniformly on the outer surface of inner conductor and inner surface of outer conductor. The field intensity at P is then

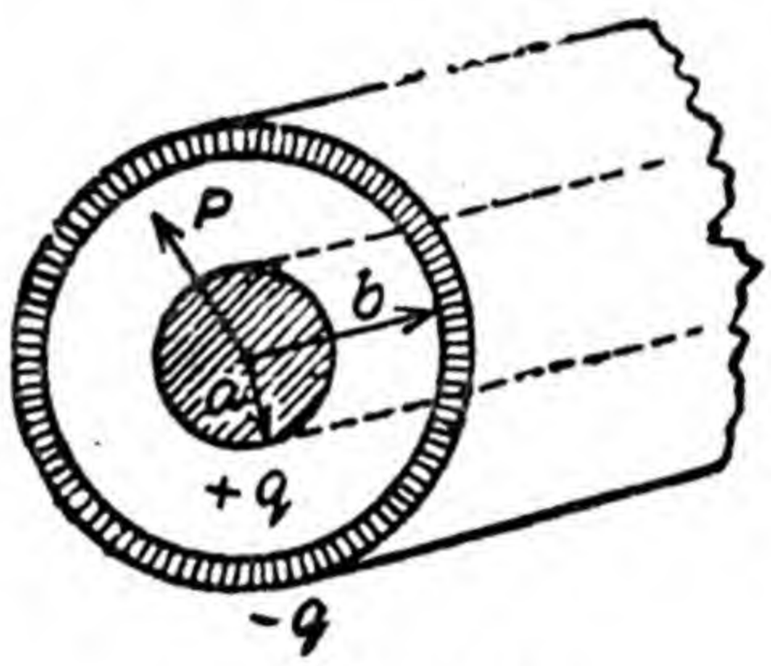


Fig. 1 (g)

$$E = \frac{q}{2\pi r\epsilon} \text{ volt/meter}$$

and the potential difference from outer to the inner conductor is

$$V = -\int_b^a \frac{q}{2\pi r\epsilon}\, dr = \frac{q}{2\pi\epsilon} \log_e \frac{b}{a} \text{ volts.}$$

The capacitance of the coaxial line will be

$$C = \frac{q}{V} = \frac{2\pi\epsilon}{\log_e \frac{b}{a}} \text{ farads/meter}$$

$$= \frac{24{\cdot}14\ \epsilon_r}{\log_{10} \frac{b}{a}}\ \mu\mu f/\text{meter}$$

$$= \frac{3{\cdot}886 \times 10^4\ \epsilon_r}{\log_{10} \frac{b}{a}}\ \mu\mu f/\text{mile.}$$

LINE PARAMETERS AT HIGH FREQUENCIES : OPEN WIRE LINE :

Parameters of the line are modified by the presence of high frequency currents. Skin effect comes into play as discussed below :

We know that when an alternating current flows in a conductor, the alternating magnetic flux within the conductor induces an e.m.f. This e.m.f. causes the current density to decrease in the interior of the wire and to increase at the outer surface. This effect is known as skin effect. It is more marked at radio frequencies and therefore the current in the wire of a moderate size is concentrated in a thin skin at the surface. The effective thickness of the surface layer of current may be given by

$$\delta = \sqrt{\left(\frac{\rho}{\pi f \mu}\right)} \text{ meters}$$

$$= \frac{1}{\sqrt{(\pi f \mu \sigma)}} \text{ meters}$$

where ρ is the resistivity of the conductor in ohm meters

σ is the conductivity of conductor in mhos per meter

f is the frequency of alternating current in c/s

and μ is the absolute magnetic permeability of the conductor and for free space, its value is $4\pi \times 10^{-7}$ henry meter.

For copper $\mu = 4\pi \times 10^{-7}$, $\sigma = 5{\cdot}75 \times 10^{7}$ mhos/meter at 20°C. Then

$$\delta = \frac{0{\cdot}0664}{\sqrt{f}} \text{ meters.}$$

Thus the nominal depth of penetration for copper conductor is about 86×10^{-4} meters at 60 c/s and is only 66×10^{-6} meters at 1 mc/s.

R, the resistance per unit length of the line : The re istance of a round conductor of radius a meters to direct current is inversely proportional to the area as

$$R_{d.c.} = \frac{K}{\pi a^2},$$

while that of a round conductor with alternating current flowing in a skin of thickness δ is

$$R_{a.c.} = \frac{K}{2\pi a \delta}$$

Therefore

$$\frac{R_{a.c.}}{R_{d.c.}} = \frac{a\sqrt{(\pi f \mu \sigma)}}{2}$$

$$= 7\,63 a\sqrt{f} \text{ for copper,}$$

where a is in meters and f in c/s Relation implies that for the conductor of large radius increase in resistance with increase of frequency will be greater compared in the conductor of small radius. For an open wire line of copper with spacing greater than 20 a, it can be computed that

$$R_{a.c} = \frac{8{\cdot}33 \times 10^{-8}\sqrt{f}}{a} \text{ ohms/meter.}$$

L, the inductance per unit length of the line : We have already shown that for two wire line, when internal impedance becomes negligible at high frequencies, the inductance is given by

$$L=\frac{\mu_0}{2\pi}\log_e\frac{d}{a}$$

$$=4\times10^{-7}\log_e\frac{d}{a}\text{ henrys/meters.}$$

$$=9{\cdot}21\times10^{-7}\log_{10}\frac{d}{a}\text{ henrys/meters.}$$

C, capacitance per unit length of the line : The value of capacitance of a line is not affected by skin effect or frequency. Therefore the relation

$$C=\frac{\pi\epsilon}{\log_e\frac{d}{a}}\text{ farads/meter}$$

$$=\frac{12{\cdot}07}{\log_e\frac{d}{a}}\,\mu\mu f/m$$

is still applicable.

LINE PARAMETERS AT HIGH FREQUENCIES : COAXIAL LINE

Due to the skin effect, resistance of a copper coaxial line is given by

$$R_{a.c.}=4{\cdot}16\times10^{-8}\sqrt{f}.\left(\frac{1}{b}+\frac{1}{a}\right)\text{ ohms/meter,}$$

where a and b are outer radius of inner conductor and inner radius of the outer conductor in meters.

Because of skin effect, the current flows in the outer surface of the inner conductor and inner surface of the outer conductor. Due to this effect flux linkage due to inner conductor flux is eliminated and the inductance of the coaxial line is given by

$$L=4{\cdot}60\times10^{-7}\log_e\frac{b}{a}\text{ henrys/meter.}$$

Capacitance of coaxial line is not affected by frequency and is given by

$$C=\frac{24{\cdot}14\ \epsilon_r}{\log_{10}\frac{b}{a}}\,\mu\mu f/\text{meter.}$$

For the transmission of energy at high frequencies, it is posisible to assume negligible losses or zero dissipation in the analysis of performance of transmission lines.

26·1. VOLTAGE AND CURRENT RELATIONS ON RADIO FREQUENCY TRANSMISSION LINE :

At frequencies of operation so high that the wavelength becomes comparable to physical dimensions of the apparatus, analysis cannot be made with circuit concept, *e g.* in terms of lumped parameter equivalent circuits and hence it is preferable to consider directly electric and magnetic fields. Here we shall analyse the line in terms of voltage and current.

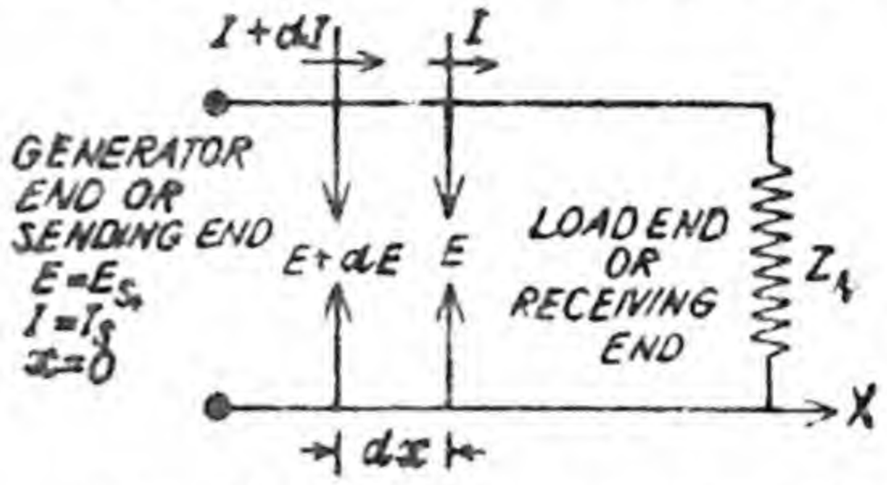


Fig. 2. (*a*) Transmission line

The voltage and current relations in a transmission line may be studied by considering a basic type of line which consists of two straight parallel wires of uniform size, separated by air. Consider the voltage and current relations that exist in a very short length, dx, of the transmission line as shown in fig. 2 (*a*).

Let R = loop resistance per unit length of line
= sum of resistance of *both* the wires for unit line length,
L = loop inductance per unit length of line
= sum of inductance of *both* the wires for unit line length,
G = shunt conductance between wires per unit length,
C = shunt capacitance between wires per unit length.

In travelling this short length dx, the voltage between the wires changes by an amount dE due to the voltage drop produced by a line current through the impedance Z consisting of the resistance $R\,dx$ and the inductive resistance $j\omega L\,dx$ of the length dx. Likewise the current changes by an amount dI due to the flow of current through the admittance Y having components, the capacitive reactance $j\omega C\,dx$ and the conductance $G\,dx$ of the length dx of the line as a result of the voltage E existing between them. Thus we have

$$Z=(R+j\omega L)\,dx,$$

$$*Y=(G+j\omega C)\,dx$$

Therefore fall in voltage,

$$dE=(R+j\omega L)\,dx\,I. \qquad \ldots(1)$$

and fall in current.

$$dI=\frac{E}{1/(G+j\omega\ldots)\,dx}$$

$$=(G+j\omega C)\,dx\,E. \qquad \ldots(2)$$

Differentiating eq (1) with respect to x, we get

$$\frac{d^2E}{dx^2}=(R+j\omega L)\frac{dI}{dx}$$

$$*Y=\frac{1}{\text{Corresponding impedance}}=\frac{1}{1/G}+\frac{1}{1/j\omega C}=G+j\omega C.$$

$$=(R+j\omega L)\,(G+j\omega C)\,E \qquad \text{(from eqn. 2)}$$
$$=\gamma^2 E$$

or
$$\frac{d^2E}{dx^2}-\gamma^2 E=0, \qquad \ldots(3)$$

where $\gamma=\sqrt{\{(R+j\omega L)\ (G+j\omega C)\}}$, called propagation constant, The solution of equation (3) will be of the form

$$E=Ae^{\gamma x}+Be^{-\gamma x}, \qquad \ldots(4)$$

where A and B are the constants to be determined with the help of boundary conditions.

Further from equation (1), we have

$$I=\frac{1}{(R+j\omega L)}\cdot\frac{dE}{dx}$$
$$=\frac{\gamma}{(R+j\omega L)}\ (Ae^{\gamma x}-Be^{-\gamma x}) \qquad \{\text{from eqn. (4)}\}$$
$$=\sqrt{\left\{\frac{G+j\omega C}{R+j\omega L}\right\}}\ (A^{\gamma x}-Be^{-\gamma x})$$
$$=\frac{1}{Z_0}\ (Ae^{\gamma x}-Be^{-\gamma x}), \qquad \ldots(5)$$

where $Z_0=\sqrt{\left(\frac{R+j\omega L}{G+j\omega C}\right)}$, called *characteristic impedance* of the line and wlll be defined later on.

The voltage and current equations for the line can be determined with reference to load end or sending end (generator end). First, we shall take the sending end as the origin of coordinates.

Voltage and Current Relations with distance from Sending End :

The boundary conditions are

At $x=0$ $\qquad I=I_s,$
$\qquad E=E_s.$

Also $\qquad E_s=Z_sI_s.$

With these conditions equations (4) and (5) give

$$E_s=A+B.$$
$$I_s=\frac{1}{Z_0}(A-B),$$

so that

$$A=\frac{1}{2}\ (E_s+Z_0I_s)$$
$$=\frac{1}{2}\ I_s\ (Z_s+Z_0)$$

and
$$B=\frac{1}{2}\ (E_s-Z_0I_s).$$

$$=\frac{1}{2} I_s (Z_s - Z_0).$$

Putting these values of constants, equations (4) and (5) give the desired relations for voltage and current on any point on the line. That is

$$E=\frac{1}{2} I_s (Z_s+Z_0) e^{\gamma x}+\frac{1}{2} I_s (Z_s-Z_0) e^{-\gamma x}$$

$$=I_s.Z_s\left(\frac{e^{\gamma x}+e^{-\gamma x}}{2}\right)+I_s Z_0\left(\frac{e^{\gamma x}-e^{-\gamma x}}{2}\right)$$

$$=E_s \cosh \gamma x + I_s Z_0 \sinh \gamma x. \quad \ldots(6)$$

Similarly,

$$I=\frac{1}{Z_0}\left[\frac{1}{2} I_s (Z_s+Z_0) e^{\gamma x}-\frac{1}{2} I_s (Z_s-Z_0) e^{-\gamma x}\right]$$

$$=I_s\left(\frac{e^{\gamma x}+e^{-\gamma x}}{2}\right)+\frac{E_s}{Z_0}\left(\frac{e^{\gamma x}-e^{-\gamma x}}{2}\right)$$

$$=I_s \cosh \gamma x+\frac{E_s}{Z_0} \sinh \gamma x. \quad \ldots(7)$$

If we put $\gamma=\alpha+j\beta$ in equation (4), then

$$E=Ae^{(\alpha+j\beta)x}+Be^{-(\alpha+j\beta)x}$$

$$=Ae^{\alpha x} e^{j\beta x}+Be^{-\alpha x} e^{-j\beta x}. \quad \ldots(8)$$

The second term in equation (8) reperesents a voltage component travelling towards the receiving or load end of the line and is called the incident wave; this decays exponentially in the positive direction of x. The first term represents a wave similar to the second term but travelling from the load towards the generator end of the line; it is termed as reflected wave.

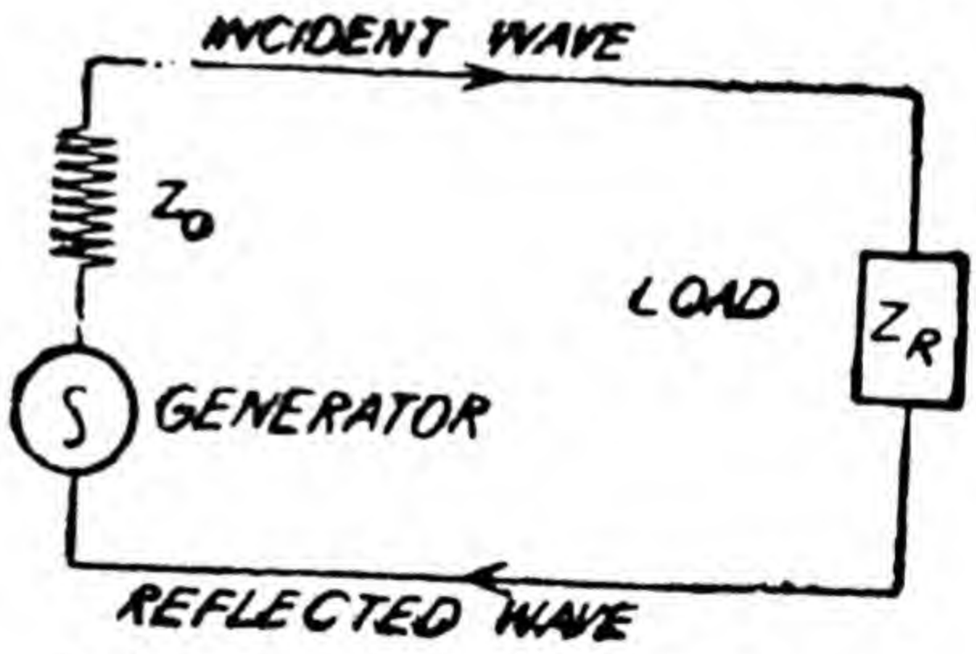


Fig. 2 (*b*) Travelling waves

Voltage and Current Equations with distance referred from Load End or Receiving End :

In this case boundary conditions will be

At $x=0$, $I=I_R$,

$E=E_R$,

Hence $E_R=A+B$,

$Z_0 I_R=A-B$.

$$\therefore \quad A=\tfrac{1}{2}(E_R+Z_0 I_R)$$

$$=\frac{I_R}{2}(Z_R+Z_0)$$

and $$E=\frac{I_R}{2}(Z_R-Z_0).$$

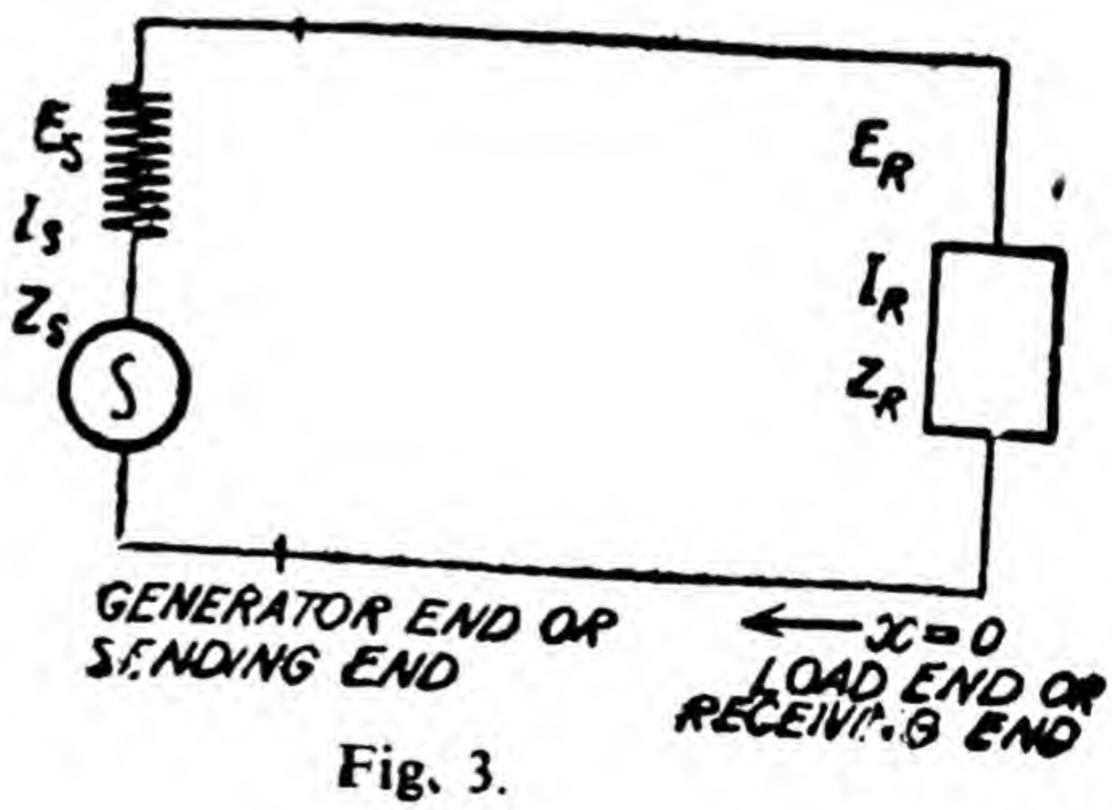


Fig. 3.

Thus $(A+B)=I_R Z_R=E_R,$

$(A-B)=I_R Z_0.$

So $E=E_R \cosh \gamma x+I_R Z_0 \sinh \gamma x,$

$I=I_R \cosh \gamma x+E_R/Z_0 \sinh \gamma x.$

These equations are referred to the load end as the origin of co-ordinates.

$$E=Ae^{\alpha x}.e^{j\beta x}+Be^{-\alpha x}.e^{-j\beta x}.$$

When x is measured from load end, amplitude of reflected wave would decrease with an increase in x and hence $Be^{-\alpha x}$ is the amplitude of reflected wave and $Ae^{\alpha x}$ that of incident wave. The ratio of reflected voltage to the incident voltage is termed as reflection coefficient, *i e.*

$$|\rho|=\frac{Be^{-\alpha x}}{Ae^{\alpha x}}=\frac{B}{A}e^{-2\alpha x}$$

$$=\frac{Z_R-Z_0}{Z_R+Z_0}.e^{-2\alpha x}.$$

If we put $Z_R=Z_0$, then $|\rho|=0$; also if we put $x=\infty$, again $|\rho|=0$, *i.e.* in both the cases no reflected wave is produced, instead whole of the incident wave is absorbed. On viewing the two conditions together, we arrive at that a line of finite length, terminated in a load equivalent to its characteristic impedance, appears to the sending end as an infinite line. Further

$$Z=\frac{E}{I}$$

$$=\frac{Ae^{\gamma x}+B^{-\gamma x}}{\frac{1}{Z_0}(Ae^{\gamma x}-Be^{-\gamma x})}$$

For sending end $x=l$, he length of line, so that at the generator end

$$Z_S=\frac{E_s}{I_s}=\frac{Ae^{\gamma l}+Be^{-\gamma l}}{\frac{1}{Z_0}(Ae^{\gamma l}-Be^{-\gamma l})}.$$

Substituting,

$$\frac{B}{A}=\frac{Z_R-Z_0}{Z_R+Z_0},$$

$$Z_S=Z_0.\frac{\left\{e^{\gamma l}+\left(\frac{Z_R-Z_0}{Z_R+Z_0}\right)e^{-\gamma l}\right\}}{e^{\gamma l}-\left(\frac{Z_R-Z_0}{Z_R+Z_0}\right)e^{-\gamma l}}$$

When $Z_R=Z_0$ or $l=\infty$,

then $Z_S=Z_0,$

and therefore, the characteristic impedance is the input impedance

offered by a line of infinite length. Z_0 is called *characteristic impedance*, because it depends only on the parameters which are characteristic of the unit length of transmission line (not as a whole) and the frequency of operation, and *not* on any terminal conditions.

Example. Preliminary calculations estimate that in a transmitting system installation, 4 miles of a coaxial transmission line of characteristic impedance 400 ohms would be needed. But when installation is made then only 3·5 miles of cable is used. How does this affect the characteristic impedance of the line ?

Obviously, there would be no affect on the characteristic impedance of the line because it depends only on the line constants L, C, R, and G, and not on the the line length.

PROPAGATION CONSTANT :

The relation between current and voltage for a transmission line is given by

$$E = E_s \cosh \gamma x + I_s Z_0 \sinh \gamma x,$$

$$I = I_s \cosh \gamma x + \frac{E_s}{Z_0} \sinh \gamma x,$$

where γ is known as the propagation constant. The quantity $\gamma = \sqrt{(ZY)}$ is seen to govern the manner in which E and I vary with x, *i.e., it governs the way in which the waves are propagated and hence the name.*

The propagation constant of a line per unit length is defined as the natural log of the vector ratio of the steady state currents entering and leaving the structure

$$\gamma = \log \frac{I_s}{I_R}.$$

Attenuation constant and phase constant : The propagation constant is a complex quantity, *i.e.*, $\gamma = \alpha + j\beta$;

so
$$\frac{I_s}{I_R} = e^{\gamma} = e^{\alpha} e^{j\beta},$$

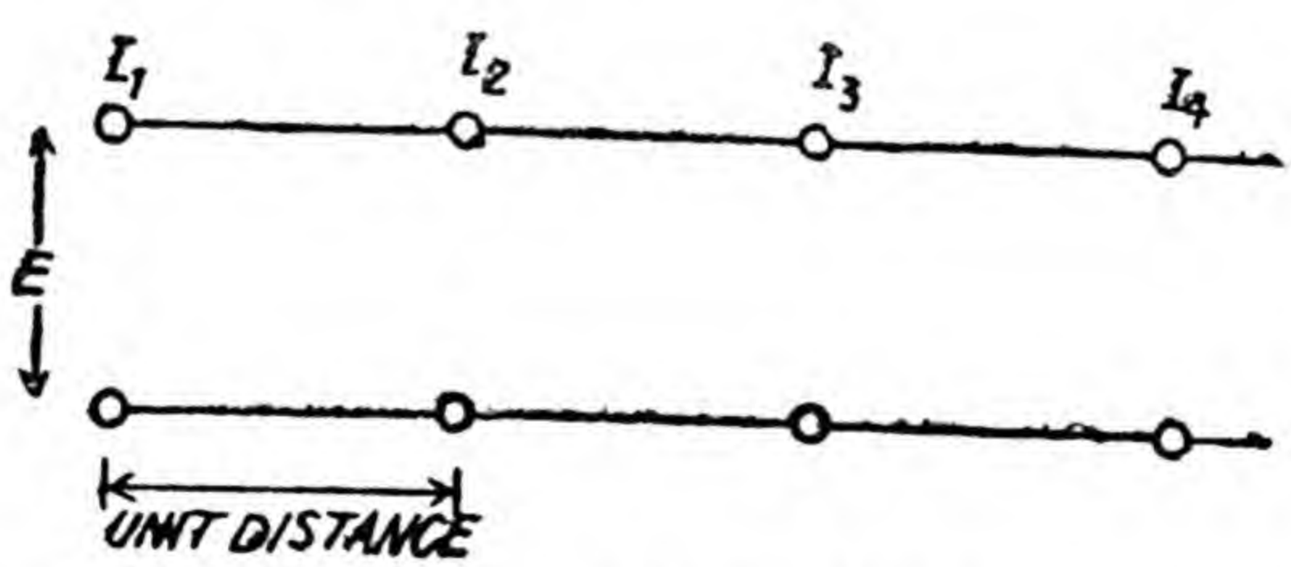


Fig. 4 (*a*) Transmission Line.

Thus if I_1 is the current entering the line in fig. 4 (*a*), the current I_2 at a unit distance farther along the line would be attenuated in magnitude by an amount α and retarded in phase by an amount β radians per unit length.

The locus of the current vector will be a long spiral as shown in the figure 4 (*b*) and in the absence of attenuation, figure will be a circle.

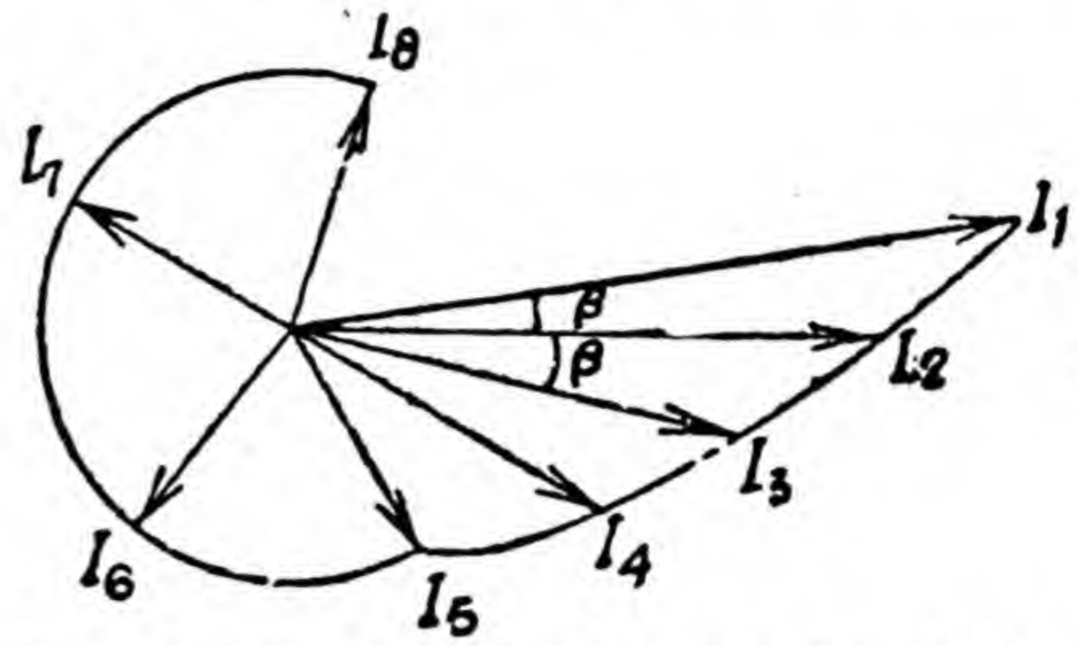


Fig. 4 (b). Locus of current vector.

Both incident and reflected waves suffer a change e^{α} in amplitude for unit length of transmission line, *where* α *is termed as attenuation constant as it represents the rate at which the incident wave decays along the line. Since in unit length the wave suffers a phase change of* β *radians, the latter is called the phase constant.*

If we consider incident wave only, then

$$E=E_s\, e^{-\gamma x},$$

$$I=I_s\, e^{-\gamma x}.$$

Putting $\gamma=\alpha+j\beta$, we write

$$E=E_s\, e^{-\alpha x}\, e^{-j\beta x},$$

$$I=I_s\, e^{-\alpha x}\, e^{-j\beta x}.$$

From these equations it is obvious that voltage and current become progressively smaller in magnitude because of the factor $e^{-\alpha x}$. Apart from this decrease in amplitude, phases of voltage and current also lag progressively as the distance x increases.

Time dependence of above equations of E and I is implicit because the sending end values will be functions of time, *viz.*

$$E_s=E_{s0}\, e^{j\omega t},$$

$$I_s=I_{s0}\, e^{j\omega t}.$$

This means E and I, *i.e.* voltage and current on the line are functions of both distance and time.

The wavelength λ of the wave travelling on the line is equal to the distance between successive crests of the wave at any moment, *i.e.*, it is equal to the change in x which makes the angle βx to increase by 2π radians. Therefore $\beta\lambda=2\pi$ or $\beta=2\pi/\lambda$. Since a phase change of 2π radians represents one cycle in time and involves a distance of one wavelength, then

$$x=vT$$

or $$\lambda=\frac{v}{f} \qquad \left(\text{putting } x=\lambda,\ T=\frac{1}{f}\right)$$

or $$\frac{2\pi}{\beta}=\frac{v}{f}$$

or $$\omega=2\pi f=\beta v$$

or $$v=\frac{\omega}{\beta}.$$

This is the *velocity of propagation along the line and is based on the observations of the phase on the line.*

Expression for phase constant β and attenuation constant (α) : The propagation constant γ is of the form

$$\gamma = \alpha + j\beta,$$

where α is the attenuation constant, *i.e.*, rate at which the incident wave decays along the line and β is the phase constant. γ is also equal to $\sqrt{\{(R+j\omega L)\ (G+j\omega C)\}}$.

So $$\alpha + j\beta = \sqrt{\{(R+j\omega L)\ (G+j\omega C)\}}. \quad \ldots(1)$$

$$\alpha - j\beta = \sqrt{\{(R-j\omega L)\ (G-j\omega C)\}}, \quad \ldots(2)$$

Multiplying (1) and (2),

$$\alpha^2 + \beta^2 = \sqrt{\{(R^2+\omega^2L^2)\ (G^2+\omega^2C^2)\}}. \quad \ldots(3)$$

From (1), on squaring,

$$\alpha^2 - \beta^2 + 2j\alpha\beta = (R+j\omega L)\ (G+j\omega C)$$

$$= (RG-\omega^2LC) + j\omega\ (RC+LG), \quad \ldots(4)$$

So $$\alpha^2 - \beta^2 = RG - \omega^2LC. \quad \ldots(5)$$

$$2\alpha\beta = \omega\ (RC+LG). \quad \ldots(6)$$

Adding (3) and (5), we get

$$2\alpha^2 = (RG-\omega^2LC) + \sqrt{\{(R^2+\omega^2L^2)\ (G^2+\omega^2C^2)\}},$$

$$\alpha = \pm[\tfrac{1}{2}\ (RG-\omega^2LC) + \tfrac{1}{2}\sqrt{\{\ R^2+\omega^2L^2)\ (G^2+\omega^2C^2)\}}]^{1/2}.$$

Subtracting (5) from (3), we get

$$2\beta^2 = (-RG+\omega^2LC) + \sqrt{\{(R^2+\omega^2L^2\ \ (G^2+\omega^2C^2)\}},$$

$$\beta = \pm[\tfrac{1}{2}(\omega^2LC-RG) + \tfrac{1}{2}\sqrt{\{(R^2+\omega^2L^2)\ (G^2+\omega^2C^2)\}}]^{1/2}.$$

26·2. LINE DISTORTION AND ATTENUATION :

For no distortion attenuation constant α *should be independent of frequency and phase constant* β *should vary linearly with frequency*. Suppose

(*a*) α is dependent on frequency and waveform which is the sum of a fundamental f and a third harmonic $3f$ is being transmitted down a line in the steady state. Then attenuation for the two frequencies will be different and the resultant waveform will be frequency distorted. Therefore a complex applied voltage, such as voice voltage containing many frequencies, will not have all frequencies transmitted with equal attenuation, and the received waveform will not be identical with the input waveform at the sending end. *Thus to avoid frequency distortion*, α *should be independent of frequency*.

(*b*) β does not vary linearly with frequency; then the time relationship of fundamental f and harmonic $3f$ will be destroyed and delay or phase distortion will be produced. Thus all frequencies applied to a transmission line will not have the same time of transmission, some frequencies being delayed more than the others. Consequently, applied voice voltage wave form atthe receiving end will not be identical to the waveform at the sending end.

From expressions for attenuation constant α and phase constant β, we see that α is dependent on frequency and β does not vary linearly with frequency. Their variations are shown in fig. 5(*a*) and fig. 5(*b*). Hence transmission line introduces both frequency and delay or phase distortion.

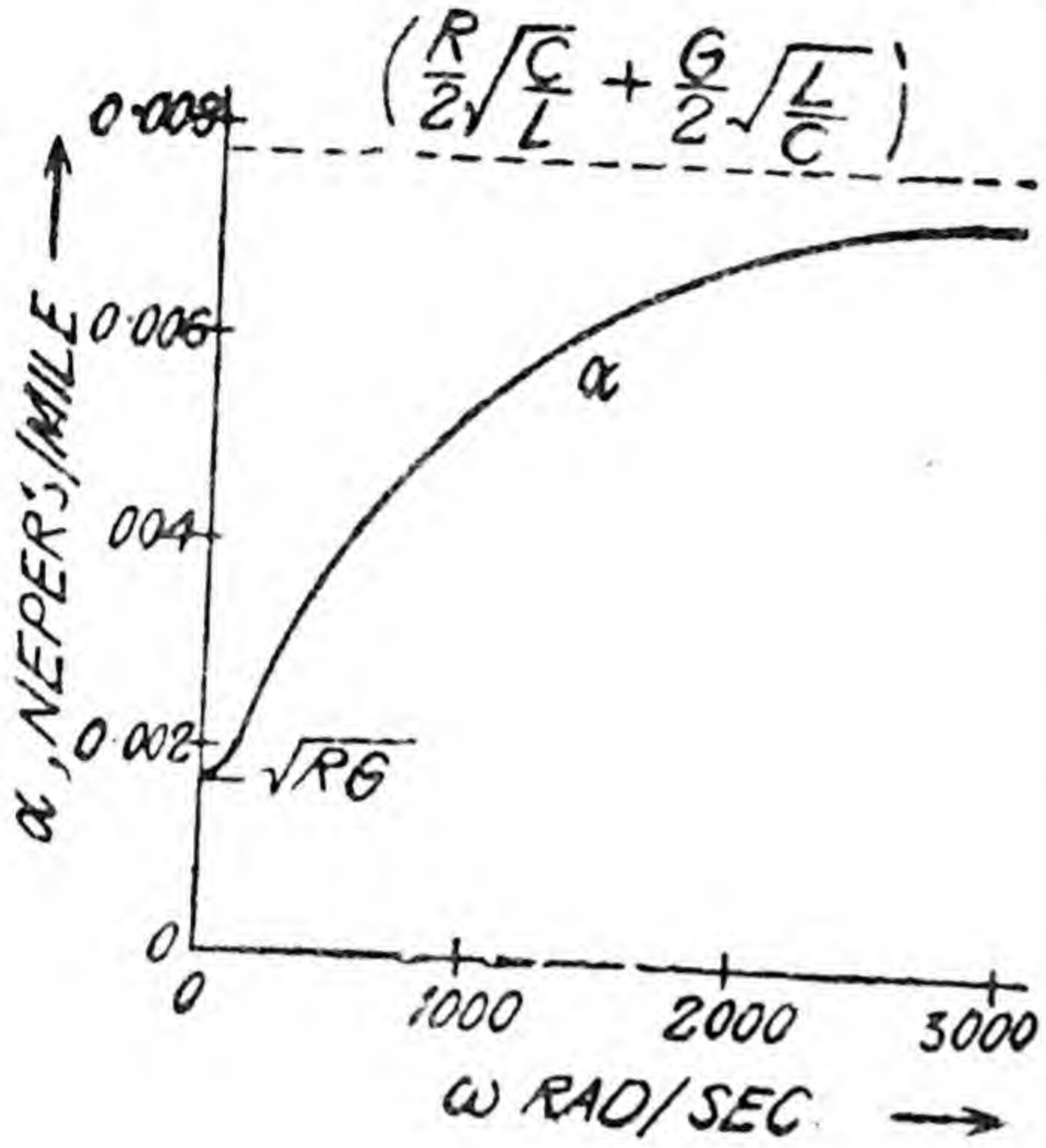


Fig. 5. (*a*)
Variation of attenuation constant as a function of ω for open wire line (composed of No. 10 copper conductor spaced 12 nches apart). The high frequency value (dotted line) is shown for comparison.

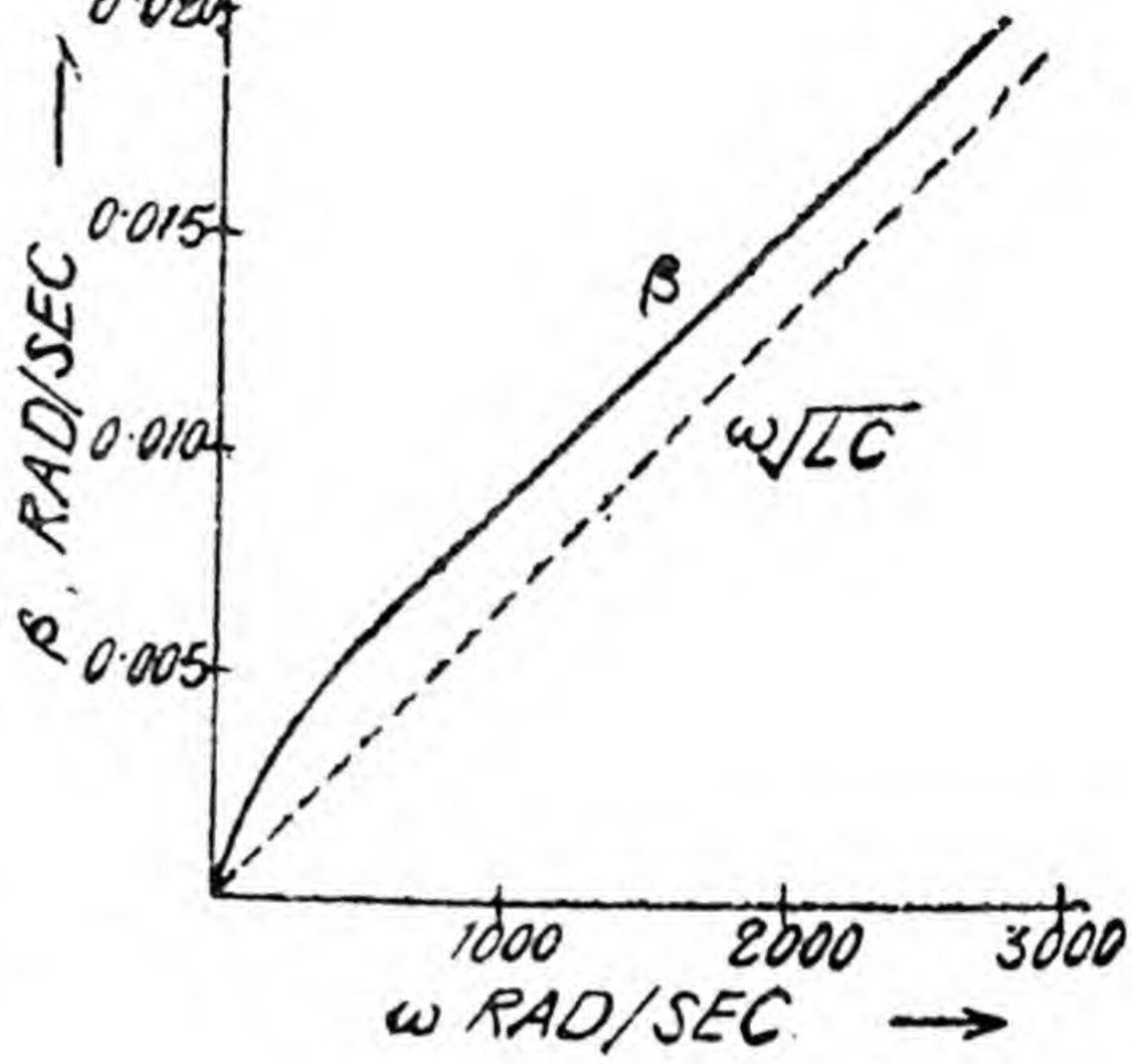


Fig. 5 (*b*)
Variation of phase constant with ω for the same line described in fig. 5 (*a*). The frequency asymptote $\omega\sqrt{(LC)}$ is plotted for comparison.

Condition for no distortion. The conditions for no distortion are ideal ones and cannot be achieved in practice. Even then we are to discuss such a possibility. It is obvious that for no distortion α should be independent of frequency. For the satisfaction of this condition there are two possibilities :

(*i*) Line resistance and conductance be zero, *i.e.*, $R=G=0$, so that from

$$\alpha^2=\tfrac{1}{2}\left[RG-\omega^2LC+\sqrt{\{(R^2+\omega^2L^2)\ (G^2+\omega^2C^2)\}}\right]$$
$$=0,$$

i.e., $\alpha=0,$

which implies zero attenuation. This condition is purely ideal because line will certainly have some resistance though conductance may be reduced to minimum (by using a perfect dielectric).

(*ii*) If the expression on solving the radical consists of a term ω^2LC then this term will be cancelled by other ω^2LC term outside the radical and thus the expression for α will be independent of ω.

i.e., if $\sqrt{\{(R^2+\omega^2L^2)\ (G^2+\omega^2C^2)\}}=K+\omega^2LC$

or $(R^2+\omega^2L^2)\ (G^2+\omega^2C^2)=(K+\omega^2LC)^2$

or $R^2G^2+\omega^4L^2C^2+\omega^2(R^2C^2+L^2G^2)=K^2+\omega^4L^2C^2+2K.\omega^2LC.$

Equating the coefficients of the power of ω. we get

$$R^2G^2=K^2$$

or $$R^2C^2+L^2G^2=2K.LC$$

or $$R^2C^2+L^2G^2-2.RG.LG=0$$

or $$(RC-LG)^2=0 \quad \text{or} \quad \frac{R}{G}=\frac{L}{C}$$

This is the condition for no distortion. Then the value of α will be

$$\alpha=\frac{1}{\sqrt{2}}[RG-\omega^2LC+\sqrt{\{(R^2+\omega^2L^2)(G^2+\omega^2C^2)\}}]^{1/2}$$

$$=\frac{1}{\sqrt{2}}\left[RG-\omega^2LC+RG.\sqrt{\left\{\left(1+\frac{\omega^2L^2}{R^2}\right)\left(1+\frac{\omega^2C^2}{G^2}\right)\right\}}\right]^{1/2}$$

Putting $\frac{L}{R}=\frac{C}{G}$, we get

$$\alpha=\frac{1}{\sqrt{2}}\left[RG-\omega^2LC+RG.\sqrt{\left\{\left(1+\frac{\omega^2C^2}{G^2}\right)\left(1+\frac{\omega^2C^2}{G^2}\right)\right\}}\right]^{1/2},$$

$$=\frac{1}{\sqrt{2}}\left[RG-\omega^2LC+RG.\left(1+\frac{\omega^2C^2}{G^2}\right)\right]^{1/2}$$

$$=\frac{1}{\sqrt{2}}\left[2RG-\omega^2LC+\omega^2C^2\,\frac{R}{G}\right]^{1/2}$$

Putting $\frac{R}{G}=\frac{L}{C}$, we get

$$\alpha=\frac{1}{\sqrt{2}}[2RG-\omega^2LC+\omega^2LC]^{1/2}$$

$$=\sqrt{(RG)}=\sqrt{\left(\frac{C}{L}\right)}.R. \qquad \text{...(1)}$$

Expression for β is

$$\beta=\frac{1}{\sqrt{2}}\left[\omega^2LC-RG+RG\left(1+\frac{\omega^2C^2}{G^2}\right)\right]^{1/2}$$

$$=\omega\sqrt{(LC)}. \qquad \text{...(2)}$$

Expression for characteristic impedance

$$Z_0=\sqrt{\left(\frac{R+j\omega L}{G+j\omega C}\right)}=\sqrt{\left(\frac{R}{G}\right)}.\sqrt{\left(\frac{1+j\omega L/R}{1+j\omega C/G}\right)}=\sqrt{\left(\frac{R}{G}\right)},$$

since $\frac{L}{R}=\frac{C}{G}$.

From expressions of α and β, we see that the conditions for no distortion are fully satisfied. Since Z_0, the characteristic impedance, is independent of frequency, it is purely resistive. The conditions, however, remain satisfactory if line constants R, G, L and C do not change with frequency. But in practice, because of skin

effect, they undergo changes and hence perfect freedom from distortion is not achieved.

For practical purposes, it would be better to discuss some conditions which can be achieved in actual line or cables, *i e.* conditions for low distortion.

Condition for Low Distortion. Let us consider the case where $\omega L \gg R$ and $\omega C \gg G$. Then

$$Z=R+j\omega L=j\omega L=\omega L \angle (\tfrac{1}{2}\pi),$$
$$Y=G+j\omega C=j\omega C=\omega C \angle (\tfrac{1}{2}\pi).$$

Though, for such a line, the values of R and G are so small that they will not affect the amplitude of Z and Y but still there will be a marked effect on the *phase* of Z and Y and hence instead of $\frac{1}{2}\pi$, it is resonable to write

$$Z=\omega L \angle \theta.$$
$$Y=\omega C \angle \phi,$$

where θ and ϕ represent the phase angles of Z and Y.

We know that angle between R and ωL will be $\tan\theta=\omega L/R$ so that $\cot\theta=R/\omega L$. Similarly, angle between G and ωC will be $\tan\phi=\omega C/G$ or $\cot\phi=G/\omega C$. Now when a certain angle, say ψ, is nearly equal to $\frac{1}{2}\pi$ then we can write

$$*\cos\psi=\cot\psi=\tfrac{1}{2}\pi-\psi.$$

Since the angles θ and ψ are nearly equal to $\frac{1}{2}\pi$, we can put

$$\theta=\tfrac{1}{2}\pi-\cot\theta$$
$$=\tfrac{1}{2}\pi-R/\omega L$$

and

$$\phi=\tfrac{1}{2}\pi-G/\omega C.$$

Therefore $\dagger Z=\omega L \angle (\frac{1}{2}\pi-R/\omega L)=\omega L e^{j(\pi/2-R/\omega L)}$,

$$Y=\omega C \angle (\tfrac{1}{2}\pi-G/\omega C)=\omega C e^{j(\pi/2-G/\omega C)}$$

Therefore propagation constant will be

$$\gamma=\sqrt{(Z.Y)}.$$
$$=\sqrt{\{\omega^2 LC.e^{j(\pi/2-R/\omega L)+j(\pi/2-G/\omega C)}\}}$$
$$=\omega\sqrt{(LC)}.e^{j/2\,[\pi-(R/\omega L+G/\omega C)]}$$
$$=\omega\sqrt{(LC)} \angle \left\{\tfrac{1}{2}\pi-\tfrac{1}{2}\left(\frac{R}{\omega L}+\frac{G}{\omega C}\right)\right\}.$$

But propagation constant $\gamma=\alpha+j\beta$, so that

$$\gamma=\alpha+j\beta$$
$$=\omega\sqrt{(LC)}\, e^{j\,[\pi/2-1/2\,(R/\omega L+G/\omega C)\,]}$$

---

*If you put $\psi=\pi/2$ then $\pi/2-\psi=0$ so that $\cos\psi=\cot\psi=0$ which is true.

†Let $R+j\omega L=A\cos\theta+jA\sin\theta=|A|e^{j\theta}$,

where $A=\sqrt{(R^2+\omega^2L^2)}$ $\theta=\tan^{-1}\dfrac{\omega L}{R}$

$$=\omega\sqrt{(LC)}\left[\cos\left\{\tfrac{1}{2}\pi-\tfrac{1}{2}\left(\frac{R}{\omega L}+\frac{G}{\omega C}\right)\right\}\right.$$
$$\left.+j\sin\left\{\tfrac{1}{2}\pi-\tfrac{1}{2}\left(\frac{R}{\omega L}+\frac{G}{\omega C}\right)\right\}\right]$$
$$=\omega\sqrt{(LC)}\left[\sin\tfrac{1}{2}\left(\frac{R}{\omega L}+\frac{G}{\omega C}\right)+j\cos\tfrac{1}{2}\left(\frac{R}{\omega L}+\frac{G}{\omega C}\right)\right].$$

But $\omega L \gg R$ and $\omega C \gg G$ and hence $(R/\omega L+G/\omega C)$ is a small quantity. We know that when angle is small $\sin\psi=\psi$ and $\cos\psi=1$. Thus

$$\alpha+j\beta=\omega\sqrt{(LC)}\left[\tfrac{1}{2}\left(\frac{R}{\omega L}+\frac{G}{\omega C}\right)+j.1\right]$$

Equating real and imaginary parts, we get

$$\alpha=\omega\sqrt{(LC)}\left(\frac{R}{\omega L}+\frac{G}{\omega C}\right)=\tfrac{1}{2}\left[R\sqrt{\left(\frac{C}{L}\right)}+G\sqrt{\left(\frac{L}{C}\right)}\right]$$

and $$\beta=\omega\sqrt{(LC)}.$$

So characteristic impedance will be

$$Z_0=\sqrt{\left(\frac{Z}{Y}\right)}=\sqrt{\left[\frac{\omega L.\epsilon^{j(\pi/2-R/\omega L)}}{\omega C\,\epsilon^{j(\pi/2-G/\omega C)}}\right]}$$
$$=\sqrt{\left(\frac{L}{C}\right)}\,\angle[e^{j(G/\omega C-R/\omega L}]$$
$$=\sqrt{(L/C)}\,\angle[\tfrac{1}{2}\,(G/\omega C-R/\omega L)]\text{ radians.}$$

Here too we see that α, the attenuation constant, is independent of frequency and β varies linearly with frequency. Z_0, the characteristic impedance, is also independent of frequency and the angle is very small, *i.e.* we can take it as nearly resistive. Now taking $|Z_0|=\sqrt{(L/C)}$, and putting in expression of α, we get

$$\alpha=\tfrac{1}{2}\left(\frac{R}{|Z_0|}+G|Z_0|\right)$$

If G is negligible,

$$\alpha=\tfrac{1}{2}\,R/|Z_0|.$$

Achievement of the Condition of Low Distortion : It is obvious that α depends upon all the four line constants R, C, L and G. Therefore with reference to their values, possibilities of low distortion can be discussed. Since

$$\alpha=\tfrac{1}{2}\{R\sqrt{(C/L)}+G\sqrt{(L/C)}\}.$$

It can be seen that if here too we apply the condition for no distortion, *i.e.*, $L/C=R/G$, then

$$\alpha=\tfrac{1}{2}\{R.\sqrt{(G/R)}+G\sqrt{(R/G)}\}=\sqrt{(RG)}=R\sqrt{(C/L)},$$

which is the same as obtained in the case of distortionless line. Therefore, to have low distortion, we should decrease the value of R and C and increase that of L. Now to reduce R, we require large conductors and to reduce C we require an increase of spacing between the conductors. But these result in an increase of cable

size and cost. The best approach to improve the cable response is to increas the value of inductance for which the line can be wrapped uniformly with a tape of some ferrous material such as iron although this will be an expensive treatment. The way is used for the cables which remain immersed in water. For the land cables, inductance can be added in lumps at regular intervals. We cannot increase inductance indefinitely because alongwith inductance, resi tance of the line also increases.

Lines of Low Loss : In the conditions of low distortion, we have completely neglected R. But inorder to be close to practical conditions, one must account for R. Therefore in this condition, we shall make due consideration of R, neglecting G. Since R is dissipative factor, this conditon is termed as the condition of *low loss*. If the frequency of operation is very high, then $\omega L \gg R$ and $\omega C \gg G$ and we can write

$$Z \approx j\omega L,$$
$$Y \approx j\omega C$$

and hence propagation constant.

$$\gamma = \alpha + j\beta = \sqrt{Z.Y} = \sqrt{(j^2\omega^2 LC)} = j\omega\sqrt{(LC)};$$

so

$$\alpha = 0,\ \beta = \omega\sqrt{(LC)}$$

and

$$Z_0 = \sqrt{\left(\frac{Z}{Y}\right)} = \sqrt{\left(\frac{L}{C}\right)}$$

Thus line is lossless since $\alpha = 0$; further β varies linearly with frequency and hence no phase distortion ; characteristic impedance is purely resistive.

If we do not consider such high frequencies then in that case approximation can be limitted so as to neglect conductance only. Hence

$$Z = R + j\omega L,$$
$$Y = j\omega C,$$

though ωL will increase rapidly as compared to R which increases as the square root of frequency because of the skin effect.

The propagation constant will be

$$\gamma = \alpha + j\beta = \sqrt{(ZY)} = \sqrt{\{(R + j\omega L)\, j\omega C\}}$$

$$= j\omega\sqrt{(LC)}.\sqrt{\left(1 - j\frac{R}{\omega L}\right)}$$

$$= j\omega\sqrt{(LC)}.\left[1 - j\frac{R}{2\omega L} + \frac{1}{8}\left(\frac{R}{\omega L}\right)^2 + j\frac{3}{48}\left(\frac{R}{\omega L}\right)^3 + \ldots\right].$$

Since $\omega L \gg R$, the term $R/\omega L$ is small and higher power terms can be neglected.

$$\alpha+j\beta=\frac{R}{2}\sqrt{\left(\frac{C}{L}\right)}\left[1-\frac{1}{8}\left(\frac{R}{\omega L}\right)^2\right]+j\omega\sqrt{(LC)}\left[1+\frac{1}{8}\left(\frac{R}{\omega L}\right)^2\right]$$

So
$$\alpha=\frac{R}{2}\sqrt{\left(\frac{C}{L}\right)}\left[1-\frac{1}{8}\left(\frac{R}{\omega L}\right)^2\right],$$

$$\beta=\omega\sqrt{(LC)}\left[1+\frac{1}{8}\left(\frac{R}{\omega L}\right)^2\right]$$

As α depends upon ω and β does not vary linearly with ω, conditions of low loss do not fit the conditions of low distortion. If, however, R be neglected completely, then conditions of lossless line coincide with those of distortionless line.

In ordinary telephone cable, wires are insulated with paper ($G=0$) and twisted in pair ($L=0$), so that $Z=R$, $Y=j\omega C$ giving

$\gamma=\sqrt{(j\omega CR)}=\sqrt{(\omega CR)}\,(\cos 45°+j\sin 45°)$, *i.e.* $\alpha=\sqrt{\left(\frac{\omega CR}{2}\right)}=\beta$.

so that
$$Z_0=\sqrt{\left(\frac{Z}{Y}\right)}=\sqrt{\left(\frac{R}{j\omega C}\right)}=\sqrt{\left(\frac{-jR}{\omega C}\right)}=\sqrt{\left(-\frac{Re^{j\pi/2}}{\omega C}\right)}$$

$$=\sqrt{\left(\frac{-R}{\omega C}\right)}.e^{j\pi/4}$$

and velocity of propagation $=\frac{\omega}{\beta}=\sqrt{\left(\frac{2\omega}{RC}\right)}$.

Ex. *A cable pair is loaded with* 6 *mH coils, at intervals of* 0·9 *km. for operation at frequencies upto* 16 *kc/s. At this frequency, the primary constants of the cable are :*

$$R=72\ ohm,\ L=1{\cdot}0\ mH,\ C=0{\cdot}065\ \mu F,\ G=50\ \mu\ mho,$$

all per loop kilometer and the effective resistance of each loading coil is 4·5 *ohm. Estimate the approximate attenuation per mile of the loaded cable.*

The attenuation constant is given by

$$\alpha=\frac{R}{2}\sqrt{\left(\frac{C}{L}\right)}+\frac{G}{2}\sqrt{\left(\frac{L}{C}\right)}$$

where
$R=72+$effective resistance of each loading coil per km.

$$=72+\left(\frac{4{\cdot}5}{0{\cdot}9}\right)$$

$=77$ ohms.

$L=1{\cdot}0+$inductance of loading coil taken per km.

$$=1{\cdot}0+\left(\frac{6}{0{\cdot}9}\right)$$

$=7{\cdot}67\ mH=7{\cdot}67\times10^{-3}\ H.$

$$G = 50 \times 10^{-6} \text{ mho}$$
$$G = 0{\cdot}065 \times 10^{-6} \text{ Farad}$$

Putting these values in the expression for attenuation constant, we get

$$\alpha = \frac{77}{2}\sqrt{\left(\frac{0{\cdot}065 \times 10^{-6}}{7{\cdot}67 \times 10^{-3}}\right)} + \frac{50 \times 10^{-6}}{2}\sqrt{\left(\frac{7{\cdot}67 \times 10^{-3}}{0{\cdot}065 \times 10^{-6}}\right)}$$
$$= 0{\cdot}1121 + 0{\cdot}0086$$
$$= 0{\cdot}1207 \text{ neper} \quad \text{or} \quad 1{\cdot}048 \text{ } db.$$

26·3. LINE TERMINATION :

The purpose of transmission line is to transfer the maximum of energy, fed to it, to the load. The conditions required for maximum transference are fundamentally based on two factors : the first factor is that the line attenuation should be minimum and has been discussed in last article ; the second factor is the termination of the line with a load of proper value. Thus the value and nature of the load largely affects the basic function of transmission line. We shall discuss the line termination with different loads, *e g.* zero load (short-circuited), infinite load (open circuited), complex load, etc. The important of this discussion is to know that what are the voltages and currents at different points of the line and how the input impedance of the line varies with different loads.

We know that reflection coefficient is

$$|\rho| = \frac{Z_R - Z_0}{Z_R + Z_0} = \frac{\text{reflected voltage}}{\text{incident voltage'}}$$

so that when load impedance Z_R equals Z_0, the characteristic impedance reflection, coefficient, ρ, is zero, *i.e.*, reflected voltage is zero, which means that whole of the energy incident on the load is absorbed by it and no energy is reflected back. *Thus if a line is terminated by characteristic impedance, then best matching conditions are obtained since whole of the input energy is transferred to the load.* For such a matching, characteristic impedance being dependent upon line constants is fixed for a line, load impedance is to be controlled. However, in most of the cases, it is not possible to control the load impedance. To obtain matching conditions in such cases, stub line is connected in parallel to the main line. This way of matching will also be discussed towards the end of this article.

Furthermore, greater is the difference between load impedance Z_R, and characteristic impedance, Z_0, greater will be the magnitude of reflected voltage, *i e.*, reflected wave is then sufficiently strong and the same amount of attenuation will be caused by the reflected wave while travelling on transmission line as is caused by the

incident wave. Thus to avoid losses due to attenuation, suppression of reflected wave is almost necessary.

In the discussion of the termination of a transmission line, we shall analyse a purely lossless line, *i.e.*, for which $\alpha=0$. The lines not terminated by characteristic impedance are called resonant lines while those terminated in characteristic impedance are called non-resonant lines. At radio frequencies $\omega L \gg R$ and $\omega C \gg G$ and therefore, we can take $R=G=0$ approximately. When $R=G=0$, attenuation constant α is zero. Thus we are, infact, considering here the *lines at radio frequencies*.

Since we are to discuss conditions on the line with reference to load end, it will be most convenient to measure the distance from the load end, *i.e.*, equations of voltage and current, which refer to load end as the origin of coordinates, will be employed.

(i) Line termination by zero load, *i.e.* short circuited line ($Z_R=0$) : Consider a line short-circuited at the load end and connected, at the sending end, to a generator of internal impedance Z_0, equal to the characteristic impedance of the line. Since the line is terminated by an impedance other than Z_0, there will be a reflected wave *i.e.* resonant line.

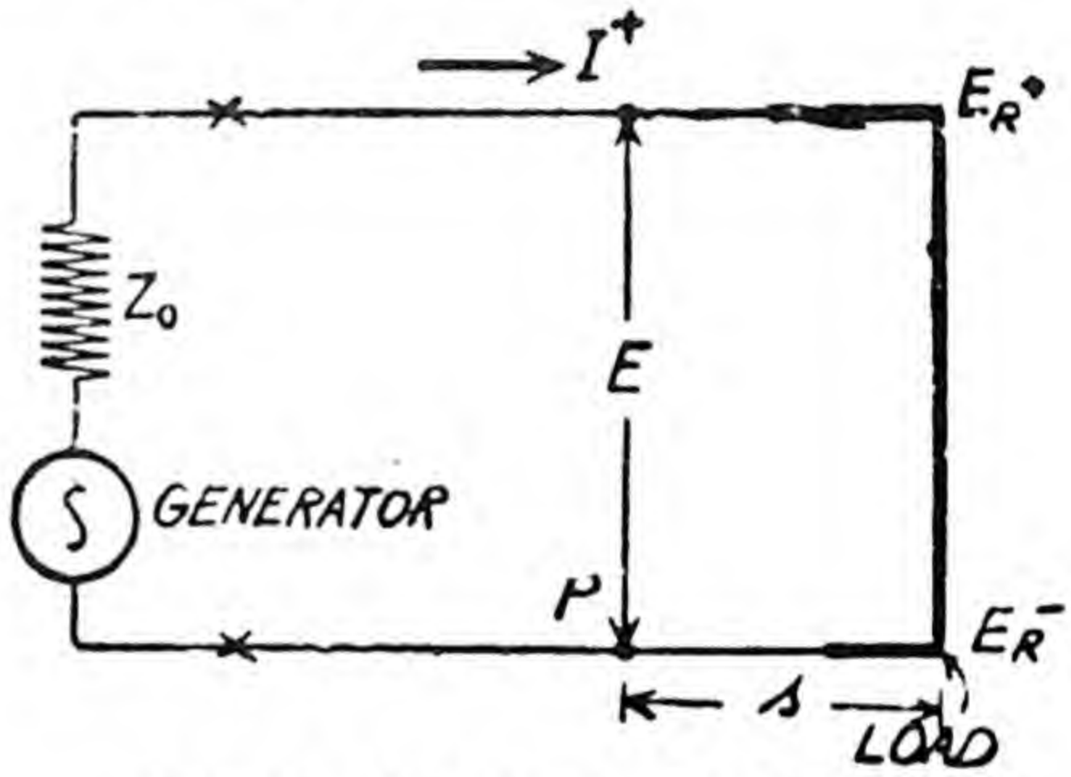


Fig. 6. Short-circuited line.

Since at input terminals the load is Z_0, the line will behave with respect to reflected wave exactly like a line terminated in characteristic impedance and hence *steady state* conditions are reached as the reflected wave reaches input terminals.

If E_R, I_R, Z_R refer to load end, then obviously $Z_R=0$ and hence from $Z_R=E_R/I_R$. $E_R=0$† That is, at short-circuited end *voltage must be zero* which means the voltage developed at the load due to incident wave and that due to reflected wave must be equal in magnitude and opposite in polarity. Therefore, if we denote E_R^+ as the voltage at the load due to incident wave and E_R^- that due to reflected wave, then

$$E_R^+ = -E_R^-$$

---

†At the load end $s=0$. So that $E_R=E_R^+.e^{j\beta s}=E_R^+$. Now $E_R=0$ because $Z_R=0$, I_R is thus not zero. Further $I_R=I_R^+ + I_R^-$ at the load end and reflection coefficient is $|\rho| = \dfrac{Z_R - Z_0}{Z_R + Z_0} = -1 = 1.e^{j\cdot 180}$, but for currents, since $I_R \neq 0$, there is no phase reversal and hence $I_R^+ = |\rho| I_R^- = I_R^-$

The conclusion follows from calculation of reflection coefficien which in this case is -1, so that $E_R^+ = -E_R^-$.

Since the current is not zero, hence they will not be of opp-site polarity *i.e.*,

$$I_R^+ = I_R^-.$$

The above conclusion can also be reached at as follows : The energy in the electromagnetic field is equally divided between the dielectric field (electrostatic field) and the magnetic field, that is, $(\frac{1}{2}\epsilon_0 E^2 + \frac{1}{2}\mu_0 H^2)$. With the voltage at the shorted end reduced to zero, the electrostatic field collapses. This energy is not lost nor can it be dissipated in the short circuit. There is only one possible explanation—all energy must now be in the magnetic field. In order for the magnetic field strength to increase, the current at the shorted end must increase. This implies that the current wave must reflect back in *the same phase* as it would have continued.

Now we consider a point P at a distance s from the load end. Let voltage due to incident wave at P be E^+ and that due to reflected wave be E^-. Then in travelling to a distance, s, from P to the short circuit end, the incident wave is delayed by a phase βs, *i.e.*,

$$E^+ = E_R^+ e^{j\beta s}.$$

Similariy, reflected wave, in returning to point P, is delayed by an equal amount,

i.e.,

$$E^- = E_R^- e^{-j\beta s}.$$

Thus total voltage at P

$$E = E^+ + E^- = E_R^+ e^{j\beta s} + E_R^- e^{-j\beta s},$$

$$= E_R^+ (e^{j\beta s} - e^{-j\beta s})$$

$$= j.2E_R^+ \sin \beta s. \qquad \ldots(1)$$

Similarly, the current at P,

$$I = I^+ + I^- = I_R^+ e^{j\beta s} + I_R^- e^{-j\beta s}$$

$$= I_R^+ (e^{j\beta s} + e^{j\beta s})$$

$$= 2I_R^+ \cos \beta s = 2 \frac{E_R^+}{Z_0^*} \cos \beta s. \qquad \ldots(2)$$

Occurrence of j in equation (1) and absence in eqn. (2) shows

---

•We know that

$$\frac{E}{I} = \frac{Z_0 [Ae^{\gamma x} + Be^{-\gamma x}]}{Ae^{\gamma x} - Be^{-\gamma x}}.$$

For *incident* wave at the load $x=0$, so that

$$\frac{E_R^+}{I_R^+} = Z_0 \frac{Ae^{\gamma \cdot 0}}{Ae^{\gamma \cdot 0}} = Z_0.$$

that E and I will be always 90° out of phase and E is leading I by 90°. Now we shall plot phases of E and I, amplitudes of E and I, and voltage·current additions (complex number diagram) at different points of the line (for different s values, being measured from load end).

To plot $|E|$ *and* $|I|$: From equation (1),

$$E=2jE_R^+ \sin \beta s.$$

So

$\beta s=0$,	*i.e.*,	$s=0$,	$E=0$,
$\beta s=\frac{1}{4}\pi$,		$s=\frac{1}{8}\lambda$,	$E=j\sqrt{2}\,E_R^+$,
$\beta s=\frac{1}{2}\pi$		$s=\frac{1}{4}\lambda$,	$E=j2\,E_R^+$,
$\beta s=3\pi/4$,		$s=3\lambda/8$,	$E=j\sqrt{2}\,E_R^+$,
$\beta s=\pi$,		$s=\frac{1}{2}\lambda$,	$E=0$.
... ...			

From equation (2),

$$I=2\,\frac{E_R^+}{Z_0}\cos \beta s,$$

$\beta s=0$,	*i.e.*,	$s=0$,	$I=2I_R^+$
$\beta s=\frac{1}{4}\pi$,		$s=\frac{1}{8}\lambda$,	$I=\sqrt{2}I_R^+$
$\beta s=\frac{1}{2}\pi$,		$s=\frac{1}{4}\lambda$,	$I=0$,
$\beta s=3\pi/4$,		$s=3\lambda/8$	$I=-\sqrt{2}\,I_R^+$
$\beta s=\pi$,		$s=\frac{1}{2}\lambda$,	$I=-2I_R^+$

To plot phases of E and I : We know that from $\beta s=0$ to 180° sine does not change its sign and hence E will not change its phase till 180°. But cosine changes at $\beta s=90°$ so upto $\beta s=0$ to 90° phase of I will be same but after 0° its phase will be reversed, *i.e.*, a phase change of 180° till $\beta s=270°$. Further initially E is leading I by 90° and hence the plot (fig. 7).

Complex Number Diagram : With knowledge of amplitude and phase, complex number diagram can be plotted. For example, at $\beta s=\frac{1}{4}\pi$, $E=\sqrt{2}E_R^+$, for which it is necessary that E^+ and E^- should be inclined at 45° with resultant E etc.

It should be noted here that upto quarter wavelength of the line ($s=\frac{1}{4}\lambda$) voltage is leading current which is an inductive effect in which voltage leads the current.

The input impedance of this line will be

$$Z_s=\frac{E^+}{I^+}=\frac{j\,2E_R^+ \sin \beta s}{2E_R^+/Z_0 \cos \beta s}=j.Z_0 \tan \beta s.$$

Since Z_s does not contain any real part, the input impedance of short-circuited line is purely reactive. The plot of Z_s with βs is

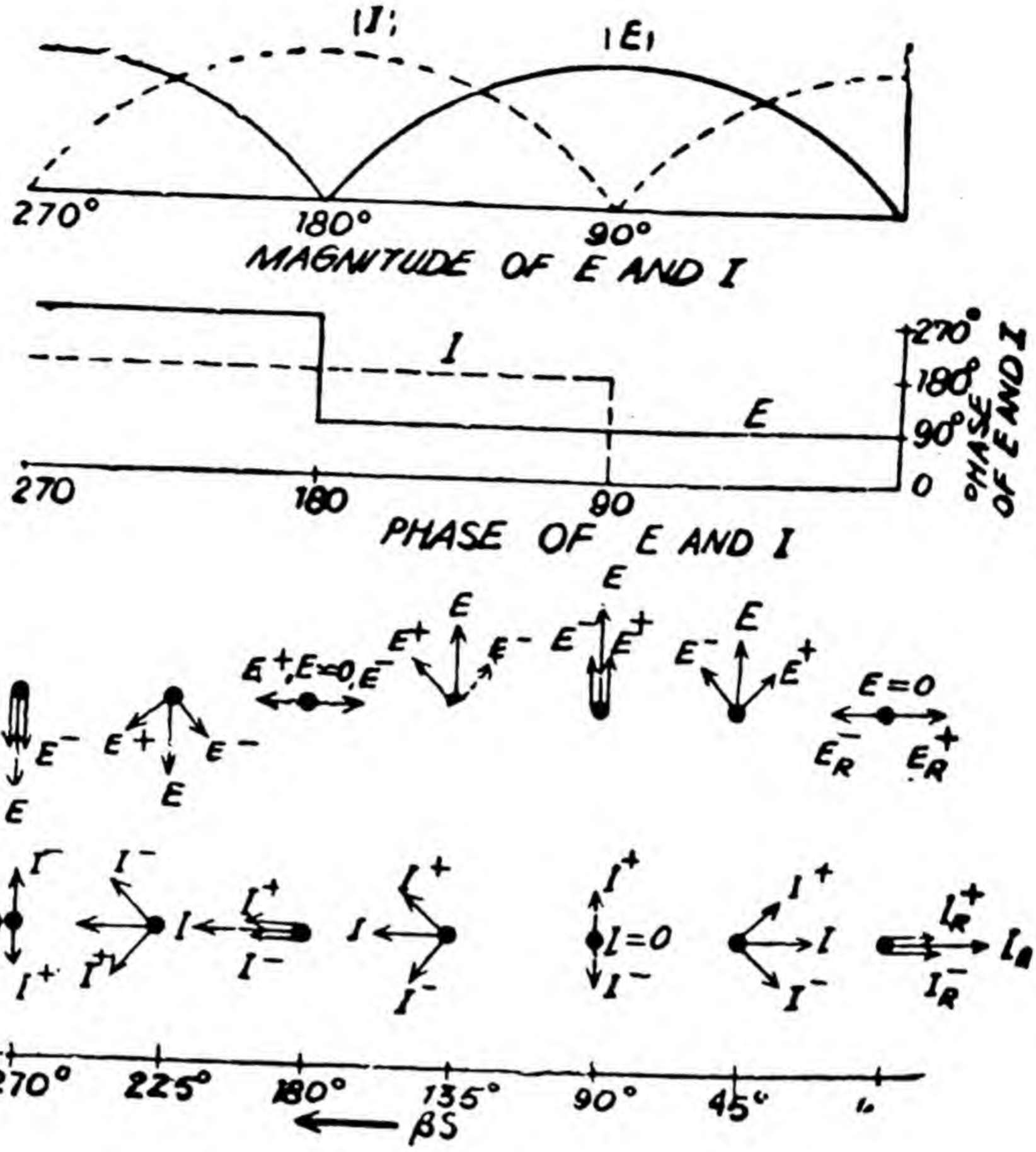


Fig. 7. Variations in the magnitude and phase of voltage and current on a shorted end line.

shown in fig. 8. From phase diagram, E leads I from $\beta s=0$ to 90° or $s=0$ to $\frac{1}{4}\lambda$ and therefore reactance is inductive. But from $\beta s=90°$ to 180°, *i.e.*, $s=\frac{1}{4}\lambda$ to $\frac{1}{2}\lambda$, current I leads voltage E *i.e.*, a capacitive effect and hence reactance is capacitive and so on.

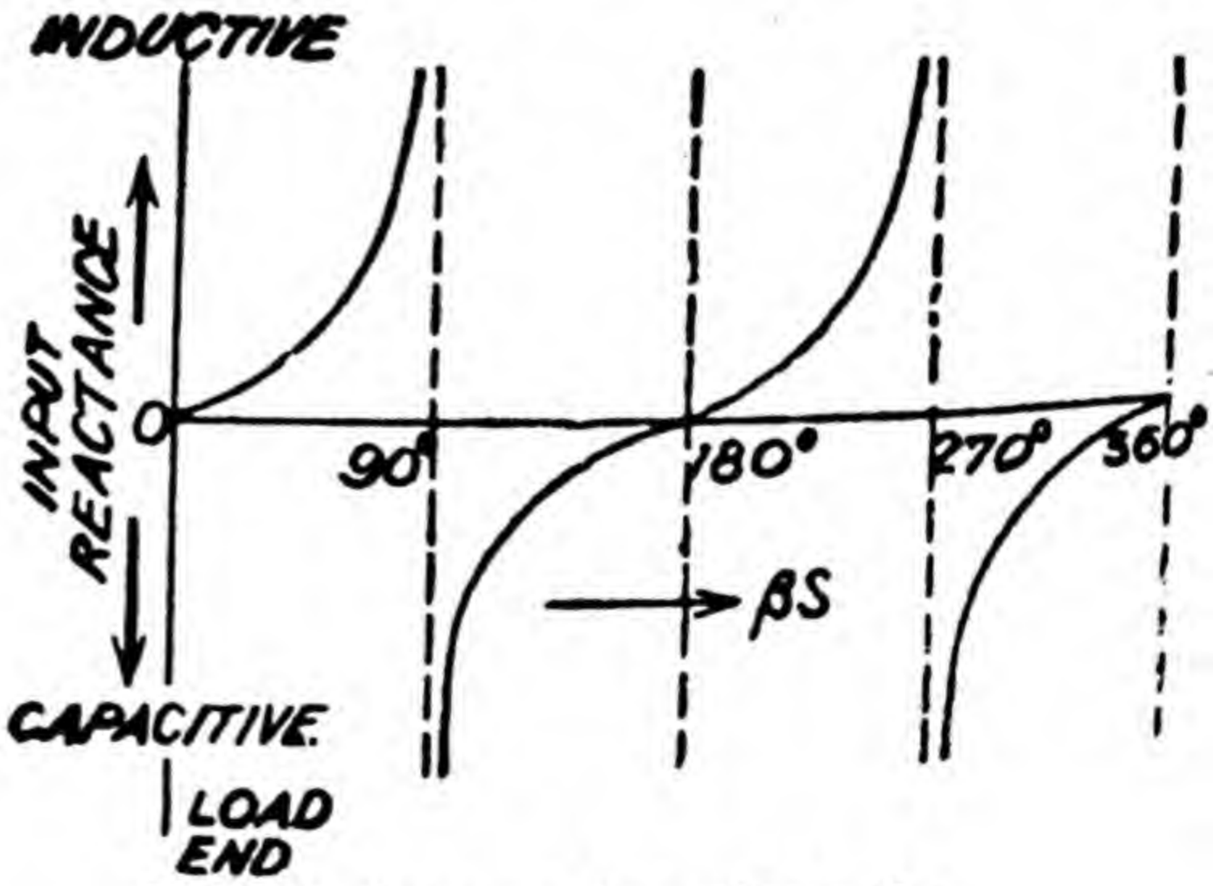


Fig. 8. Z_s as a functions βs.

We see that minimum voltage E_{min} on the line is zero. If we define a quantity, S, to be the ratio of maximum voltage to the minimum volage on the line, then

$$S=\frac{E_{max.}}{E_{min}}=\infty.$$

For such a line, S is termed as *standing wave ratio.*

(ii) Line termination by infinite impedance, i.e. open circuited line ($Z_R=\infty$) : Line is shown in fig. 9. Since load is different from Z_0, a reflected wave will be produced. The internal resistance of the generator is assumed to be Z_0. This assumption is made for steady standing wave pattern to exist on the line.

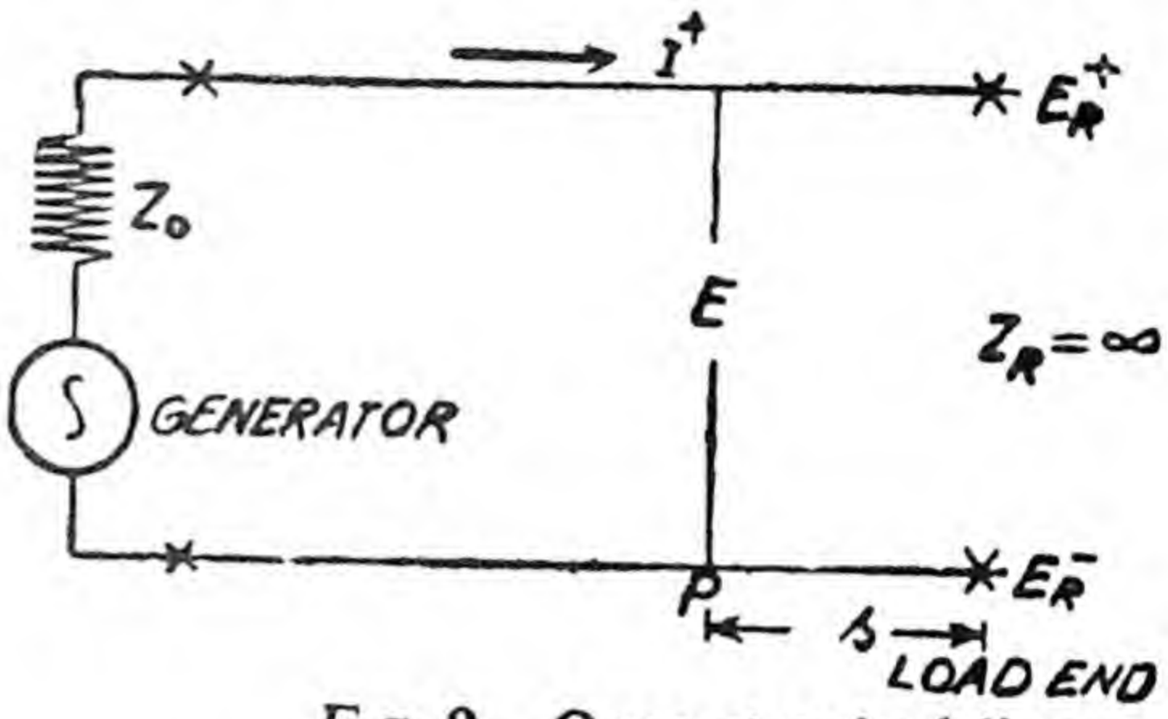


Fig. 9. Open circuited line.

At load $\frac{E_R}{Z_R}=I_R$.

Since $Z_R=\infty$, I_R is zero. Therefore currents at the load due to reflected and incident waves will be same in magnitude but opposite in polarity. Hence here,

$$I_R^+=-I_R \ .$$

But voltages $E_R^+=E_R^-$; we can reach at this conclusion as follows : The energy in the electromagnetic field is $(\frac{1}{2}\epsilon_0 E^2+\frac{1}{2}\mu_0 H^2)$. With current at the open end reduced to zero, the magnetic field collapses. This energy is not lost, nor can it be dissipated in open circuit. Therefore, the whole energy must now be in electrostatic field. In order for the electrostatic field strength to increase, the voltage wave must reflect back in the *same phase* from open end and hence at point P, .

$$E=E^+ + E^- = E_R^+ e^{j\beta s} + E_R^- e^{-j\beta s}$$

$$=2E_R^+ \cos \beta s \qquad \text{.. (1)}$$

and

$$I=I^+ + I^- = I_R^+ e^{j\beta s} - I_R^- e^{-j\beta s}$$

$$=2jI_R^+ \sin \beta s = 2j\frac{E_R^+}{Z_0} \sin \beta s. \quad \text{...(2)}$$

Thus, as compared to previous case, the conditions of voltage and current are interchanged, *i.e.*, which was true for voltage in previous case is now true for current and the plots of magnitude and phase of E and I will be as shown in fig. 10.

$$Z_s=\frac{E}{I}=-jZ_0 \cot \beta s.$$

When Z_s is plotted as a function βs, then the curves of the form shown in the fig. 11, are obtained.

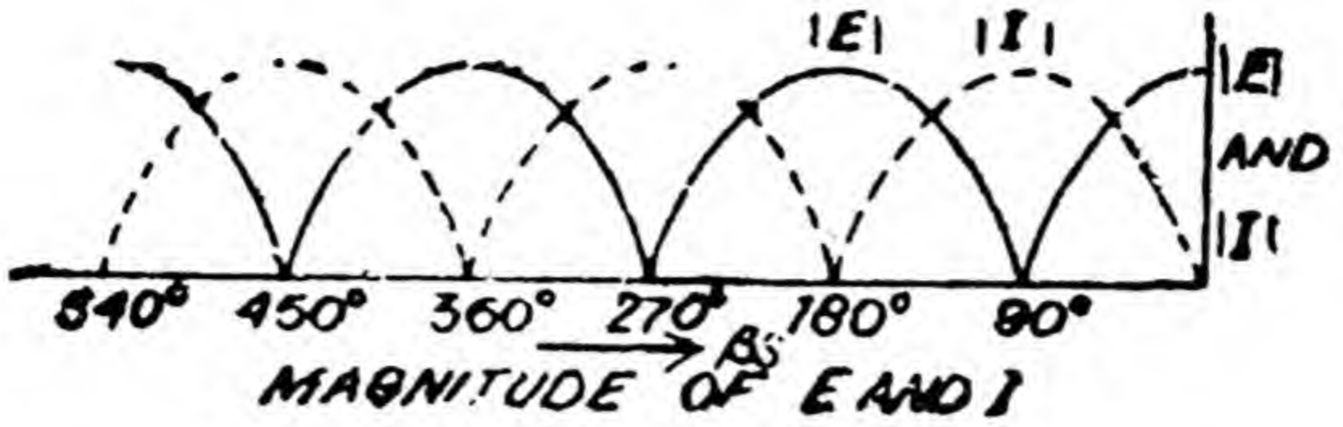


Fig. 10. Magnitudes and phases of E and I on an open end line

It should be noted from the phase diagram that from $\beta s=0$ to 90°, *i.e.* upto $s=\frac{1}{4}\lambda$, I leads E, *i.e*, capacitive effect and hence Z_s is capacitive : but from $s=\frac{1}{4}\lambda$ to $\frac{1}{2}\lambda$, E leads I and Z_s is inductive and so on. It is quite obvious that input impedance of an open circuited line is purely reactive.

The value of standing ratio, S, will be again infinite for such a line since $E_{min.}$ is zero.

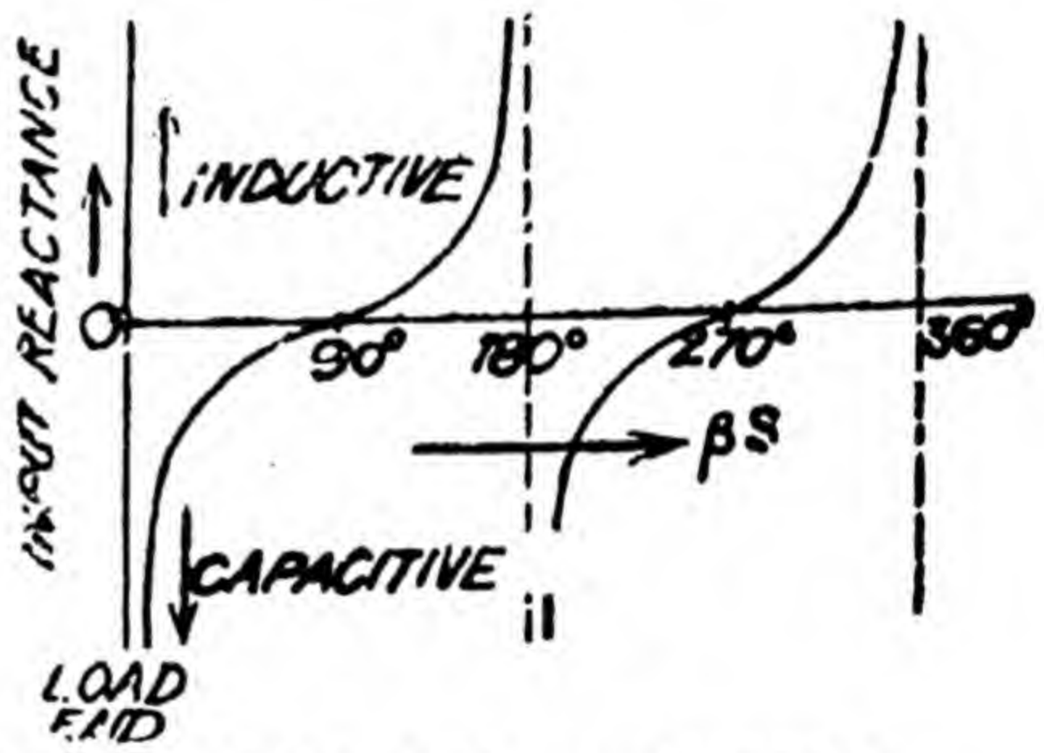


Fig. 11. Plot of Z_s as a function of βs.

Example : What length of line is needed and how should it be terminated to act as a capacitance of $2pF$ at 500 mc/s. The line impedance (Z_0) is 30 ohms.

For open circuited lines, input impedance offered by the line is

$$Z_s=Z_o \cot \beta s.$$

It is given that

$$Z_s=\text{capacitive reactance}$$

$$=\frac{1}{2\pi f X_c}=\frac{1}{6{\cdot}28\times 5\times 10^8\times 10^{-12}}$$

$$=159 \text{ ohms.}$$

Therefore,

$$159=300 \cot\left(\frac{2\pi}{\lambda}\cdot s\right)$$

or $$\frac{2\pi}{\lambda}.\ s=\tan^{-1}\frac{300}{159}$$

$$=\tan^{-1}1.9$$

$$\frac{360^\circ\ s}{\lambda}=62{\cdot}3^\circ$$

so that $$s=\frac{62{\cdot}3}{360}\lambda$$

But $$\lambda=\frac{3{\cdot}10^{10}}{500{\cdot}10^{6}}=60\text{ cm.}$$

Therefore line length is

$$s=\frac{62{\cdot}3}{360}\times\lambda=0\ 173\lambda=10{\cdot}38\text{ cm.}$$

Similarly for short circuited lines we can show that

$$s=0{\cdot}423\lambda$$

$$=25{\cdot}38\text{ cm.}$$

Problem 1. *For low loss transmission line section which are much shorter than $\lambda/4$, show that the input reactance can be represented by*

$$Z_S=\omega Ls \qquad \textit{when line is shorted}$$

and $$Z_S=1/\omega Cs, \qquad \textit{when it is opened,}$$

where $L=Z_0/v$ and $C=1/Z_0v$ are the inductances and capacitances respectively, v is the velocity of the wave on the line.

For short circuited line sections :

Refer to fig. 8. For $\beta s \lessdot 90^\circ$, or $s<\lambda/4$, input reactance is inductive. It is given by

$$Z_S=jZ_0\tan\beta s$$

when $s \ll \lambda/4$, $\beta s \ll 90^\circ$ or $\beta s=0^\circ$, we have

$$Z_S=jZ_0\,\beta s$$

$$=jZ_0\frac{2\pi}{\lambda}.\ s$$

$$=j.\frac{Z_0}{v}.\frac{2\pi v}{\lambda}.\ s$$

$$=j.\frac{Z_0}{v}.\ 2\pi f.s=j\omega Ls.$$

or $$|\ Z_s\ |=\omega Ls.$$

For open circuited line sections :

Refer to fig. 11. For $\beta s \ll 90^\circ$ or $s<\lambda/4$, input reactance is capacitive. It is given by

$$Z_s=-jZ_0\cot\beta s$$

$$=-j\frac{Z_0}{\tan\beta s}$$

$$=-j.\frac{Z_0}{\beta s} \qquad (\text{when } \beta s \ll \lambda/4)$$

$$= -j.\frac{Z_0.\lambda}{2\pi s}$$

$$= -j.\ Z_0 v.\ \frac{\lambda}{2\lambda v}.\ \frac{1}{s}$$

$$= -j.\ \frac{1}{C}.\ \frac{1}{2\pi f}.\ \frac{1}{s} = -\frac{j}{\omega C s}$$

or $$|\ Z_s\ | = \frac{1}{\omega C s}.$$

Problem 2. *State the type of transmission line termination by which complete reflection is produced ? Is it possible to replace reactive load by fictitious open and short circuited extension of the line ? If so, find the equivalent short circuit extension of the reactive load of $j.600$ and equivalent open circuit extension of the reactive load of $-j.500$. What is the use of finding such eauivalent extensions ?*

Following terminations produce complete reflection

(*i*) short circuited, $|\ \rho\ | = 1$, $S = \infty$ $(P = -1)$

(*ii*) open circuited, $|\ \rho\ | = 1$, $S = \infty$ $(P = +1)$

(*iii*) purely inductive (I will lag E by 90° so $P = EI \cos 90 = 0$)

(*iv*) purely capacitive (I will lead E by 90° so $P = EI \cos 90 = 0$)

The effect of reactive load can be observed by replacing it with fictitious open or short circuited extension of the line. The length of the fictitious extension is so chosen that the reactance seen at the actual load end is the same (input impedance of the extended section is equal to the load).

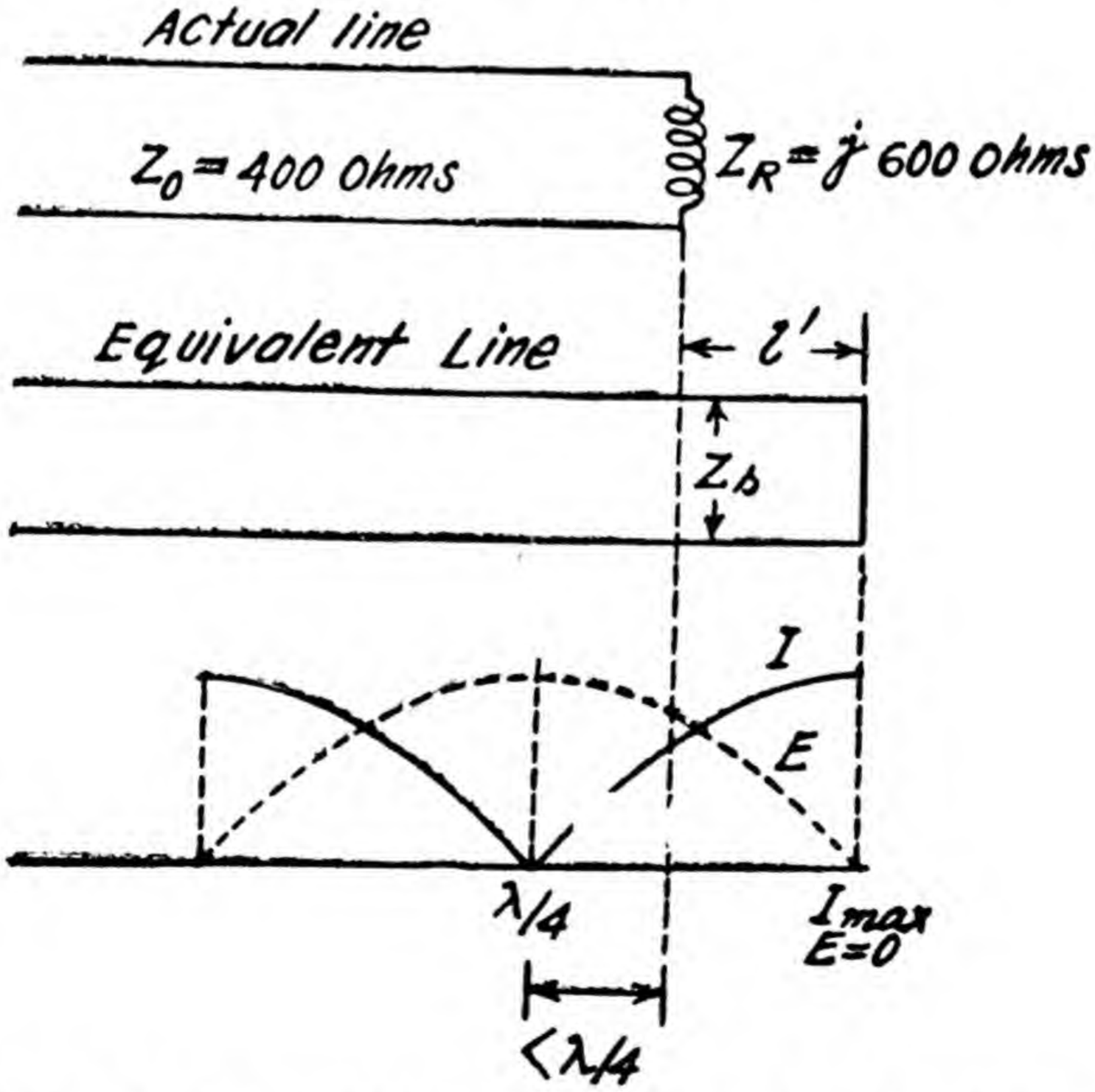


Fig. 12 (a) Equivalent short circuit extension of reactive load. Voltage maximum occurs at length $< \lambda/4$ from Z_R end.

Case 1 : *Equivalent short circuit extension of reactive load j.600 ohms :* Extension is shown in the fig. 12 (*a*). Let its length be l'. Then we should have

$$Z_s = Z_R = j.600$$

But $$Z_s = jZ_0 \tan \beta l'$$

so that $$j.o00 = jZ_0 \tan \beta l'$$

$$= j.400 \tan\left(\frac{2\pi}{\lambda} l'\right)$$

$$\frac{3}{2} = \tan\left(\frac{360 l'}{\lambda}\right)$$

or $$56{\cdot}3° = \frac{360}{\lambda}\, l'$$

or $$0{\cdot}156\lambda = l'$$

Since $l' < \lambda/4$ so Z_s is inductive.

It implies that an inductive load is equivalent to a short circuited extension with a length smaller that $\lambda/4$. Since E is zero at the fictitious shorted end, this places a voltage maximum at less than $\lambda/4$ distance from an inductive load. Larger inductances are equivalent to longer extensions and the voltage maximum moves towards load end, until at infinite inductance, the effect is that ot an open circuit.

Case 2 : *Equivalent open circuited extension of reactive load $-j.500$ ohms :* For open circuited line input impedance is

$$Z_s = -j\, \frac{Z_0}{\tan \beta l'}$$

or $$-j.500 = -j.\, \frac{400}{\tan \beta l'}$$

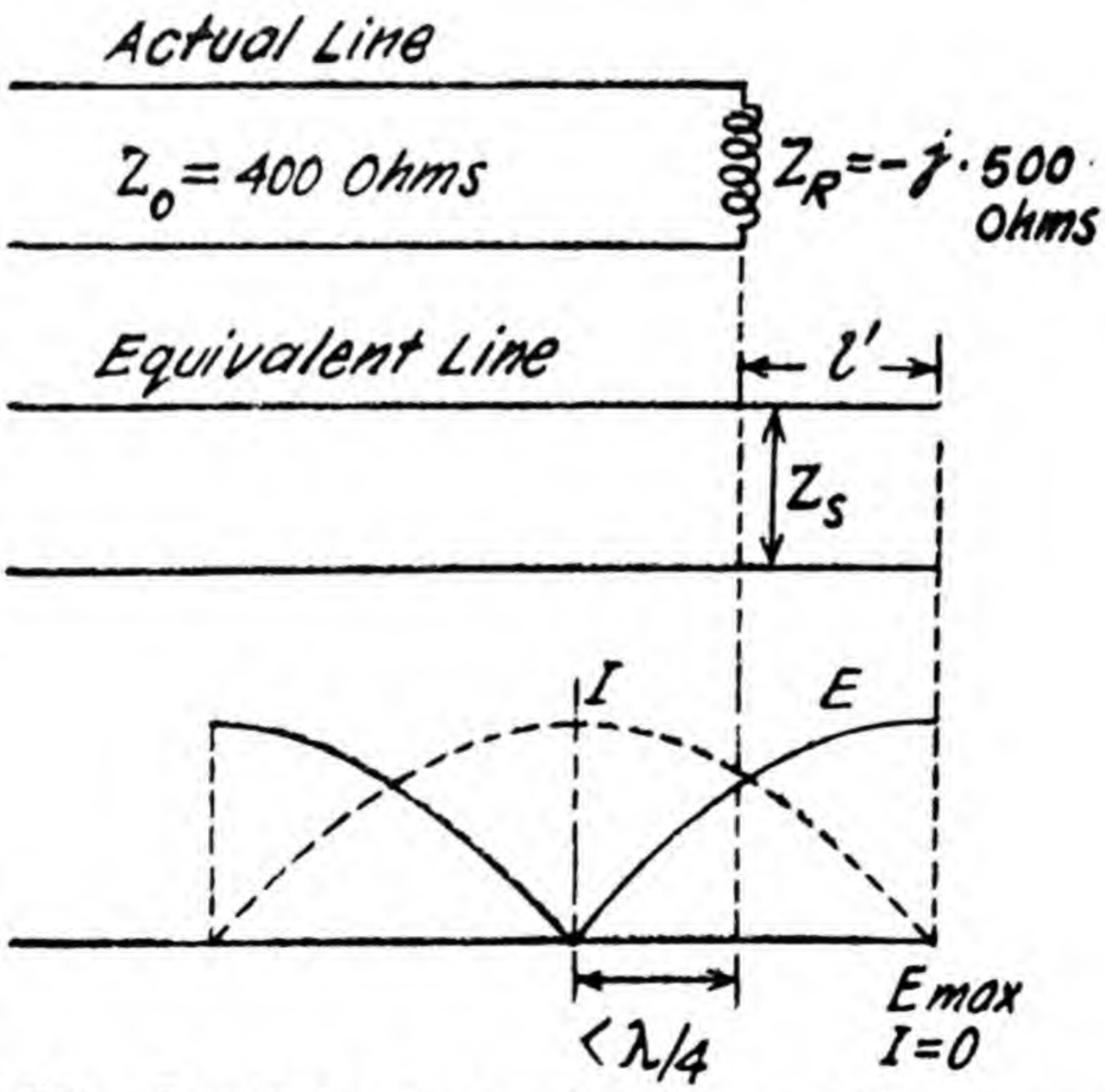


Fig. 12 (b). Equivalent open circuit extension of reactive load. Current maximum occurrs at length $<\lambda/4$ from Z_R end.

or $$\tan\left(\frac{2\pi}{\lambda}l'\right)=\frac{4}{5}$$

or $$l'=0{\cdot}10 > \lambda$$

Since $l' < \lambda/4$, Z_s is capacitive.

It implies that a capacitive load is equivalent to an open-ended extension with a length smaller than $\lambda/4$. This leads to a voltage minimum or current maximum) at less than $\lambda/4$ distance from the load. Larger capacitances are equivalent to longer extensions thus causing the voltage minimum to move closer to load. A very large capacitance gives the effect of a short circuit, with $E=0$ at the load.

The use of these extensions help us to locate maxima and minima of voltage and current on the line. Thus standing wave pattern and hence the impedances at various points is found.

(iii) **Line Terminated by Some Resistance :** This case is of much practical significance.

Let us now consider that a transmission line is terminated by a resistance R and suppose the characteristic impedance, Z_0. of the line is R_0. Consider, when switch S is not closed, the initial value of current is I_1 and the voltage at the input terminals of the line is e_1. Now when S is closed, I_1 becomes I_2 (say) and e_1 becomes e_2. Then $e^+=(e_2-e_1)$ voltage wave will travel towards the load end and therefore corresponding value of current will be $\frac{e^+}{R_0}=i^+$.

Thus under such conditions, the equivalent circuit can be represented as shown in fig. 13 (*b*).

This equivalent circuit can be replaced by another by applying Thevenin's theorem as shown in fig. 13 (*c*),

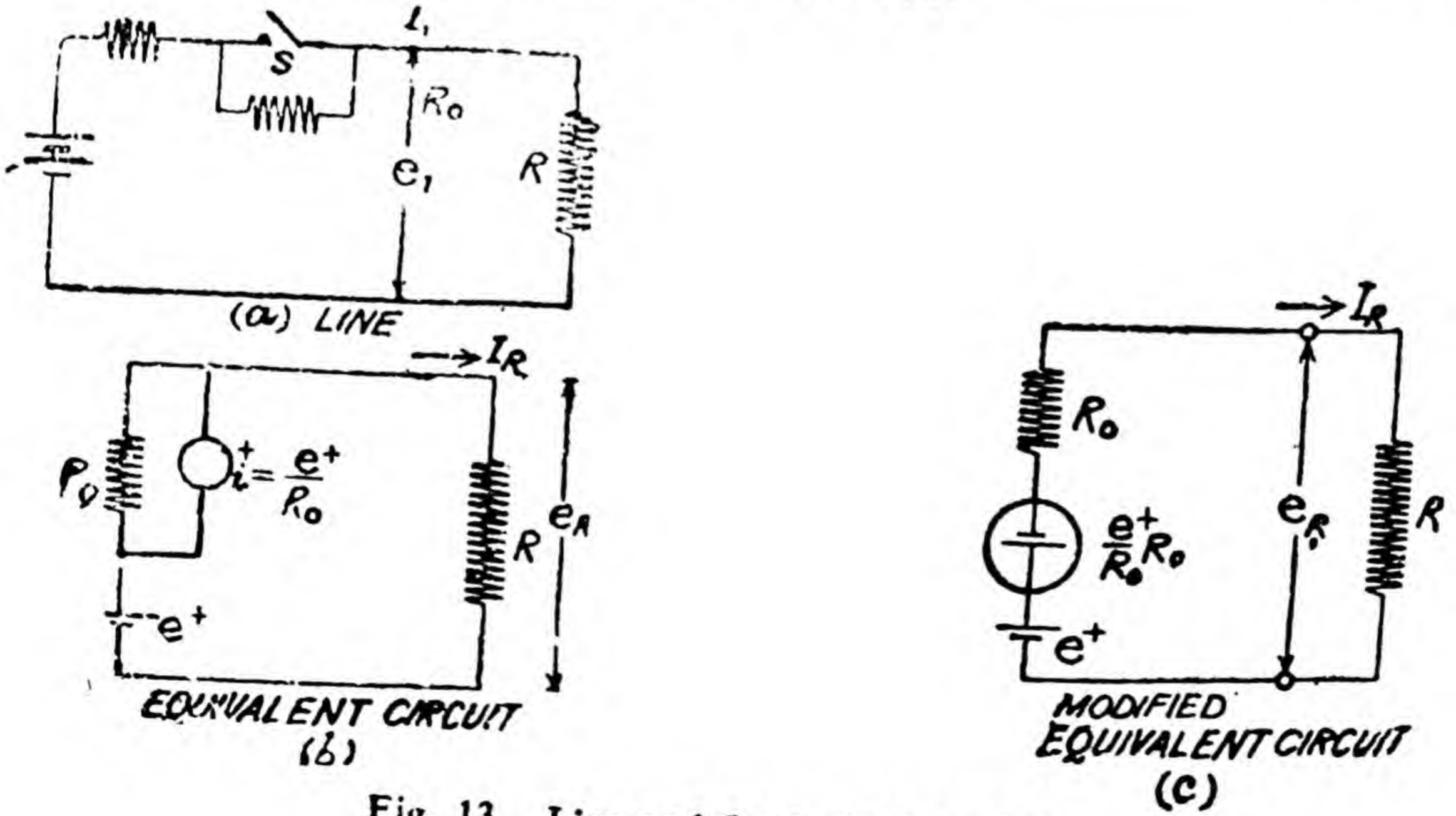


Fig. 13. Line and Equivalent Circuit

Now in this overall series circuit total voltage is $2e^+$ and total resistance is $(R+R_0)$ and hence current will be

$$I_R=\frac{2e^+}{R+R_0}$$

Therefore $e_R=I_R.R=\dfrac{2e^+R}{R+R_0}$

Voltage e_R at the load will be due to the sum of incident voltage e^+ and reflected voltage e^- and hence

$$e^-=e_R-e^+=\frac{2e^+.R}{R+R_0}-e^+=e^+.\frac{R-R_0}{R+R_0}$$

Therefore, reflection coefficient,

$$\rho=\frac{\text{reflected voltage}}{\text{incident voltage}}=\frac{e^-}{e^+}=\frac{R-R_0}{R+R_0}.$$

If $R=R_0$, $\rho=0$, *i e.*, no reflected wave,

$R=0$, $\rho=-1$, wave in reflected with full amplitude but reversed in polarity.

Thus the amplitude and polarity of the incident wave on reflection depends upon the ratio R/R_0. If $R>R_0$, ρ is positive, *i.e.*, incident wave and reflected wave are of the same polarity. If $R<R_0$, ρ is negative *i.e.* polarity of the reflcted wave is reversed.

Now we shall discuss a particular case when the line is terminated by resistance equal to 3 times the characteristic resistance of the line. *i.e.*, $R=3R_0$.

Obviously $$\rho=\frac{3R_0-R_0}{3R_0+R_0}$$
$$=0{\cdot}5.$$

Since ρ is +ve, reflected wave will be of the same polarity but of half the amplit de of incident wave. So if E_R^+ is the voltage at the load due to incident wave and E_R^- that due reflec'ed wave, then

$$\frac{E_R^-}{E_R^+}=\rho=0{\cdot}5$$

$\therefore$ $E_R^-=0{\cdot}5\ E_R^+$.

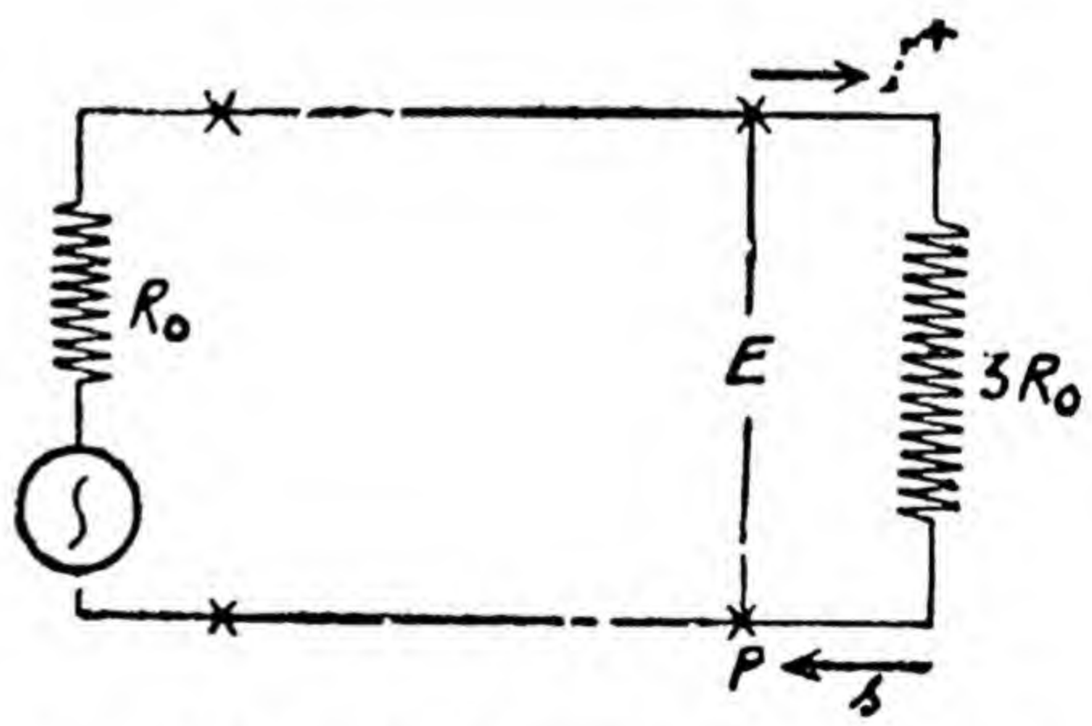


Fig. 13 (d). Line terminated in thrice the characteristic resistance.

Further, the incident and reflected current components are given by $I_R^+=\dfrac{E_R^+}{R_0}$

$$I_R^- = -\frac{E_R^-}{R_0} = -\frac{\rho E_R^+}{R} = -\frac{0{\cdot}5E_R^+}{R_0}.$$

The explanation for minus sign can be had as follows ; since voltage is reflected back from load in the same phase, though with half the incident wave amplitude, the strength of electrostatic field increases. This implies a corresponding decrease in current so that $I_R^- = -0{\cdot}5\ I_R^+$ at the load.

Now at the distance s from the load end, the voltage will be

$$E = E^+ + E^-$$
$$= E_R^+ e^{j\beta s} + E_R^- e^{-j\beta s}$$
$$= E_R^+ e^{j\beta s} + \rho E_R^+ e^{-j\beta s}$$
$$E = E_R^+ (1+\rho) \cos \beta s + jE_R^+ (1-\rho) \sin \beta s$$
$$= 1{\cdot}5E_R^+ \cos \beta s + j0{\cdot}5E_R^+ \sin \beta s,$$

so that at $\beta s = 0$, *i.e.* $s = 0$ $\quad E = 1{\cdot}5E_R^+$
$\beta s = 90$, *i.e.* $s = \lambda/4$, $E = j\ 0{\cdot}5E_R^+$ } Both are perpendicular to each other
$\beta s = 180$, *i.e.* $s = \lambda/2$, $E = -1{\cdot}5E_R^+ \rightarrow$ phase is reversed.

... ...

Hence we can plot $| E |$ and express the complex number diagrams as a function of βs as shown in fig. 14.

Now for the plot of current we see that current I at P,

$$I = I_R^+ e^{j\beta s} + I_R^- e^{-j\beta s}$$
$$= \frac{E_R^+}{R_0} e^{j\beta s} - \frac{E_R^-}{R_0} e^{-j\beta s}$$
$$= \frac{E_R^+}{R_0} e^{j\beta s} - \frac{\rho E_R^+}{R_0} e^{-j\beta s}$$
$$= \frac{E_R^+}{R_0} (1-\rho) \cos \beta s + j\,\frac{E_R^+}{R_0} (1+\rho) \sin \beta s$$
$$= 0{\cdot}5\ I_R^+ \cos \beta s + j.\ 1{\cdot}5\ I_R^+ \sin \beta s,$$

so that at $\beta s = 0$, *i.e.*, $s = 0$, $\quad I = 0{\cdot}5\ I_R^+$
$\beta s = 90$, *i.e.*, $s = \lambda/4$, $I = j\ 1{\cdot}5\ I_R^+$,
$\beta s = 180$, *i.e.*, $s = \lambda/2$, $I = -0{\cdot}5\ I_R^+$.

These data help to plot magnitude and complex number diagrams for current.

Here we note that $E_{min.} = 0{\cdot}5\ |\ E_R^+\ |$
$E_{max.} = 1{\cdot}5\ |\ E_R^+\ |$.

Therefore standing wave ratio

$$S = \frac{\frac{|\ E_{max.}\ |}{\sqrt{2}}}{\frac{|\ E_{min.}\ |}{\sqrt{2}}} = \frac{E_{max.}}{E_{min.}}$$

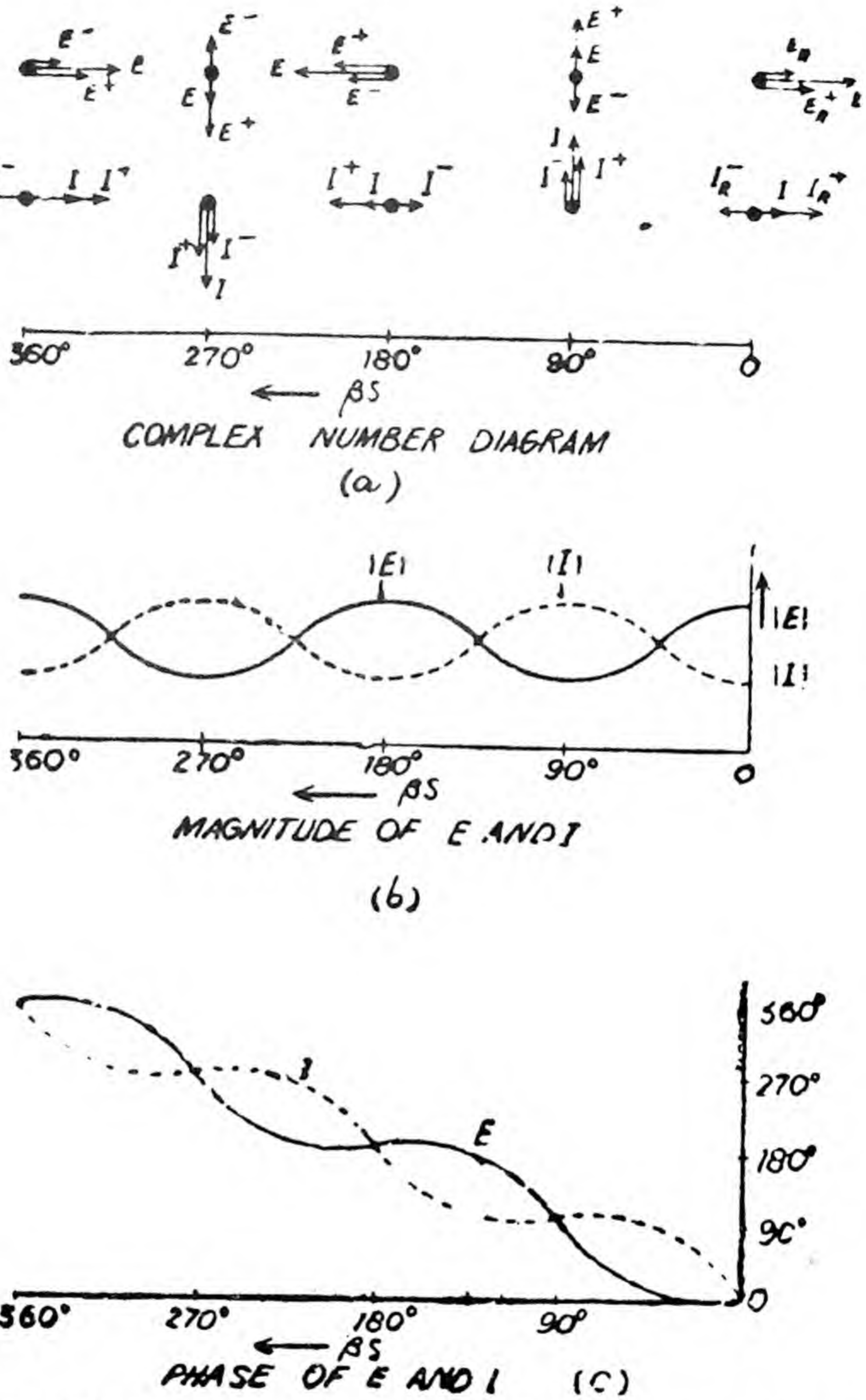


Fig. 14. Voltage and current relations for a line terminated in a resistance $R=3R_0$.

i.e. the amount of variation of r.m.s. voltage in such a case will be

$$S=\frac{1\cdot5}{0\cdot5}\left|\frac{E_R^+}{E_R^+}\right|=3.$$

From complex number diagrams, we observe that maximum voltage occurrs at points where reflected and incident voltages are in phase, while minimum voltage occurrs where the two are in opposite phase and hence

$$|E_{max.}| = |E_R^+| + |E_R^-|$$
$$= |E_R^+| + |\rho|\,|E_R^+|$$

$$= (1 + |\rho|)\,|E_R^+|$$

and
$$|E_{min.}| = |E_R^+| - |E_R^-|$$
$$= |E_R^+| - |\rho|\,|E_R^+|$$
$$= (1 - |\rho|)\,|E_R^+|.$$

Hence S.W.R., $$S = \frac{|E_{max.}|}{|E_{min.}|} = \frac{1 + |\rho|}{1 - |\rho|}.$$

The relation clearly shows that when reflection coefficient tends to zero, standing wave ratio tends to unity ; while when reflection coefficient is unity, standing wave ratio tends to infinity The latter case is with short circuited or open circuited lines. High value of standing wave ratio is always undesirable since it results in the reduction of maximum power that can be transmitted by the line, *i.e.*, the reduction of *power capacity* of the line —the limit to the power is set by the maximum voltage between the conductors that the insulation can withstand. Let this maximum or limiting r.m.s. value of the voltage between the conductors be $|E_{lim}|$. The power transmitted from the source to the load by a loss less line is

$$P = |E_{max.}|\,|I_{min.}|,$$

since at a point where voltage is maximum, current will be minimum, and phase difference between voltage and current is zero in case of resistive load. Substituting $|I_{min}| = |E_{min.}|/Z_0$, we get

$$P = \frac{|E_{max.}|\,|E_{min.}|}{Z_0},$$

where Z_0 is the characteristic resistance of the line. Further the standing wave ratio is

$$S = \frac{|E_{max.}|}{|E_{min.}|} = \frac{|E_{lim.}|}{|E_{min.}|},$$

so that

$$E_{min} = \frac{|E_{lim}|}{S}.$$

Therefore power transmitted becomes

$$P = \frac{|E_{lim}|\,|E_{lim}|/S}{Z_0}$$
$$= \frac{|E_{lim}|^2}{Z_0 S},$$

which shows that standing waves cause a reduction of power capacity. In short or open circuited lines S is infinity so that P is zero.

(iv) **Termination by complex impedance :** We have noted that in the case of short circuited and open circuited lines, wave is reflected with opposite and same polarity respectively but with full amplitude. However, in the case of resistance termination there is a change of both amplitude and phase depending upon the value of load resistance Similarly in the case of complex impedance change in amplitude and phase depends upon the value of the impedance. Suppose the load is such that $\rho = 0{\cdot}5e^{j\cdot 60}$.

Then obviously,

$$\frac{E_R^-}{E_R^+} = 0{\cdot}5e^{j\cdot 60},$$

$$E_R^- = 0{\cdot}5E_R^+ e^{j\cdot 60}.$$

Thus phase between reflected and incident waves will be 60° at the load and the amplitude of reflected wave is half the amplitude of incident wave.

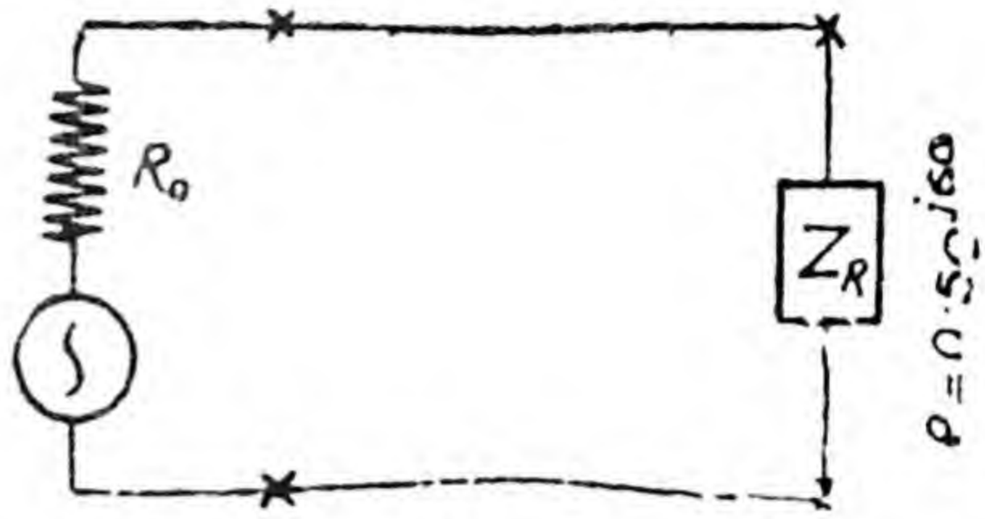


Fig. 15.

Total voltage at the load

$$= E_R^+ + E_R^-$$
$$= E_R^+ + 0{\cdot}5e^{j\cdot 60}\, E_R^+$$
$$= E_R^+ + 0{\cdot}5E_R^+ (\cos 60 + j \sin 60)$$
$$= 1{\cdot}32\, E_R^+ e^{j19\cdot 1}. \quad \ldots (1)$$

Total load current $= I_R^+ - 0{\cdot}5I_R^+ e^{j\cdot 60}$

$$= \frac{E_R^+}{R_0}[1 - 0{\cdot}5\cos 60 - 0{\cdot}5j \sin 60]$$

$$= 0{\cdot}866\, \frac{E_R^+}{R_0} e^{-j\,30}. \quad \ldots (2)$$

The *general equations* for voltage at any distance s from the load end will be

$$E = E_R^+ e^{j\beta s} + E_R^- e^{-j\beta s}.$$

Let $E_R^- = E_R^+ \rho$

$$= E_R^+ \mid \rho \mid e^{j\phi}$$

$$E = E_R^+ e^{j\beta s} + \mid \rho \mid e^{j\phi} . E_R^+ e^{-j\beta s}$$

$$= E_R^+ (1 + \mid \rho \mid e^{j\phi}) \cos \beta s$$

$$+ jE_R^+ (1 - \mid \rho \mid e^{j\phi}) \sin \beta s. \quad \ldots (3)$$

Similarly for current the expression will be

$$I = I_R^+ e^{j\beta s} + I_R^- e^{-j\beta s}$$

$$= \frac{E_R^+}{R_0} e^{j\beta s} - \frac{E_R^-}{R_0} e^{-j\beta s}$$

$$= \frac{E_R^+}{R_0} e^{j\beta s} - \mid \rho \mid e^{j\phi} . \frac{E_R^+}{R_0} . e^{-j\beta s}$$

$$= I_R^+ (1 - \mid \rho \mid e^{j\phi}) \cos \beta s + jI_R^+ (1 + \mid \rho \mid e^{j\phi}) \sin \beta s. \quad \ldots (4)$$

Like the previous case, here too we can plot the complex number diagrams, amplitude of voltage and current at different points of the line with the help of equations (3) and (4).

Since equations (1) and (2) correspond to the quantities at the load end, the load impedance that produces the assumed value of reflection coefficient is therefore

$$Z_R = \frac{E_R}{I_R} = \frac{1{\cdot}32}{0{\cdot}866}\, R_0 e^{j\,(19\cdot 1 + 30)}$$

$$= 1{\cdot}53\, R_0 . e^{j49\cdot 1}.$$

Angle of the reflection coefficient determines the position of standing wave pattern along the line and magnitude determines the shape of the pattern.

(v) **Stub matching** : When it is not possible to control the terminating impedance of the line, the characteristic impedance being fixed, an approach to secure matching condition is, however, made by connecting a stub line in parallel with the transmission line. The length of the stub line and its point of contact both are important factors to be considered.

Let us consider a line with characteristic resistance R_0 and terminated by a resistance R. Suppose at distance l from load end, *i.e.*, at AA points a stub line is connected. Then if the reflected wave on the transmission line at AA, becomes equal and opposite in polarity to the wave on stub line at AA, then both will cancel out and there will be no reflected wave on the transmission line in the portion XA. Such a behaviour is expected only when a line is terminated in characteristic resistance or in other words, the net impedance at AA points should be resistive and equal to R_0, the characteristic resistance. It means the stub line must offer only a resistance equal and opposite to the reactance of the transmission line at AA, so that both cancel out and only resistive component of the line impedance at AA is left. Thus stub line should be such a line whose input impedance is purely reactive (no resistive component is present) for example, short-circuited line.

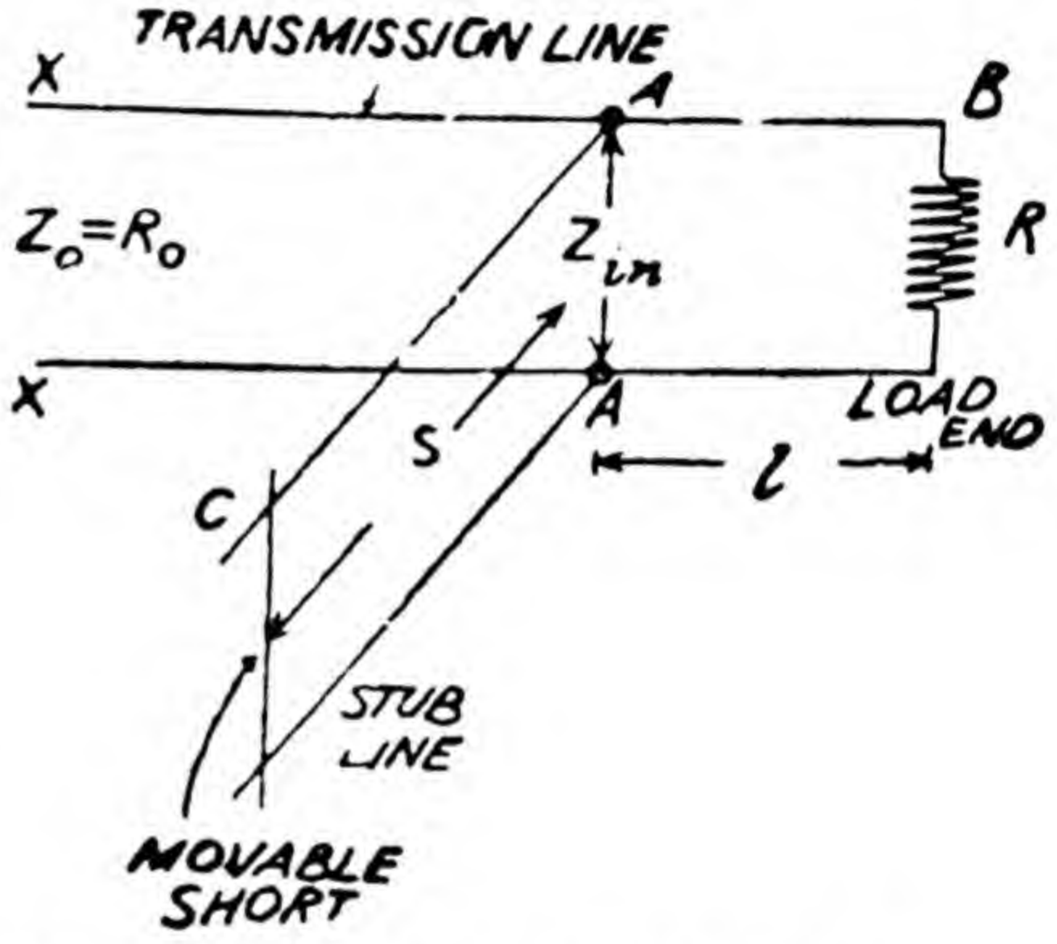


Fig. 16. Stub connected line

Thus there are two considerations :

(*a*) length of the line to the points at which stub line is connected from the load end should be such as to offer a resistive component R_0 in the impedance

(*b*) length of the stub line should be such as to cancel the reactive component of line impedance by its input reactance.

Mathematically, input impedance at AA of the line will be

$$*Z_{in}=\frac{R_0\,(R\cos\beta l+jR_0\sin\beta l)}{(R_0\cos\beta l+jR\sin\beta l)}$$

$$*Z_{in}=\frac{E_l}{I_l}=\frac{E_R\cosh\gamma l+I_R/Z_0\sinh\gamma l}{I_R\cosh\gamma l+E_R/Z_0\sinh\gamma l}$$

Now here $Z_0=R_0$ and put $E_R=I_R R$; $\gamma=j\beta$ so that $\sinh\gamma l=j\sin\beta l$; $\cosh\gamma l=\cos\beta l$.

or input admittance

$$Y_{in}=\frac{1}{R_{in}}=G_{in}+jB_{in},$$

where $G_{in}=\frac{1}{R_{in}}$=conductance, $B_{in}=\frac{1}{X_{in}}$=susceptance

$$Y_{in}=\frac{R_0\cos\beta l+jR\sin\beta l}{R_0\,(R\cos^2\beta l+jR_0\sin^2\beta l)}$$

$$=\frac{RR_0}{R_0\,(R^2\cos^2\beta l+R_0{}^2\sin^2\beta l)}$$

$$+\frac{j\,(R^2-R_0{}^2)\sin\beta l\cos\beta l}{R_0\,(R^2\cos^2\beta l+R_0{}^2\sin^2\beta l)}$$

We know *from consideration* (*a*) that I should be chosen such that

$$G_{in}=\frac{1}{R_0}$$

$$\therefore\quad \frac{RR_0}{R_0\,(R^2\cos^2\beta l+R_0{}^2\sin^2\beta l)}=\frac{1}{R_0}$$

$$\frac{R}{R_0}\cos^2\beta l+\frac{R_0}{R}\sin^2\beta l=\cos^2\beta l+\sin^2\beta l,$$

$$\cos^2\beta l.\left(1-\frac{R}{R_0}\right)=\sin^2\beta l.\left(\frac{R_0}{R}-1\right).$$

$$\tan^2\beta l=R/R_0$$

$$\tan\beta l=\sqrt{(R/R_0)}. \qquad \ldots(1)$$

This relation obviously gives the length l of the line from the load end at which stub line should be connected.

Now *for consideration* (*b*) :

input reactance of stub line$=\frac{1}{jB_{in}}$.

Now input reactance of a short circuited stub line$=jR_0\tan\beta s$ where s is the length of stub line. Therefore, to satisfy condition (*b*), we should have

$$\frac{1}{jB_{in}}=-jR_0\tan\beta s$$

or

$$jB_{in}=-\frac{1}{jR_0\tan\beta s}$$

or

$$j\cdot\frac{\sin\beta l\cos\beta l\,(R^2-R_0{}^2)}{R_0\,(R^2\cos^2\beta l+R_0{}^2\sin^2\beta l)}=\frac{j}{R_0\tan\beta s}$$

$$\therefore\quad \tan\beta s=\frac{R^2\cos^2\beta l+R_0{}^2\sin^2\beta l}{(R^2-R_0{}^2)\sin\beta l\cos\beta l}$$

$$=\frac{\frac{R}{R_0}+\frac{R}{R_0}\tan^2\beta l}{\tan\beta l\cdot\left(\frac{R}{R_0}-\frac{R_0}{R}\right)}$$

$$= \frac{\frac{R}{R_0} + \frac{R_0}{R}\cdot\frac{R}{R_0}}{\sqrt{\left(\frac{R}{R_0}\right)\left(\frac{R}{R_0} - \frac{R_0}{R}\right)}} = \frac{1+\frac{R}{R_0}}{\sqrt{\left(\frac{R}{R_0}\right)\cdot\left(\frac{R}{R_0}-\frac{R_0}{R}\right)}} = \frac{1+\frac{R}{R_0}}{\sqrt{\left(\frac{R_0}{R}\right)\cdot\left(\frac{R^2}{R_0^2}-1\right)}}$$

On further simplifying,

$$\tan \beta s = \frac{\sqrt{\left(\frac{R}{R_0}\right)}}{\left(\frac{R}{R_0}-1\right)} \qquad \ldots(2)$$

This equation gives the proper stub line length. Thus having formulated the two considerations, it is possible to match a transmission line by the application of a stub line. This causes the maximum transference of energy from input to the load end. From load end to stub connected points the line is resonant one, but from stub connected point to input end the line is non-resonant.

†**To obtain the values of the line length and stub line length in terms of standing wave ratio :**

The input impedance Z_l, looking towards the load from any point (distant l from load end) on the line, may be written as

$$Z_l = \frac{E_l}{I_l} = \frac{E_R \cosh \gamma l + I_R Z_0 \sinh \gamma l}{I_R \cosh \gamma l + E_R/Z_0 \sinh \gamma l}$$

$$= \frac{I_R\left[Z_R \cdot \frac{e^{\gamma l}+e^{-\gamma l}}{2} + Z_0 \cdot \frac{e^{\gamma l}-e^{-\gamma l}}{2}\right]}{I_R\left[\frac{e^{\gamma l}+e^{-\gamma l}}{2} + \frac{Z_R}{Z_0}\cdot\frac{e^{\gamma l}-e^{-\gamma l}}{2}\right]}$$

$$= \frac{Z_0\left[(Z_R+Z_0)\,e^{\gamma l} + (Z_R - Z_0)\,e^{-\gamma l}\right]}{(Z_R+Z_0)\,e^{\gamma l} - (Z_R - Z_0)\,e^{-\gamma l}}$$

$$= \frac{Z_0\left[e^{\gamma l} + \rho e^{-\gamma l}\right]}{e^{\gamma l} - \rho e^{-\gamma l}},$$

where $\rho = \frac{Z_R - Z_0}{Z_R + Z_0}$, called the reflection coefficient.

---

†From equations (1) and (2) we can also arrive at the same requirement. With load R, reflection coefficient is

$$|\rho| = \frac{R - R_0}{R + R_0}.$$

so that $$S = \frac{1+|\rho|}{1-|\rho|} = \frac{R}{R_0},$$

giving $$\tan \beta l = \sqrt{S}$$

and $$\tan \beta s = \frac{\sqrt{S}}{(S-1)}.$$

If the line is lossless $\alpha=0$ *i.e.* $\gamma=j\beta$ and writing $\rho=|\rho|e^{j\phi}$, we find

$$Z_l=\frac{Z_0[e^{j\beta l}+|\rho|e^{j(\phi-\beta l)}]}{e^{j\beta l}-|\rho|e^{j(\phi-\beta l)}}$$

If the characteristic impedance is purely resistive, then $Z_0=R_0$ so that

$$Z_l=R_0\left[\frac{1+|\rho|e^{j(\phi-2\beta l)}}{1-|\rho|e^{j(\phi-2\beta l)}}\right]$$

$\therefore$ Input admittance,

$$Y_l=\frac{1}{Z_l}=\frac{1}{R_0}\left(\frac{1-|\rho|\angle(\phi-2\beta l)}{1+|\rho|\angle(\phi-2\beta l)}\right)$$

Writing $\frac{1}{R_0}=G_0$, we get

$$Y_l=G_0=\left[\frac{1-|\rho|\cos(\phi-2\beta l)-j|\rho|\sin(\phi-2\beta l)}{1+|\rho|\cos(\phi-2\beta l)+j|\rho|\sin(\phi-2\beta l)}\right]$$

$$Y_l=G_0\left[\frac{1-j|\rho|\sin(\phi-2\beta l)-|\rho|\cos(\phi-2\beta l)}{1+|\rho|\cos(\phi-2\beta l)+j|\rho|\sin(\phi-2\beta l)}\right]$$

$$\times\frac{[1-j|\rho|\sin(\phi-2\beta l)+|\rho|\cos(\phi-2\beta l)]}{[1+|\rho|\cos(\phi-2\beta l)-j|\rho|\sin(\phi-2\beta l)]}$$

$$=G_0\left[\frac{\{1-j|\rho|\sin(\phi-2\beta l)\}^2-|\rho|^2\cos^2(\phi-2\beta l)}{\{1+|\rho|\cos(\phi-2\beta l)\}^2+|\rho|^2\sin^2(\phi-2\beta l)}\right]$$

$$=G_0\left[\frac{1-|\rho|^2-2j|\rho|\sin(\phi-2\beta l)}{1+|\rho|^2+2|\rho|\cos(\phi-2\beta l)}\right]$$

or $$\frac{Y_l}{G_0}=\frac{G_l+jB_l}{G_0}=\left[\frac{1-|\rho|^2-2j|\rho|\sin(\phi-2\beta l)}{1+|\rho|^2+2|\rho|\cos(\phi-2\beta l)}\right]$$

Therefore,

$$\frac{G_l}{G_0}=\frac{1-|\rho|^2}{1+|\rho|^2+2|\rho|\cos(\phi-2\beta l)} \quad \ldots(3)$$

and $$\frac{B_l}{G_0}=\frac{-2|\rho|\sin(\phi-2\beta l)}{1+|\rho|^2+2|\rho|\cos(\phi-2\beta l)} \quad \ldots(4)$$

Now our first condition is that resistive part of input impedance at a point distant l from load end must be equal to the characteristic resistance and hence

$$G_l=\frac{1}{R_0}.$$

So that $$\frac{G_l}{G_0}=1. \qquad \left(\text{Since } G_0=\frac{1}{R_0}\right)$$

Therefore, from equation (1), we get

$$1=\frac{1-|\rho|^2}{1+|\rho|^2+2|\rho|\cos(\phi-2\beta l)}$$

$$1+|\rho|^2+2|\rho|\cos(\phi-2\beta l)=1-|\rho|^2$$

or $$|\rho|+\cos(\phi-2\beta l)=0$$

or $$\cos(\phi-2\beta l)=-|\rho|.$$

Since $^{*}\cos^{-1}(-|\rho|)=-\pi+\cos^{-1}|\rho|,$

thus $$\phi-2\beta l=-\pi+\cos^{-1}|\rho|,$$

or $$l=\frac{\phi+\pi-\cos^{-1}|\rho|}{2}. \qquad ...(5)$$

That gives the length of line at which the stub should be connected.

Now second condition is that the reactive part of input impedance at the point distant l from load end, *i.e.* Z_l must be cancelled by the input impedance of quarter wave line. Hence the stub line must offer only reactive impedance. Such lines whose input impedance is purely reactive are short circuited and open circuited lines. Let us take a short circuited line. The input impedance of a short circuited line is

$$jR_0 \tan \beta s$$

where s is the length of stub line.

So the susceptance B_s is

$$\beta s=\frac{1}{jR_0 \tan \beta s},$$

i.e. $$B_s=-jG_0 \cot \beta s.$$

The second condition demands that susceptance of stub, B_s, required to cancel the line susceptance, must be negative of B_l.

$\therefore \quad B_s=-B_l$

$$=+G_0.\frac{2|\rho|\sin(\phi-2\beta l)}{1+|\rho|^2+2|\rho|\cos(\phi-2\beta l)} \text{ from (4).}$$

Putting the value of l from equation (3),

$$B_s=G_0.\frac{-2|\rho|\sin(\pi-\cos^{-1}|\rho|)}{1+|\rho|^2+2|\rho|\cos(\pi-\cos^{-1}|\rho|)}$$

$$=G_0.\frac{2|\rho|\cos\pi.\sin\cos^{-1}|\rho|}{1-|\rho|^2+2|\rho|\cos\pi.\cos\cos^{-1}|\rho|}.$$

Now putting $\cos^{-1}|\rho|=\theta,$

then $$\cos\theta=|\rho|.$$

$\therefore \quad \sin\cos^{-1}|\rho|=\sin\theta=\sqrt{(1-|\rho|^2)}$

and $$\cos\cos^{-1}|\theta|=\cos\theta=|\rho|.$$

Therefore,

$$B_s=G_0\frac{-2|\rho|\sqrt{(1-|\rho|^2)}}{1-|\rho|^2-2|\rho|^2}=\frac{-2G_0|}{1}\frac{\sqrt{(1-|\rho|^2)}}{-|\rho|^2}.$$

Putting $B_s=-G_0 \cot \beta s$, we find

$^{*}\cos^{-1}(-1)=\theta$

$\cos\theta=-1,\ \theta=\pi$

$\cos^{-1}(-1)=\pi+\cos^{-1}(+1)$

$=\pi+0$

$=\pi$ or $-\pi$

$$\cot \beta_S = \frac{2 | \rho | \sqrt{(1 - | \rho |^2)}}{1 - | \rho |^2} = \frac{2 | \rho |}{\sqrt{(1 - | \rho |^2)}}.$$

$$\therefore \quad \tan \beta_S = \frac{\sqrt{(1 - | \rho |^2)}}{2 | \rho |}. \qquad \ldots(6)$$

Putting, $| \rho | = \frac{S-1}{S+1}$, S being the standing wave ratio,

$$\tan \beta_S = \frac{\sqrt{\{(S+1)^2 - (S-1)^2\}}}{2.(S-1)} = \frac{\sqrt{S}}{(S-1)}.$$

$$\therefore \quad s = \frac{1}{\beta} \tan^{-1} \frac{\sqrt{S}}{(S-1)} \qquad \ldots(7)$$

Equation (7) gives the length of stub line in terms of standing wave ratio. Equation (5) can also be represented in terms of standing wave ratio as follows :

$$l = \frac{\phi + \pi - \cos^{-1} | \rho |}{2\beta}$$

$$= \frac{\phi + \pi - \cos^{-1} \left(\frac{S-1}{S+1} \right)}{2\beta}. \qquad \ldots(8)$$

If the phase of reflection coefficient is zero, then

$$l = \frac{\pi - \cos^{-1} \left(\frac{S-1}{S+1} \right)}{2\beta} \qquad \ldots(9)$$

Equations (7) and (8) are the desired relations for the stub line length, s, and transmission line length, l, required for proper matching.

A short circuited stub line is preferred as compared to open circuited line because of following reasons :

(*a*) a short circuited line can be easily constructed,

(*b*) at the open end of an open circuited line, the insulation, and resistance high enough cannot be maintained, *i.e.*, one cannot be sure that stub is really open circuited.

(*c*) at the open ends of an open circuited line, there will be loss of energy due to radiation while short circuited line is sealed at the load end and thus stops all field propagation.

26·4. STANDING WAVE RATIO :

As mentioned in art. 20·3 that if a line is terminated by a load other than characteristic impedance, a reflected wave exists on the line. The voltage and current at any point of such a line will be the vector sum of incident and reflected wave components. If reflected wave faces an impedance equal to the characteristic impedance of the line, then it will be completely absorbed ; consequently, a steady wave pattern will exist on that line. The

standing wave ratio is defined as the ratio of maximum to the minimum magnitude of voltage or current on the line having standing waves *i.e.*,

$$*S=\frac{|E_{max.}|}{|E_{min.}|}. \qquad ...(1)$$

Since $|E_{max.}|=Z_0|I_{max.}|$ and $|E_{min.}|=Z_0|I_{min.}|$,

$$S=\frac{|I_{max.}|}{|I_{min.}|}. \qquad ...(2)$$

The voltage at a point on the line will be maximum if the two wave components E^+ and E^- satisfy inphase condition, *i.e.*,

$$|E_{max.}|=\tfrac{1}{2}I_R(Z_R+Z_0)[1+|\rho|].$$

so that

$$S=\frac{|E_{max.}|}{|E_{min.}|}=\frac{1+|\rho|}{1-|\rho|}, \qquad ...(3)$$

expressing standing wave ratio as a function of reflection coefficient. Obviously, as for short and open circuited lines reflection coefficient $|\rho|$ is unity, standing wave ratio is infinity. For a line which has no reflected wave (a line terminated in characteristic resistance), reflection coefficient is zero and standing wave ratio is unity. Variation of S with $|\rho|$ is shown in fig. 17 (*a*).

Further, power delivered to the load, P, as a fraction of incident wave power or power delivered to a matched load can also be expressed in terms of standing wave ratio. Because

$$\frac{P}{P^+}=\frac{P^+-P^-}{P^+}=1-\frac{P^-}{P^+}=1-\frac{|E^-|^2}{|E^+|^2}=1-|\rho|^2.$$

From equation (3),

$$|\rho|=\frac{S-1}{S+1},$$

so that

$$\frac{P}{P^+}=-\left(\frac{S-1}{S+1}\right)^2=\frac{4S}{(S+1)^2}$$

when $S\to\infty$, we can write

$$\frac{P}{P^+}=\frac{4}{S}\left(1+\frac{1}{S}\right)^{-2}$$

or

$$=\frac{4}{S}\left(1-\frac{2}{S}\right)$$

or

$$P=\frac{4P^+}{S}\left(1-\frac{2}{S}\right). \qquad ...(4)$$

---

*This also termed as voltage standing wave ratio (*VSWR*). In the measurement of standing wave ratio, readings proportional to the square of voltage are sometimes obtained and the ratio of such readings is called power standing wave ratio (*PSWR*).

The relation predicts that a higher value of standing wave ratio leads to the reduction in power delivered to the load and is, therefore, not desirable. The dependence of power delivered to the load can also be displayed on proceeding with equation (1). Since a point of voltage maximum on the line corresponds to a current minimum, we can write,

$$P = |E_{max.}||I_{min.}| = \frac{|E_{max.}||E_{min.}|}{Z_0}.$$

For $$S = \frac{|E_{max.}|}{|E_{min.}|},$$

we arrive at.

$$P_{max.} = \frac{|E_{max.}|^2}{SZ_0}, \qquad \ldots(5)$$

implying that a higher value of standing wave ratio results in the reduction of the maximum power or power capacity delivered to the load by source. Refer to fig. 17 (*b*).

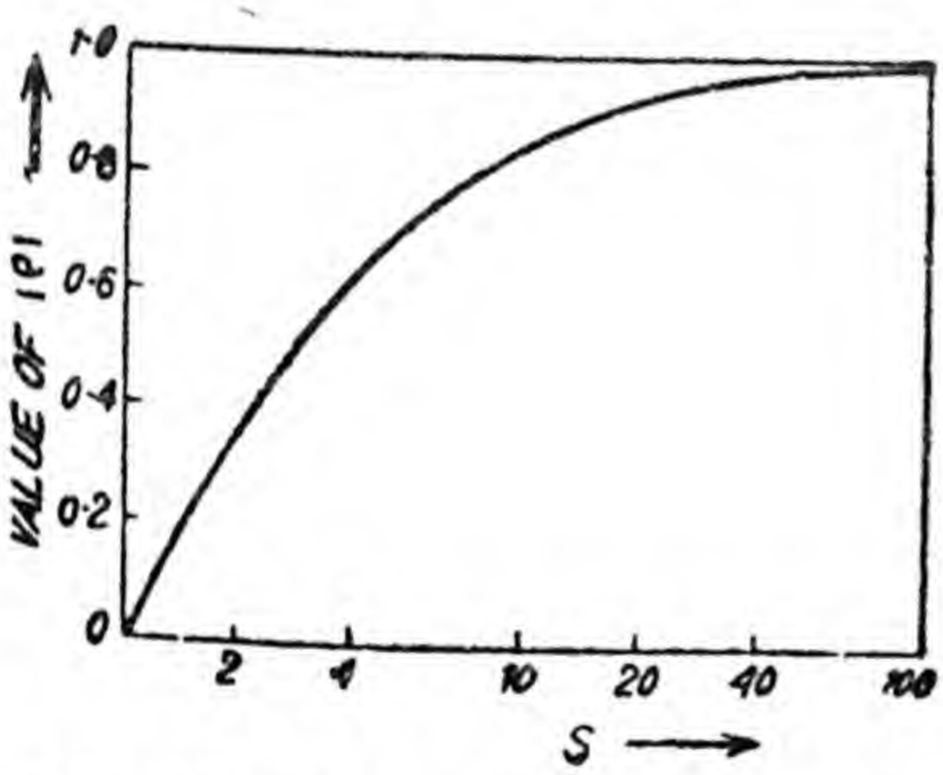


Fig. 17 (*a*) Standing wave ratio as a function of reflection coefficient.

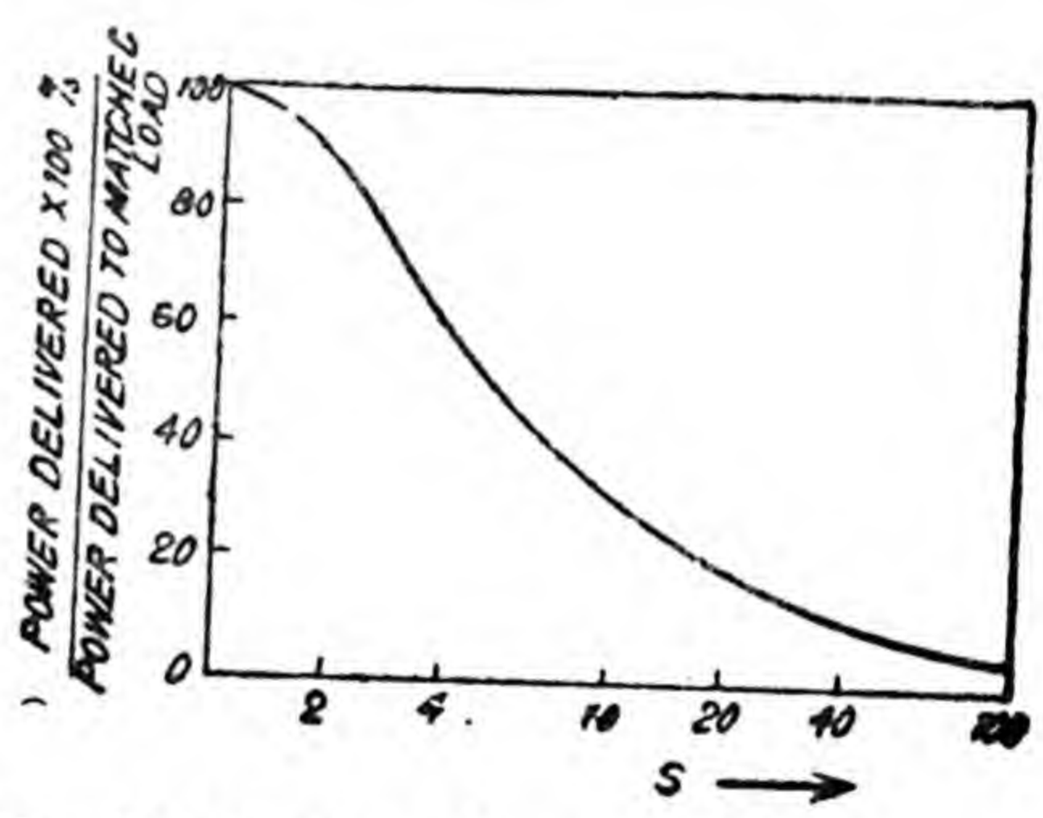


Fig. 17. (*b*) Standing wave ratio as a function of reflection losses.

If a line of characteristic resistance R_0 is terminated in R_1 such that R_1 is less than R_0, then reflection coefficient will be

$$\rho = -\frac{|R_1 - R_0|}{|R_1 + R_0|},$$

so that the standing wave ratio is

$$S = \frac{1 + \left(-\frac{|R_1 - R_0|}{|R_1 + R_0|}\right)}{1 + \frac{|R_1 - R_0|}{|R_1 + R_0|}} = \frac{R_0}{R_1} \qquad \ldots(6)$$

If the same line be terminated in R_2 such that R_2 is greater than R_0, then

$$\rho = \frac{|R_2 - R_0|}{|R_2 + R_0|}.$$

so that the standing wave ratio will come out to be

$$S=\frac{R_2}{R_0}. \qquad ...(7)$$

From equations (6) and (7), if the standing wave ratio is same in both the cases, we get

$$\frac{R_2}{R_0}=\frac{R_0}{R_1},$$
$$R_0=\sqrt{(R_1R_2)},$$

that is, characteristic resistance would be the geometric mean of the two terminating resistances.

The standing wave ratio on a transmission line can be observed by exploring along the length of the line with a pick-up arrangement that will indicate the strength of either the 'electric field, (line voltage) or the magnetic field (linear current) in the vicinity of the line. For coaxial systems a typical example of such a standing wave detector is shown in fig. 17. (*d*).

This consists of a sliding probe which extends upto a short distance into a slotted section of coaxial line. The probe is connected to a crystal

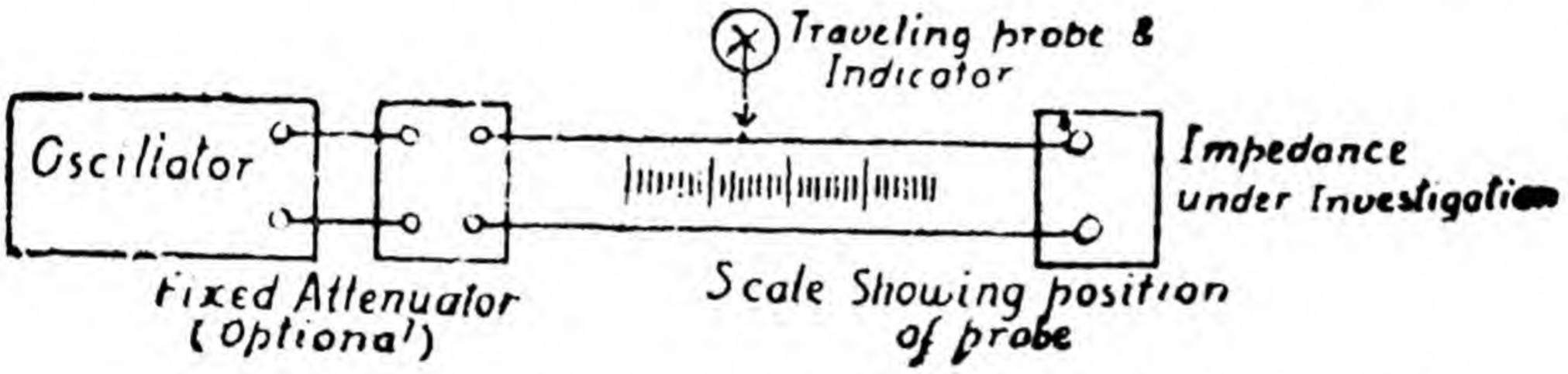


Fig. 17 (*c*) System for measuring standing wave ratio.

detector and microammeter. This device actually measures the electric intensity ; but, however, the electric intensity is proportional to the voltage between conductors, so that the standing wave indicator may be assumed to

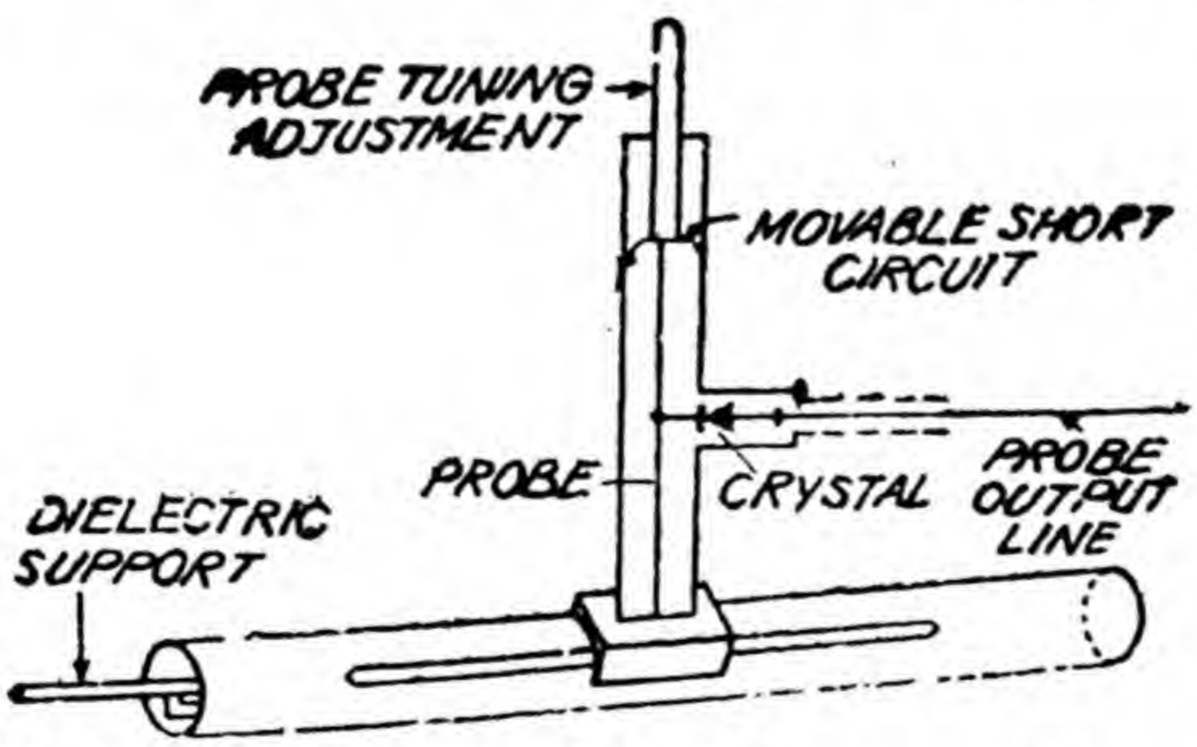


Fig. 17 (*d*). Probe and detector connection.

be a voltage measuring device. An oscillator is connected to one end of the slotted line, while the other end is connected to the unknown impedance or, alternately, to the input of the transmission line to have its standing wave ratio observation. A scale on the standing wave indicator is used to measure the distance between the terminal impedance and the voltage maximum or minimum. In this way standing wave ratio is determined.

Lecher-wire system for measuring frequency of a microwave oscillator :

A section of transmission line is connected to the resonant circuit of the oscillator. When the line is adjusted to a resonant

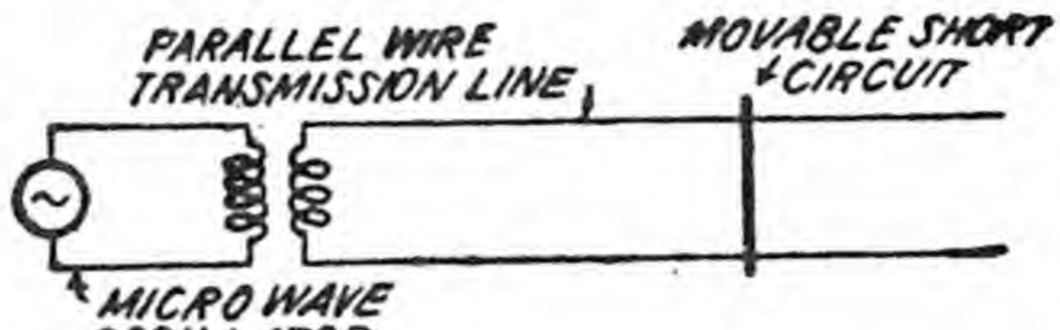


Fig. 17 (*e*). Lecher wire system for measuring frequency of a microwave oscillator.

length, the energy coupled from the oscillator is able to build up an oscillation of large amplitude on the line. The shorting bar is moved along the line. At the shorting bar the standing wave of voltage will always be zero and the current standing wave will be at its maximum. Therefore, if a current indicator is connected in the shorting bar, it will detect the positions of resonance when the bar is moved on the line. Thus a shorting bar is moved over the line and a measurement is made of the distance between two *successive* positions that produce resonance. This distance is the half wavelength. Tne frequency of the oscillator can then be found out.

26 5-1. INPUT IMPEDANCE OF DISSIPATION LESS TRANSMISSION LINE :

Here we shall discuss the impedance offered by the line at the input terminals. Load end is taken as the reference point.

Case (A) : Input impedance of a line terminated with any impedance.

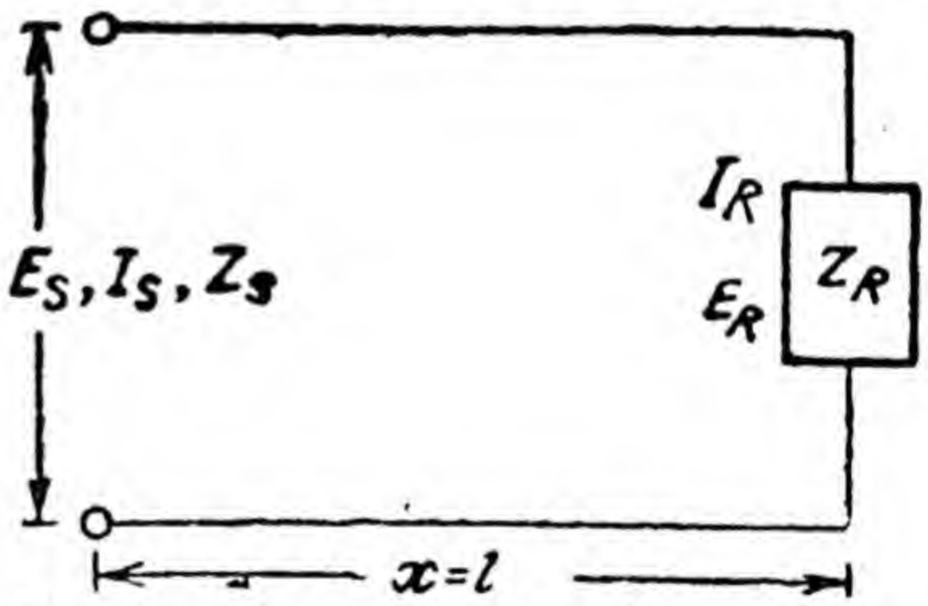


Fig. 18 (*a*), Transmission line terminated with impedance Z_R.

Voltage and current at the input terminals are

$$E_S = E_R \cosh \gamma l + I_R Z_0 \sinh \gamma l \quad \ldots(1)$$

and

$$I_S = I_R \cosh \gamma l + E_R/Z_0 \sinh \gamma l. \quad \ldots(2)$$

The sending end impedance will be

$$Z_s = \frac{E_s}{I_s} = \frac{Z\,[Z_R \cosh \gamma l + Z_0 \sinh \gamma l]}{[Z_0 \cosh \gamma l + Z_R \sinh \gamma l]}.$$

For loss-less line $R=G=0$ and $\alpha=0$. so that

$$\gamma = \alpha + j\beta = j\beta = j\, 2\pi/\lambda.$$

Input impedance (Z_s) is given by

$$Z_s' = \frac{Z_0 \left[Z_R \cosh \frac{2\pi jl}{\lambda} + Z_0 \sinh \frac{2\pi jl}{\lambda} \right]}{\left[Z_0 \cosh \frac{2\pi jl}{\lambda} + Z_R \sinh \frac{2\pi jl}{\lambda} \right]}.$$

(i) For quarter wave line : In case of quarter wave-line $\lambda=4l$.

$$\gamma=\frac{j.2\pi}{4l}=\frac{j\pi}{2l}.$$

We know that

$$Z_s=\frac{Z_0\,[Z_R\cosh\gamma l+Z_0\sinh\gamma l]}{[Z_0\cosh\gamma l+Z_R\sinh\gamma l]}$$

$$=\frac{Z_0\left[\frac{Z_R}{Z_0}+\tanh\gamma l\right]}{\left[1+\frac{Z_R}{Z_0}\tanh\gamma l\right]}$$

Hence Z_s' is given by

$$Z_s'=\frac{Z_0\left[\frac{Z_R}{Z_0}+\tanh\frac{j\pi l}{l}\right]}{\left[1+\frac{Z_R}{Z_0}\tanh\frac{j\pi l}{2l}\right]}=\frac{Z_0\left[\left(-j\cot\frac{\pi}{2}\right)\frac{Z_R}{Z_0}+1\right]}{\left[\left(-j\cot\frac{\pi}{2}\right)+\frac{Z_R}{Z_0}\right]}$$

$$=\frac{Z_0^2}{Z_R},$$

$$Z_0^2=Z_R\times Z_s',$$

$$Z_0=\sqrt{\{Z_RZ_s'\}}.$$

This equation shows that if *characteristic impedance of the line is the geometric mean of two impedances, these two impedances will be matched by the quarter wave line.*

From relation

$$Z_s'=\frac{Z_0^2}{Z_R},$$

it is obvious that the line transforms an impedance of large magnitude into one smaller magnitude and *vice versa* Thus a lossless quarter wave line acts as an impedance changing device *i e* , whatever be the terminating impedance, an inverse impedance will always appear at the input terminals Thus quarter wave line is an *impedance transformer.* Refer to solved numerical problem 9.

Uses of Quarter Wave Line :

(*a*) *As a matching device* : Since a quarter wave line acts as impedance transformer, it can be used to match a transmission line to an antenna. Refer to numerical example 9. Load impedance Z_R as well as Z_s' is given. We then connect a $\lambda/4$ line (in between Z_s' and Z_R) having $Z_0=\sqrt{(Z_RZ_s')}$. Z_0 can be calculated by the relation $Z_0=276\log s/r$ for two wire parallel line, s is separation between the centres of two wires and r is radius of the wire.

(*b*) *As a harmonic suppressor* : It can be seen that quarter wave line also suppresses the harmonics of the wave frequency. If the line is short circuited $Z_R=0$, then input impedance is infinity ($Z_s'=\infty$). Therefore, if *across* a transmission line feeding an

antenna, a quarter wave line is connected, then since $Z_s' = \infty$, there will be no effect in transmission at fundamental frequency. But the same $\lambda/4$ line will behave as half wave line at second harmonic of the frequency, and because for half wave line $Z_s' = Z_R$, Z_s' will become zero. As a result the voltage that correspond to second harmonic will be short circuited at the input terminals of quarter wave line and will not be delivered to the antenna. In other words it will suppress the second harmonic.

(c) *As an insulator to support an open wire line or the inner conductor of a coaxial line :*

Input impedance of a quarter wave shorted line is very high, voltage distribution on the transmission line will not be affected by connecting such a $\lambda/4$ line across the transmission line. Consequently they act as insulators.

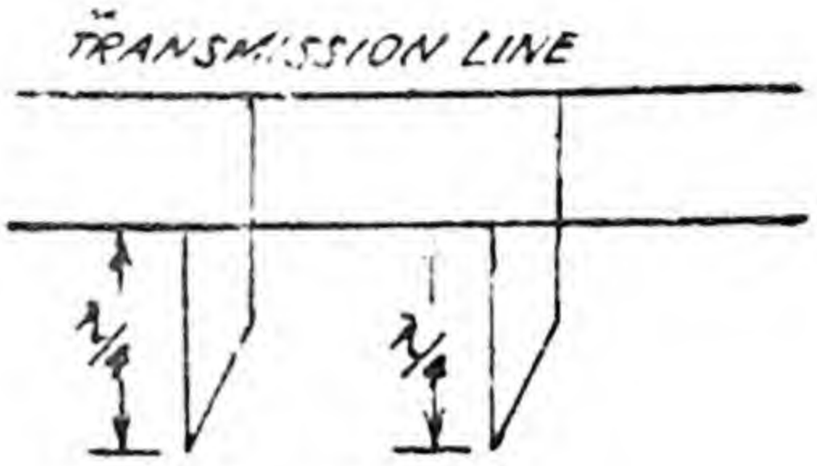


Fig. 18 (*b*).

(ii) For half wave line $\lambda = 2l$.

$$\gamma = \frac{j2\pi}{\lambda} = \frac{j\pi}{2l} = \frac{j\pi}{l},$$

$$Z_s' = \frac{Z_0 \left[\frac{Z_R}{Z_0} + \tanh \frac{j\pi l}{l}\right]}{\left[1 + \frac{Z_R}{Z_0} \tanh \frac{j\pi l}{l}\right]}$$

$$= \frac{Z_R Z_0}{Z_0} = Z_R,$$

$$Z_s' = Z_R.$$

Thus the input impedance of a half wave line is equal to the load impedance. The load voltage equals the input voltage in magnitude but is opposite in polarity†. The *behaviour is quite equivalent to that of* 1 : 1 *transformer* and hence a half wave line can be considered as 1 : 1 transformer.

Case (B) Input impedance of a line when it is short circuited :

We have derived the expresssion

$$Z_s = \frac{Z_0 [Z_R \cosh \gamma l + Z_0 \sinh \gamma l]}{[Z_0 \cosh \gamma l + Z_R \sinh \gamma l]}.$$

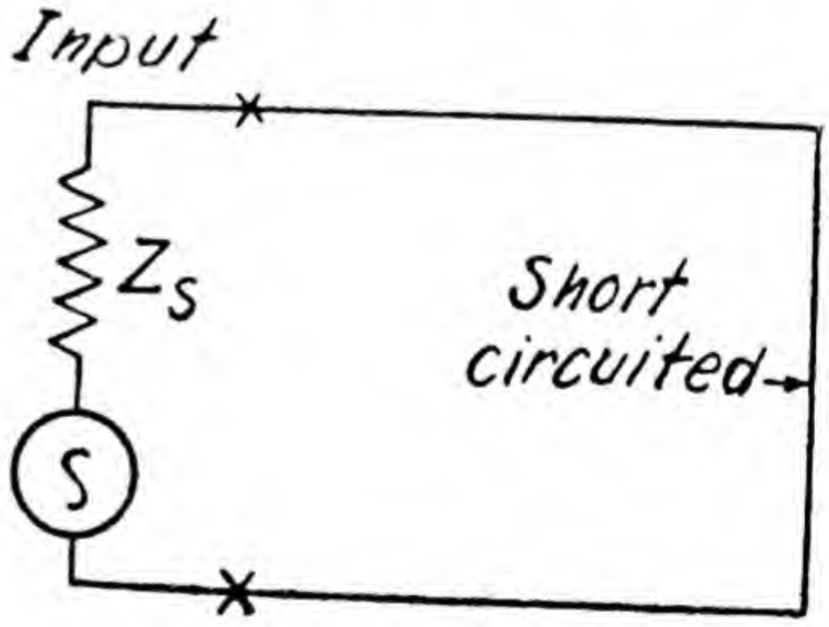


Fig. 19. Short-circuited transmission line.

† $E_s = E_R \cos \beta l + jI_R Z_0 \sin \beta l$

For $l = \lambda/2$,

$E_s = -E_R$.

When the line is short-circuited, $Z_R=0$, so that

$$\left[Z_s \right]_{Z_R=0} = \frac{Z_0 \left[+ Z_0 \sinh \gamma l \right]}{Z_0 \cosh \gamma l} .$$

$$= \frac{Z_0 \sinh \gamma l}{\cosh \gamma l}$$

If the line is loss less, then

$$\left[Z_s' \right]_{Z_R=0} = \frac{Z_0 \sinh \frac{2\pi j l}{\lambda}}{\cosh \left(\frac{2\pi j l}{\lambda} \right)} = j Z_0 \tan \left(\frac{2\pi l}{\lambda} \right)$$

Moreover the input impedance of a line when it is terminated such that $l=\lambda/4$, will be

$$\left[Z_s' \right]_{Z_R=0}^{\lambda/4} = j Z_0 \tan \left(\frac{2\pi}{\lambda} \frac{\lambda}{4} \right) = Z_0 \tan \pi/2 = \infty.$$

herefore, a loss less short circuited quarter wave line puts an infinite impedance at the generator.

The input impedance of a short-circuited line acts as a pure inductance upto the quarter wavelength of the line. The ratio Z_s'/Z_0 as a function βl is plotted in fig. 20 (*a*).

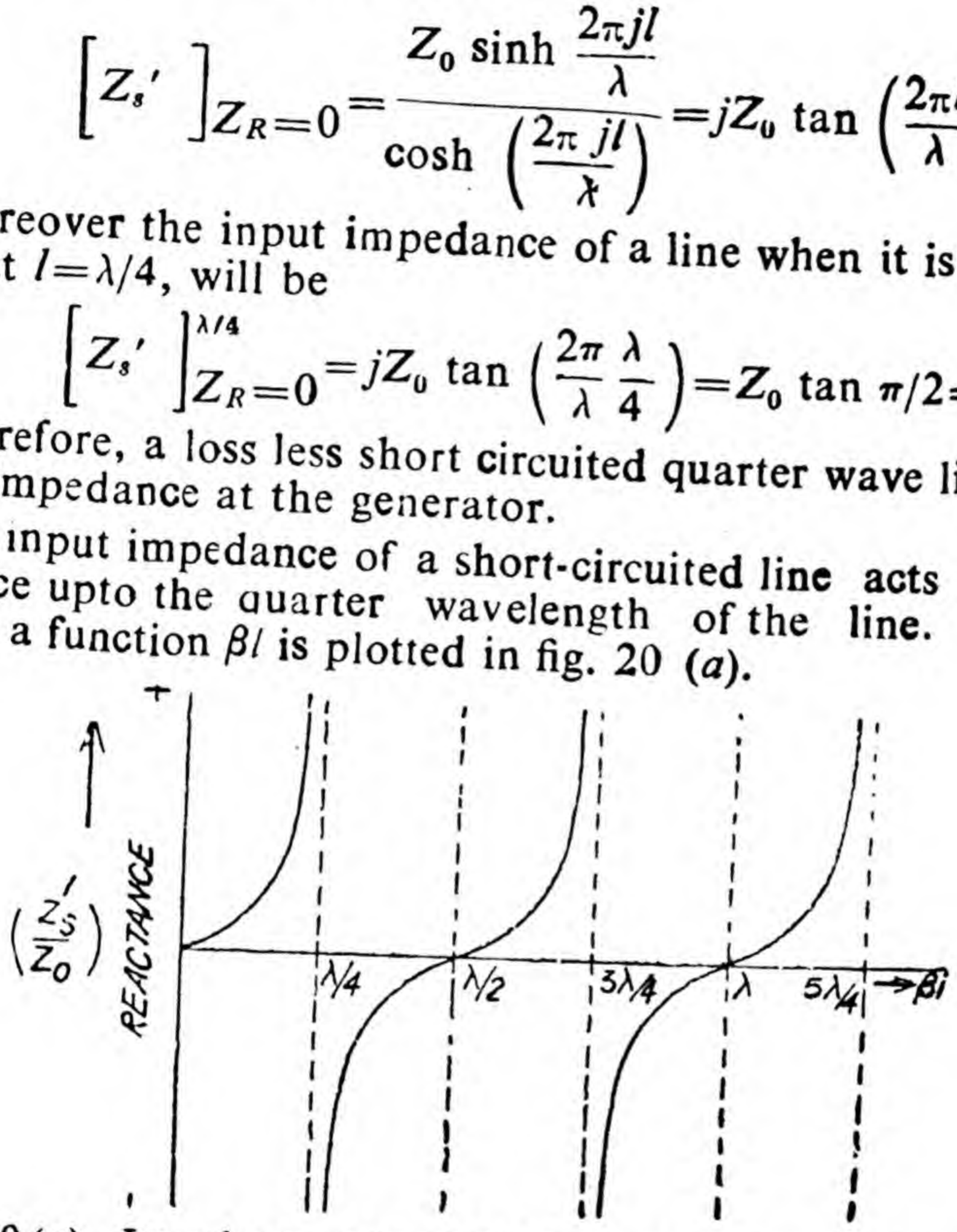


Fig. 20 (*a*) Impedance ratio Z_s'/Z_0 for a short-circuited line as a function of βl.

Case (C) : When the line is open-circuited :

If the line is open circuited at its far end $Z_R=\infty$, the input impedance can be obtained from the general expression for Z_s.

$$\left[Z_s \right]_{Z_R=\infty} = Z_0 \left[\frac{Z_R \cosh \gamma l}{Z_R \sinh \gamma l} \right] = Z_0 \coth \gamma l$$

If the line is lossless,

$$\left[Z_s' \right]_{Z_R=\infty} = Z_0 \coth \left(\frac{2\pi j l}{\lambda} \right) = -j Z_0 \cot \frac{2\pi l}{\lambda}.$$

If the line is terminated at $l=\lambda/4$,

$$\left[Z_s' \right]_{Z_R=\infty}^{\lambda/4} = -j Z_0 \cot \left(\frac{2\pi}{\lambda} \cdot \frac{\lambda}{4} \right) = 0.$$

Hence an open-circuited lossless $\lambda/4$ line presents zero impedance at the generator end.

The input impedance of an open-circuited line acts as pure capacitance upto the quarter wavelength of the line. Ratio $\frac{Z_s'}{Z_0}$ as function of βl in this case is plotted in fig. 20 (*b*).

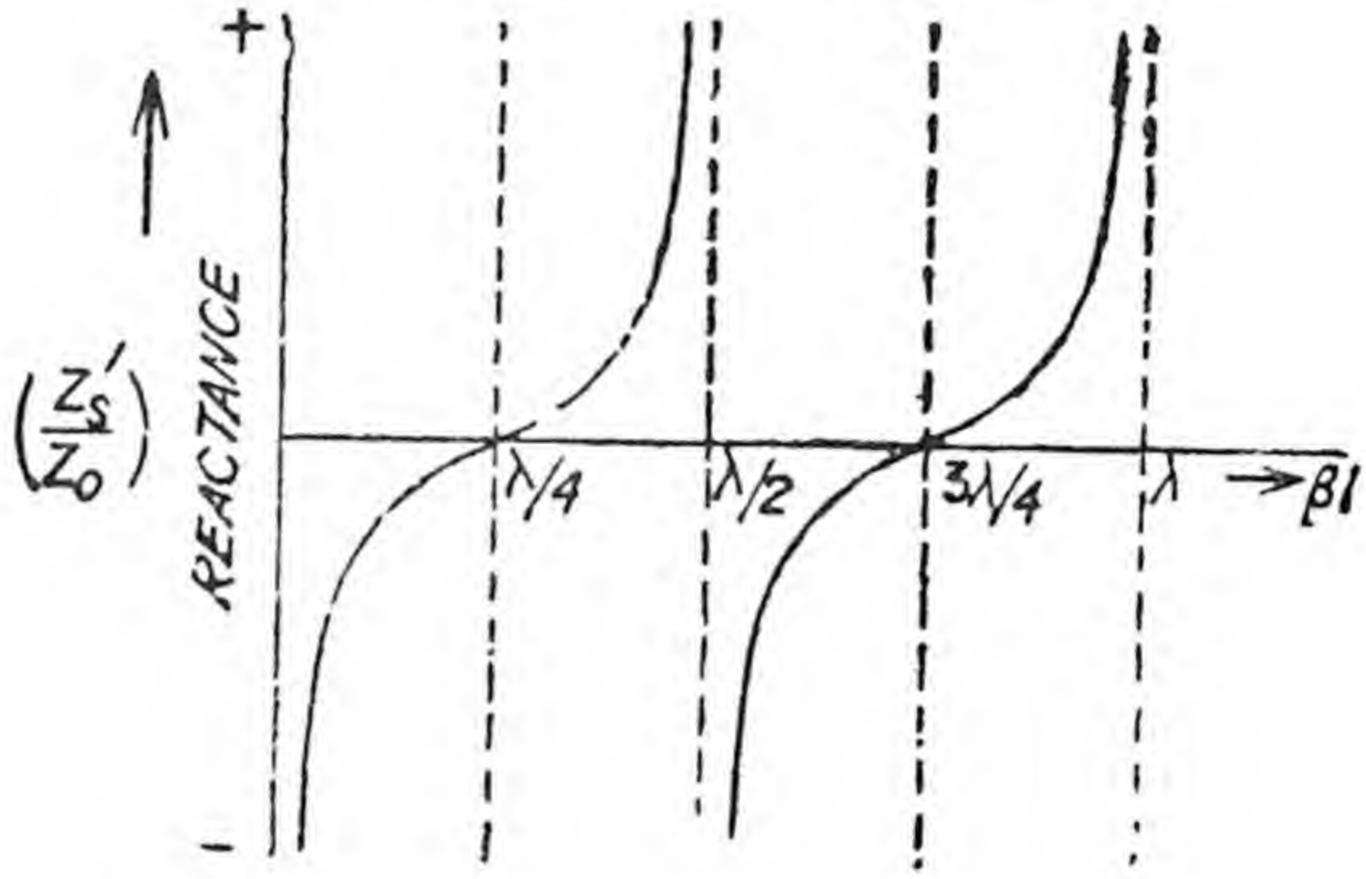


Fig. 20 (*b*). Impedance ratio $\frac{Z_s'}{Z_0}$ for open-circuited transmission line as a function of βl.

26·5-2. OPEN AND SHORT-CIRCUITED IMPEDANCES WHEN CONSIDERING DISSIPATION :

Voltage and current equations are

$$E = Ae^{\gamma l} + Be^{-\gamma l}$$

$$= A\left[e^{\gamma l} + \frac{B}{A} e^{-\gamma l} \right] = \frac{I_R}{2}(Z_R + Z_0)\left[e^{\gamma l} + \frac{Z_R - Z_0}{Z_R + Z_0} e^{-\gamma l} \right]$$

$$= \frac{E_R (Z_R + Z_0)}{2Z_R} [e^{\gamma l} + |\rho| e^{-\gamma l}]$$

and $$I = \frac{I_R (Z_R + Z_0)}{2Z_0} [e^{\gamma l} - |\rho| e^{-\gamma l}].$$

Further we have shown that

$$\alpha = \frac{R}{2Z_0},$$

then if R, the line resistance per meter length, is small with respect to $2Z_0$, and the line is short, then $e^{\gamma l}$ terms may be written as

$$e^{\gamma l} = e^{\alpha l}\, e^{j\beta l}$$

$$= (1 + \alpha l)(\cos \beta l + j \sin \beta l)$$

and $$e^{-\gamma l} = (1 - \alpha l)(\cos \beta l - j \sin \beta l),$$

so that

$$E = \frac{E_R}{2}\left(1 + \frac{Z_0}{Z_R} \right)\Big[(1 + \alpha l)(\cos \beta l + j \sin \beta l),$$

$$+ \left(\frac{Z_R - Z_0}{Z_R + Z_0} \right)(1 - \alpha l)(\cos \beta l - j \sin \beta l) \Big]$$

$$=E_R\left[\left(1+\alpha l\frac{Z_0}{Z_R}\right)\cos\beta l+j\left(\alpha l+\frac{Z_0}{Z_R}\right)\sin\beta l\right] \quad ...(1)$$

and $I=E_R/Z_0\left[\left(\alpha l+\frac{Z_0}{Z_R}\right)\cos\beta l+j\left(1+\alpha l\frac{Z_0}{Z_R}\right)\sin\beta l\right]$ ...(2)

Then input mpedance of this line of length l, will be

$$Z_s=\frac{E}{I}=Z_0\left[\frac{(Z_R+\alpha lZ_0)\cos\beta l+j(Z_R\alpha l+Z_0)\sin\beta l}{(Z_R\alpha l+Z_0)\cos\beta l+j(Z_R+\alpha lZ_0)\sin\beta l}\right] \quad ...(3)$$

SHORT CIRCUITED LINE

When the line is short circuited, equation (3) with $Z_R=0$ will become

$$(Z_s)_{short\ cct.}=Z_0\left(\frac{\alpha l\cos\beta l+j\sin\beta l}{\cos\beta l+j\alpha l\sin\beta l}\right) \quad ..(4)$$

$$=Z_0\frac{(\alpha l\cos\beta l+j\sin\beta l)(\cos\beta l+j\alpha l\sin\beta l)}{(\cos^2\beta l+\alpha^2l^2\sin^2\beta l)}$$

$$=Z_0\frac{[\alpha l\cos^2\beta l+\alpha l\sin^2\beta l+j(1-\alpha^2l^2)\sin\beta l\cos\beta l]}{(\cos^2\beta l+\alpha^2l^2\sin^2\beta l)}$$

$$=Z_0\frac{[\alpha l+j(1-\alpha^2l^2)\sin\beta l\cos\beta l]}{1-(1-\alpha^2l^2)\sin^2\beta l}$$

$$=\frac{Z_0.\alpha l}{1-(1-\alpha^2l^2)\sin^2\left(\frac{2\pi l}{\lambda}\right)}+\frac{jZ\ (1-\alpha^2l^2)\sin(2\pi l/\lambda)\cos(2\pi l/\lambda)}{1-(1-\alpha^2l^2)\sin^2\beta l}$$

$$=(R_s)_{short\ ccted}+j(X_s)_{short\ ccted}.$$

This input reactance of a short circuited line is

$$X_s=\frac{Z_0(1-\alpha^2l^2)\sin((2\pi l/\lambda)\cos(2\pi l/\lambda)}{1-(1-\alpha^2l^2)\sin^2\beta l},$$

which is inductive for l lying between 0 and $\lambda/4$, capacitive wnen l lies between $\lambda/4$ and $\lambda/2$ and so alternately. Thus it shows the same nature as shown by input reactance of a dissipationless short circuited line; tne difference is that for dissipationless line input impedance becomes infinite whereas for dissipation line input reaches a maximum but does not become infinite.

Variation of resistance along the points of short circuited quarter wave line with dissipation : We have shown above that

$$R_S=\frac{Z_o\alpha l}{1-(1-\alpha^2l^2)\sin^2\left(\frac{2\pi l}{\lambda}\right)},$$

so that the resistive impedance seen at the open end of a short circuited $\lambda/4$ line is

$$R_S=\frac{Z_0.\alpha.\lambda/4}{\left(1+\frac{\alpha^2\lambda^2}{16}\right)\sin^2\pi/2}=\frac{Z_0.\alpha\lambda/4}{\alpha^2\lambda^2/16}=\frac{4Z_0}{\alpha\lambda}.$$

The impedance at the short circuited end will be zero. There-

fore on any point between short and open ends of $\lambda/4$ line, resistive impedance will be intermediate to $\frac{4Z_0}{\alpha\lambda}$ and zero. We shall now find an expression to predict the variation of impedance along the line. We shall find Z_d at the point distant d from short end. As is obvious, we can take line to be made of two lines : one d distance long but short circuited on one end, the other open circuited on both the ends and $(\lambda/4-d)$ distance long The admittance at points XX due to short circuited line is

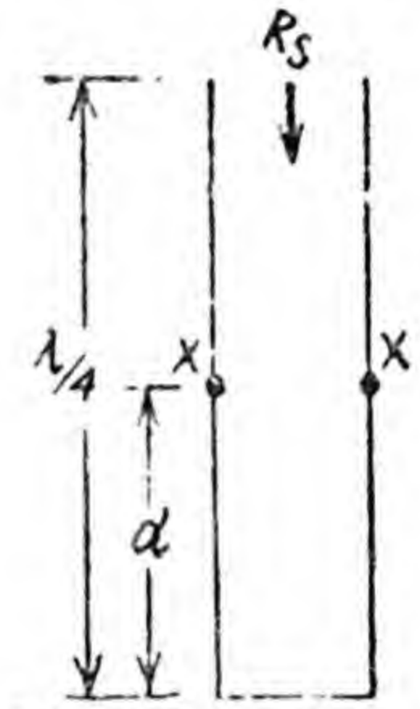


Fig. 21. (a)

$$Y_d' = \frac{1}{Z_0}\left[\frac{\cos \beta l + j\alpha l \sin \beta l}{\alpha l \cos \beta l + j \sin \beta l}\right]$$

with $l=d$, and admittance at points XX due to open circuited line is

$$Y_d'' = \frac{1}{Z_0}\left[\frac{\alpha l \cos \beta l + j \sin \beta l}{\cos \beta l + j\alpha l \sin \beta l}\right]$$

with $l=(\lambda/4-d)$. The latter relation we shall calculate in the last of this article.

The total admittance at point XX will be then

$$Y_d = Y_d' + Y_d''$$

$$= \frac{1}{Z_0}\left[\frac{1+j\alpha d \tan\left(\frac{2\pi d}{\lambda}\right)}{\alpha d + j \tan\left(\frac{2\pi d}{\lambda}\right)} + \frac{\alpha(\lambda/4-d)+j\tan\left(\frac{2\pi}{\lambda}\right)(\lambda/4-d)}{1+j\alpha(\lambda/4-d)\tan\left(\frac{2\pi}{\lambda}\right)(\lambda/4-d)}\right]$$

But

$$\tan\frac{2\pi}{\lambda}\left(\frac{\lambda}{4}-d\right) = \cot\frac{2\pi d}{\lambda},$$

$$Y_d = \frac{1}{Z_0}\left[\frac{1+j\alpha a \tan\left(\frac{2\pi d}{\lambda}\right)}{\alpha d + j\tan\left(\frac{2\pi d}{\lambda}\right)} + \frac{\alpha(\lambda/4-d)+j\cot\left(\frac{2\pi d}{\lambda}\right)}{1+j\alpha(\lambda/4-d)\cot\left(\frac{2\pi d}{\lambda}\right)}\right]$$

Further

$$\tan\frac{2\pi d}{\lambda} + \cot\frac{2\pi d}{\lambda} = \frac{2}{\sin(4\pi d/\lambda)},$$

so that above expression reduces to

$$Y_d = \frac{1}{Z_0}\cdot\frac{j\alpha\lambda/2 \sin(4\pi d/\lambda)}{\alpha(2d-\lambda/4)+j\left[\tan\left(\frac{2\pi d}{\lambda}\right)+\alpha^2 d(\lambda/4-d)\cot\left(\frac{2\pi d}{\lambda}\right)\right]}$$

Since α will be small, we ignore the terms in the denominator that involve it. Then

$$Y_d = \frac{1}{Z_0} \cdot \frac{\alpha\lambda/2 \sin (4\pi d/\lambda)}{\tan (2\pi d/\lambda)}$$

or $$Z_d = Z_0 \cdot \frac{2 \sin (4\pi d/\lambda) \tan (2\pi d/\lambda)}{\alpha\lambda}$$

$$= \frac{4Z_0}{\alpha\lambda} \sin^2 \left(\frac{2\pi d}{\lambda}\right)$$

$$= R_S \sin^2 \left(\frac{2\pi d}{\lambda}\right)$$

This variation is plotted in the figure 21 (*b*). Any load R_L connected at the open end of the line will parallel the line impedance R_S so that resultant will be

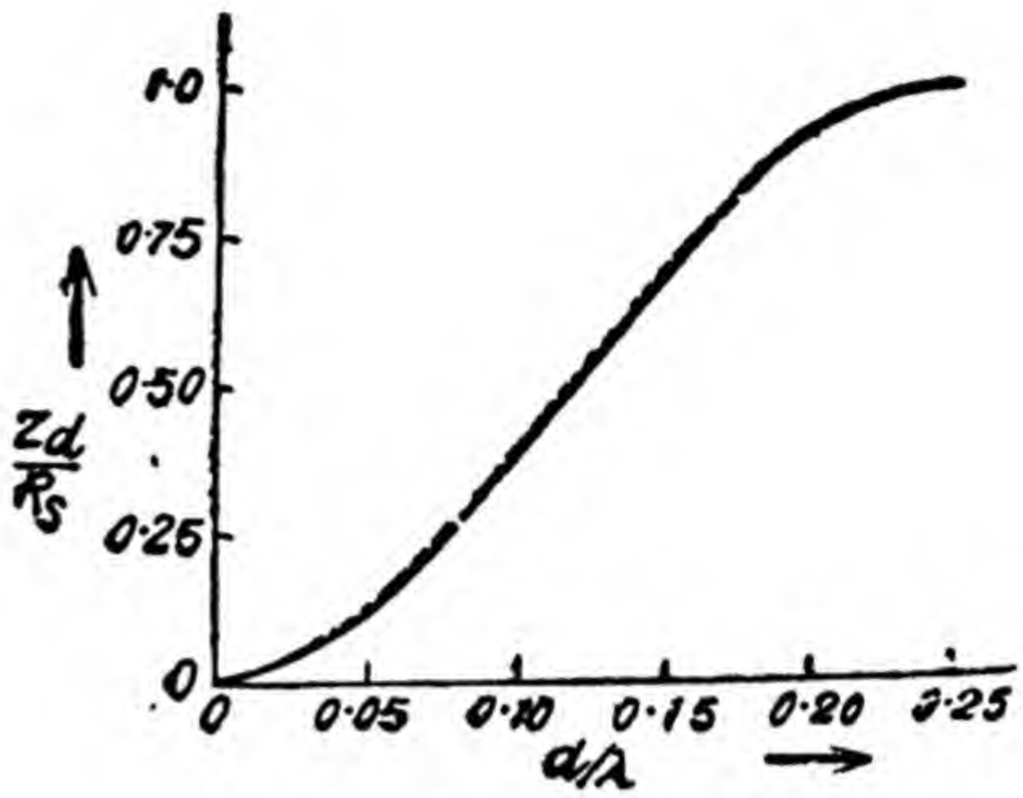


Fig. 21 (b)

$$\frac{1}{Z} = \frac{1}{R_L} + \frac{1}{R_S}$$

or $$\frac{1}{Z} = \frac{R_L + R_S}{R_L \times R_S} = \frac{(R_S + 4Z_0/\alpha\lambda)}{R_L \cdot \frac{4Z_0}{\alpha\lambda}}.$$

$$= \frac{\alpha\lambda\,(R_L + 4Z_0/\alpha\lambda)}{4Z_0 R_L}.$$

For any point along the line, this resultant would show same type of variation as predicted in the figure. Therefore the resultant resistive impedance seen at the top at any distance *d* will vary as

$$Z_d = Z \sin^2 \left(\frac{2\pi d}{\lambda}\right).$$

$$= \frac{4Z_0 R_L}{\alpha\lambda\,(R_L + 4Z_0/\alpha\lambda)} \sin^2 \left(\frac{2\pi d}{\lambda}\right)$$

$$= \frac{4Z_0}{\alpha\lambda} \left[\frac{1}{1 + \frac{4Z_0}{\alpha\lambda R_L}}\right] \sin^2 (2\pi d/\lambda).$$

The tapped $\lambda/4$ is employed for impedance matching between an antenna, or several antennas, and a transmission line.

OPEN CIRCUITED LINE :

We divide the numerator and denominator of equation (3) by Z_R and then put $Z_R = \infty$ to give

$$(Z_S)_{open\ ccted} = Z_0 \left(\frac{\cos \beta l + j\alpha l \sin \beta l}{\alpha l \cos \beta l + j \sin \beta l}\right)$$

$$= Z_0 \left[\frac{\alpha l - j\,(1 - \alpha^2 l^2) \sin (2\pi l/\lambda) \cos (2\pi l/\lambda)}{1 - (1 - \alpha^2 l^2) \cos^2 (2\pi l/\lambda)}\right]$$

$$= \frac{Z_0 \alpha l}{1-(1-\alpha^2 l^2)\cos^2(2\pi l/\lambda)} + \frac{-jZ_0(1-\alpha^2 l^2)\sin(2\pi l/\lambda)\cos(2\pi l/\lambda)}{1-(1-\alpha^2 l^2)\cos^2(2\pi l/\lambda)}$$

$$= (R_S)_{open\ ccted} + j\,(X_S)_{open\ ccted}.$$

where

$$(X_S)_{open\ ccted} = \frac{-Z_0(1-\alpha^2 l^2)\sin(2\pi l/\lambda)\cos(2\pi l/\lambda)}{1-(1-\alpha^2 l^2)\cos^2(2\pi l/\lambda}$$

For values of line lengths between 0 and $\lambda/4$ this reactance is capacitive, for values of l between $\lambda/4$ and $\lambda/2$, it is inductive, and so on alternately. It would attain a maximum value but never approaches infinity and that is main difference with open circuited dissipationless line input impedance.

26·6. CIRCLE DIAGRAM FOR DISSIPATIONLESS LINE :

Because of the simple relationship existing between reflection coefficient at the load and the standing wave pattern, a problem of importance in connection with the transmission line is the determination of reflection coefficient, ρ, when load impedance Z_R is given and *vice versa*. But since input impedance Z_s of the line is expressed as

$$Z_s = \frac{E_s}{I_s} = \frac{Ae^{j\beta s} + Be^{-j\beta s}}{\frac{1}{Z_0}(Ae^{j\beta s} - Ae^{-j\beta s})}$$

$$= Z_0 \cdot \frac{\left(1+\frac{B}{A}e^{-2j\beta s}\right)}{\left(1-\frac{B}{A}e^{-2j\beta s}\right)}$$

$$= Z_0 \cdot \frac{(1+|\rho|\,e^{j\phi}\,e^{-2j\beta s})}{(1-|\rho|\,e^{j\phi}\,e^{-2j\beta s})}$$

or

$$\frac{Z_s}{Z_0} = \frac{1+|\rho| \angle (\phi - 2\beta s)}{1-|\rho| \angle (\phi - 2\beta s)}, \qquad \ldots(1)$$

it becomes obvious that determination of reflection coefficient would mean the determination of input impedance of the line. Therefore, it is the problem that concerns the design of dissipation less lines. Further, we know that most easily measured line quantity is the standing wave ratio S. Equation (1) is

$$\frac{Z_s}{Z_0} = \frac{1+\left(\frac{S-1}{S+1}\right) \angle (\phi - 2\beta s)}{1-\left(\frac{S-1}{S+1}\right) \angle (\phi - 2\beta s)} \qquad \ldots(2)$$

which stresses that if standing wave ratio is given with the electrical length β. (or actual length s) of the line, one can compute Z_s, the input impedance.

To simplify such a design problem, we sketch a circle diagram that makes use of the input impedance equation for a dissipationless line written on a per unit basis *i.e.* equation (2). Since Z_s/Z_0 is complex, it is possible to write

$$\frac{Z_s}{Z_0}=r_a+jx_a,$$

where r_a and x_a are the values of resistance or reactance per unit of Z_0. Therefore equation (2) is

$$r_a+jx_a=\frac{1+\left(\frac{S-1}{S+1}\right)\angle(\phi-2\beta s)}{1-\left(\frac{S-1}{S+1}\right)\angle(\phi-2\beta s)}$$

or $$(r_a+jx_a)\left\{1-\left(\frac{S-1}{S+1}\right)\angle(\phi-2\beta s)\right\}=1+\left(\frac{S-1}{S+1}\right)\angle(\phi-2\beta s)$$

or $$(r_a+1+jx_a)\left(\frac{S-1}{S+1}\right)\angle(\phi-2\beta s)=r_a+jx_a-1 \qquad ...(3)$$

or $$(r_a+1+jx_a)\left(\frac{S-1}{S+1}\right)[\cos(\phi-2\beta s)+j\sin(\phi-2\beta s)]$$

$$=(r_a-1)+jx_a,$$

Equating real and imaginary parts,

$$(r_a+1)\left(\frac{S-1}{S+1}\right)\cos(\phi-2\beta s)-x_a\left(\frac{S-1}{S+1}\right)\sin(\phi-2\beta s)=(r_a-1)$$

$$\left(\frac{S-1}{S+1}\right)(r_a+1)\sin(\phi-2\beta s)+x_a\left(\frac{S-1}{S+1}\right)\cos(\phi-2\beta s)=x_a.$$

Squaring and adding we get after clearing of fractions,

$$r_a^2-r_a\left(\frac{S^2+1}{S}\right)+x_a^2=-1,$$

or $$r_a^2-2r_a\frac{S^2+1}{2S}+\left(\frac{S^2+1}{2S}\right)^2+x_a^2=\left(\frac{S^2+1}{2S}\right)^2-1$$

or $$\left[r_a-\left(\frac{S^2+1}{2S}\right)\right]^2+x_a^2=\left(\frac{S^2-1}{2S}\right)^2 \qquad ...(4)$$

which is of the form,

$$(x-c)^2+y^2=r^2,$$

that represents a family of circles of radius r and with centres shifted c units from the origin on the positive x-axis. Comparing it with equation (4), if we plot r_a along x-axis and x_a along y-axis then shift of centre of circle on the positive r_a axis will be

$$c=\frac{S^2+1}{2S}=\frac{S+\frac{1}{S}}{2} \qquad ...(5)$$

and radius of circle

$$r=\frac{S^2-1}{2S}=\frac{S-\frac{1}{S}}{2} \qquad ...(6)$$

Now we plot these circles as shown in the diagram 22.

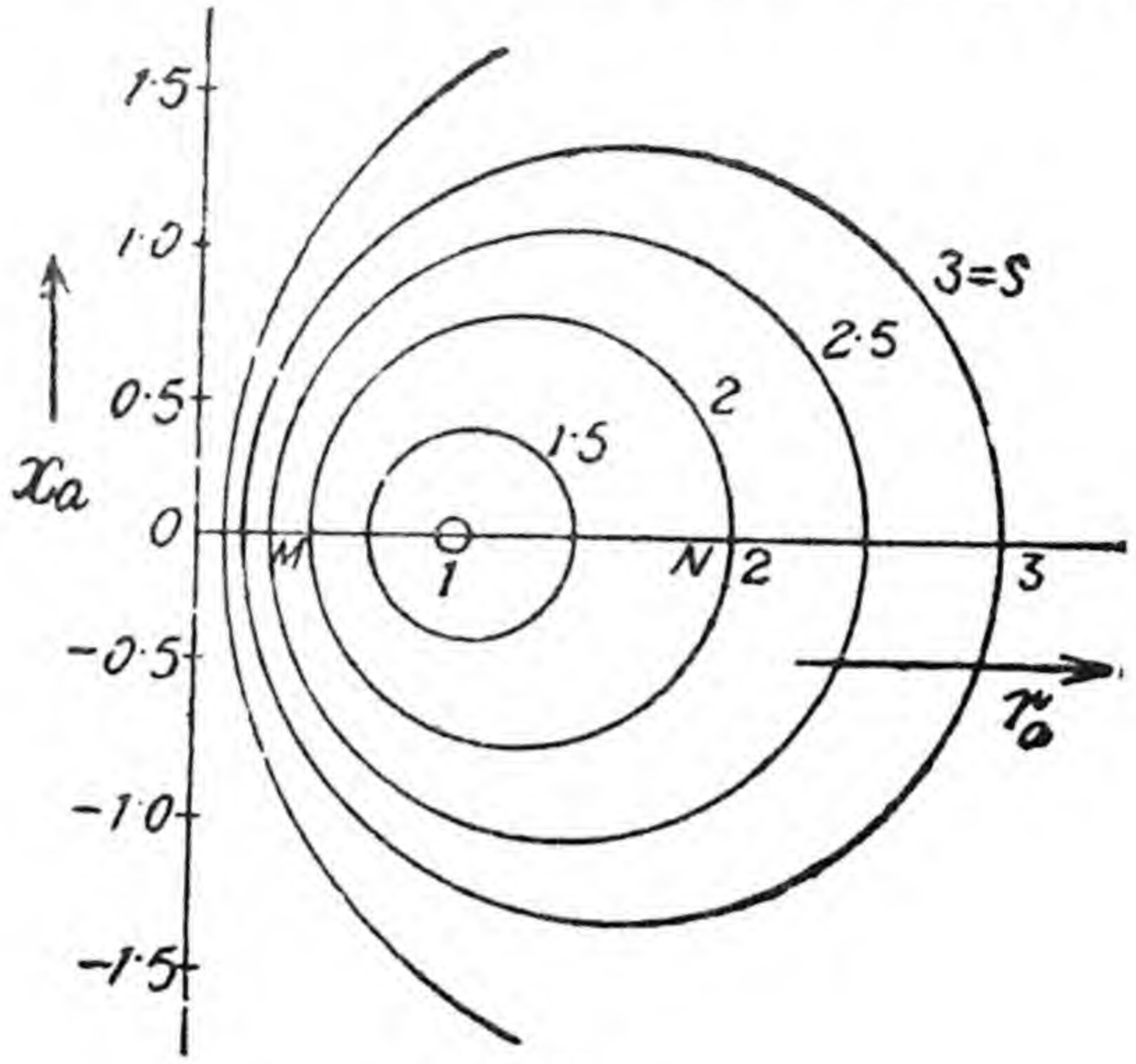


Fig. 22.

The minimum value of S is unity for which from equations (5) and (6),

$$c=1,\ r=0,$$

and lies at (1, 0) point. All the circles with greater value of S should therefore surround this circle (1, 0). The maximum value of S is infinity (open and short circuited lines). We observe that as S increases, radius of the S circle increases with its centre shifted more to the right and therefore for the limiting case $L\to\infty$, the circle becomes x_a axis. It is interesting to note that for any circle (say $S=2$) the intercept near the origin (say the M point) is at $\frac{1}{S}\left(OM=\frac{1}{2}\right)$, whereas intercept far from the origin (say N point) is at S units on the r_a axis ($ON=2$ units).

When $\frac{Z_s}{Z_0}$ lies on r_a axis, then at a point remote from the $r_a=S$, $x_a=0$ so that

$$\frac{Z_s}{Z_0}=S=\frac{1+|\rho|}{1-|\rho|},$$

which is possible only when, as compared to equation (1).

$$\phi-2\beta s=0. \qquad ...(7)$$

We find then

$$Z_s=Z_0.\ \frac{1+|\rho|}{1-|\rho|}$$

i.e. line impedance is purely resistive (it may be noted that at high frequencies or for dissipationless lines, R_0 is purely resistive) at the point $r_a=S$, $x_a=0$ on the S circle. At a point near to the origin $r_a=\frac{1}{S}$ on r_a axis, $\frac{Z_s}{Z_0}$ will assume a minimum value

$$\frac{Z_s}{Z_0}=\frac{1}{S}=\frac{1}{\frac{1+|\rho|}{1-|\rho|}}$$

$$=\frac{1-|\rho|}{1+|\rho|}.$$

We can obtain such a value of $\frac{Z_s}{Z_0}$ from equation (1) which is

$$\frac{Z_s}{Z_0}=\frac{1+|\rho|e^{j(\phi-2\beta s)}}{1-|\rho|e^{j(\phi-2\beta s)}},$$

only when phase angle $(\phi-2\beta s)$ is replaced by π, *i.e.*

$$\frac{Z_s}{Z_0}=\frac{1+|\rho|e^{+j\pi}}{1-|\rho|e^{+j\pi}}$$

$$=\frac{1-|\rho|}{1+|\rho|}$$

Now for resistive load $\rho=|\rho|$ so that $\phi=0$.

Therefore,

$$-2\beta s=\pi$$

$$\beta s=-\pi/2.$$

Thus in arriving at $\frac{Z_s}{Z_0}=\frac{1}{S}$ point, βs has undergone a change of $-\pi/2$. To represent such changes, we place a βs scale on the S circles. The βs scale increases *clockwise* or in the direction of increasing negative angles.

Rewriting equation (3),

$$\frac{(S-1)}{(S+1)}\angle(\phi-2\beta s)=\frac{r_a-1+jx_a}{r_a+1+jx_a}$$

$$=\frac{r_a^2-1+x_a^2+2jx_a}{(r_a+1)^2+x_a^2}. \qquad \ldots(8)$$

The angle ϕ may be made zero in order that βs scale may start at 0° on the abscissa (r_a axis) because it has already been pointed out that a point on r_a axis, line impedance is resistive for which ϕ is zero. Then

$$\frac{(S-1)}{(S+1)}e^{-2j\beta s}=\frac{r_a^2-1+x_a^2}{(r_a+1)^2-x_a^2}+\frac{j2x_a}{(r_a+1)^2+x_a^2},$$

so that

$$\frac{(S-1)}{(S+1)}\cos(-2\beta s)=\frac{r_a^2-1+x_a^2}{(r_a+1)^2+x_a^2}$$

and
$$\frac{(S-1)}{(S+1)}\sin(-2\beta s)=\frac{2x_a}{(r_a+1)^2+x_a^2}$$

giving
$$\tan(-2\beta s)=\frac{x_a}{r_a^2-1+x_a^2} \qquad \ldots(9a)$$

or
$$r_a^2-1+x_a^2+\frac{2x_a}{\tan 2\beta s}=0$$

or
$$r_a^2+\left(x_a+\frac{1}{\tan 2\beta s}\right)^2=1+\frac{1}{\tan^2 2\beta s}$$
$$=\frac{1}{\sin^2 2\beta s},$$

from which we infer that lines of equal βs value are circles of radius

$$=\frac{1}{\sin 2\beta s}, \qquad \ldots(9b)$$

with a shift of centre downward on x_a axis (ordinate).

$$=\frac{1}{\tan 2\beta s} \qquad \ldots(10)$$

A family of such circles is drawn in fig. 23. All βs circles pass through the point $r_a=1$, $x_a=0$.

Superposition of βs circles on S circles provides a scale of βs angles and results in the circle diagram of fig. 24.

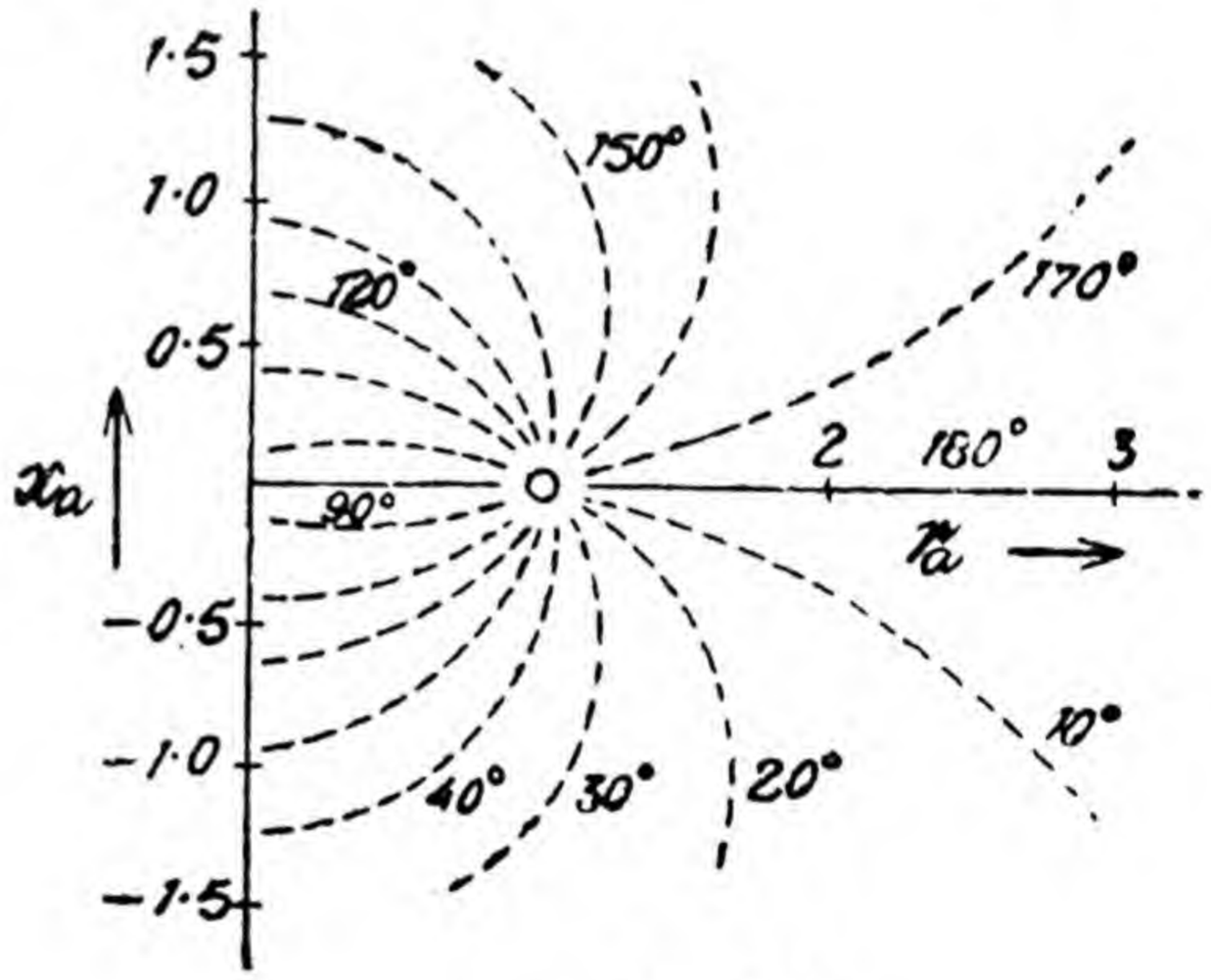


Fig. 23.

Application of circle diagram : Circle diagram may be used to find the input impedance of a line of any chosen length. For calculation, compute first the *per unit* value of the load impedance as

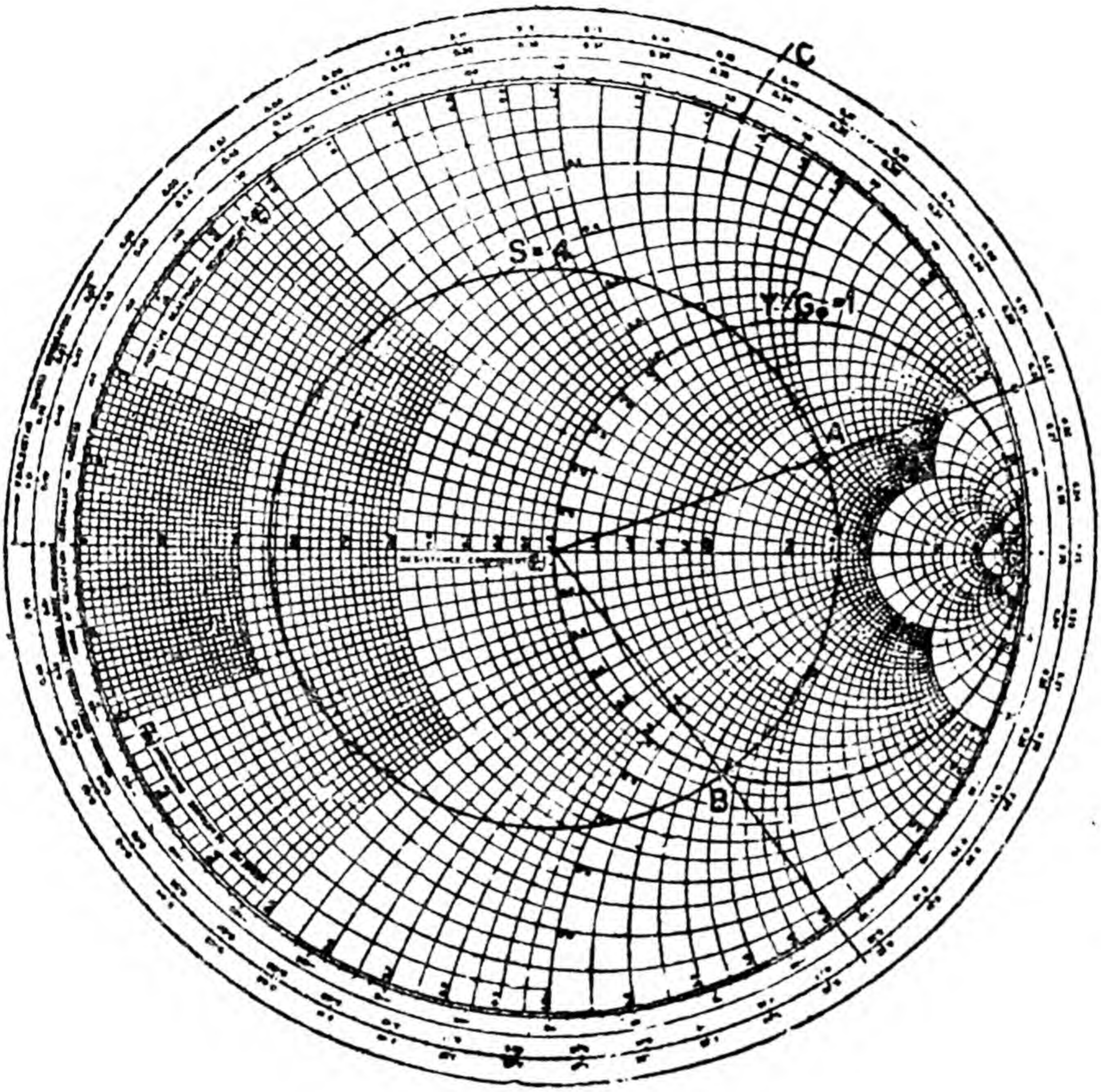
$$\frac{Z_R}{Z_0}=r_a+jx_a.$$

For example, let us take

$$\frac{Z_R}{Z_0}=2{\cdot}6+j.1,$$

the line is 28° long (βs value) and the standing wave ratio is found to be 3.

For this locate $(2{\cdot}6+j\cdot 1)$ point on $S=3$ circle (point $r_a=2{\cdot}6$, $x_a=1$). Now follow this circle in a clockwise direction through



F.g. 24.

28° equal to the electrical length of the line. Read the coordinates of the point so reached. This will provide per unit input impedance of the line, which in our case, as seen from fig. 25, is

$$\frac{Z_s}{R_0} = 1{\cdot}58 - j1{\cdot}35.$$

For an open circuited line $R=\infty$, S circle are the vertical lines appearing at the x_a axis. The input impedance of the line is then a pure reactance with different βs values provided by the intersection of βs circles with x_a axis.

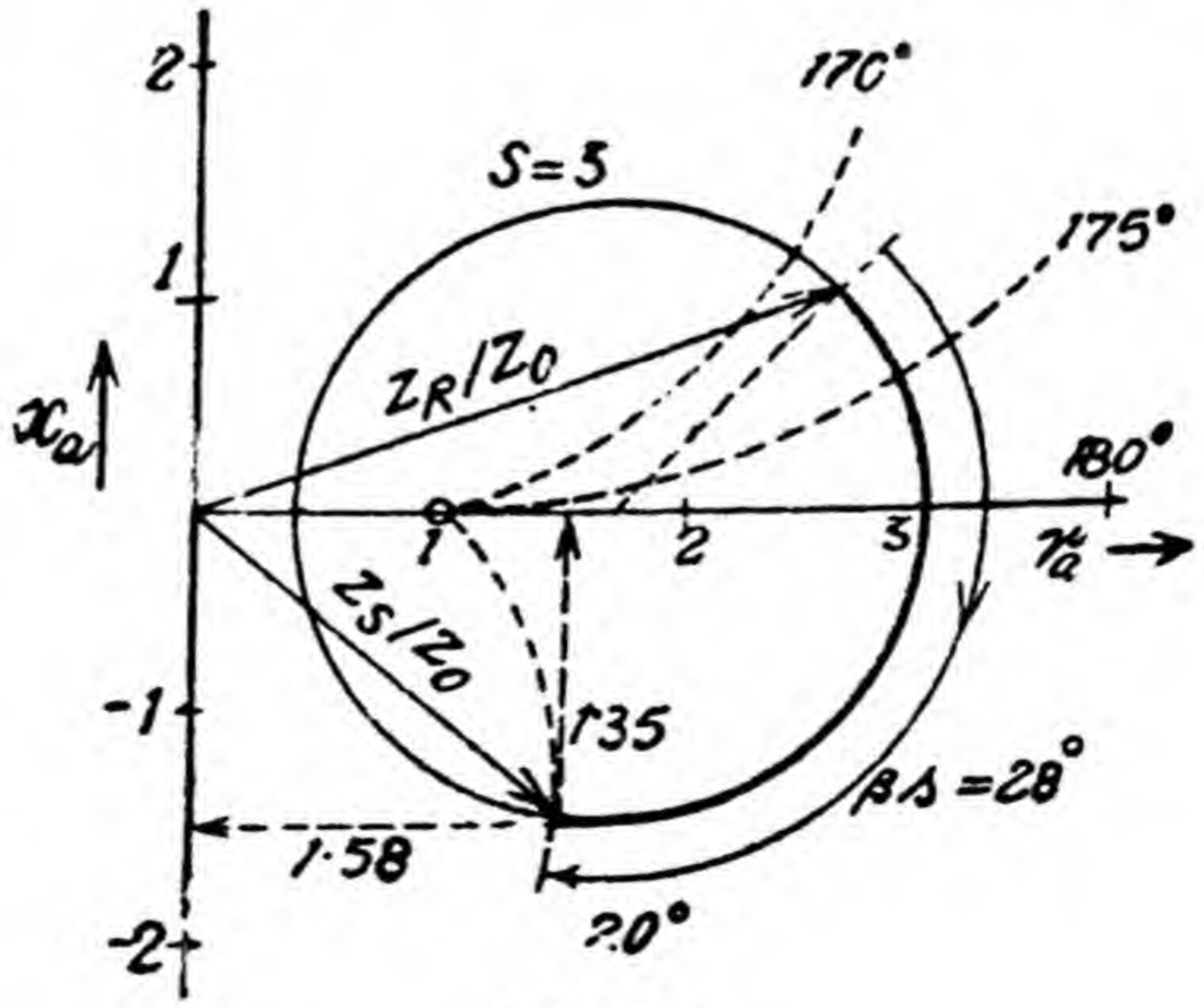


Fig. 25.

Basis of Smith circle diagram :

Smith circle diagram is a modified form of the circle diagram discussed above. From eqnation (8), we write

$$\frac{(S-1)}{(S+1)} \angle (\phi - 2\beta s) = |\rho| \angle (\phi - 2\beta s)$$

$$= \frac{r_a^2 - 1 + x_a^2 + 2jx_a}{(r_a+1)^2 + x_a^2}.$$

We introduce new variables $U+jV$ such that

$$U+jV = \frac{r_a^2 - 1 + x_a^2 + 2jx_a}{(r_a+1)^2 + x_a^2},$$

giving

$$U = \frac{r_a^2 - 1 + x_a^2}{(r_a+1)^2 + x_a^2} \qquad \ldots(11)$$

$$V = \frac{2x_a}{(r_a+1)^2 + x_a^2} \qquad \ldots(12)$$

Eliminating first x_a from (11) and then r_a from (12), we get

$$\left[U - \left(\frac{r_a}{r_a+1}\right)\right]^2 + V^2 = \frac{1}{(r_a+1)^2} \qquad \ldots(13)$$

$$(U-1)^2 + \left(V - \frac{1}{x_a}\right)^2 = \frac{1}{x_a^2}. \qquad \ldots(14)$$

Equation (13) represents a family of r_a circles with centres on U axis at $\frac{r_a}{(r_a+1)}$ and ratio of $\frac{1}{(r_a+1)}$. The second equation (14), represents a family of constant x_a circles with centres at $\left(1 + \frac{j}{x_a}\right)$ and radii equal to $\frac{1}{x_a}$. The two circle families are drawn in fig. 26. Since diagram is based on the assumption $|\rho| \angle (\phi - 2\beta s) = U + jV$, hence the maximum magnitude of $U+jV$ is fixed at unity by the maximum value of ρ. Thus all possible values of impedance are contained inside the outer circle of unit radius. The same relation between polar and rectangular coordinates fixes βs angles around $\rho = 1$ circle as shown in the fig. 26.

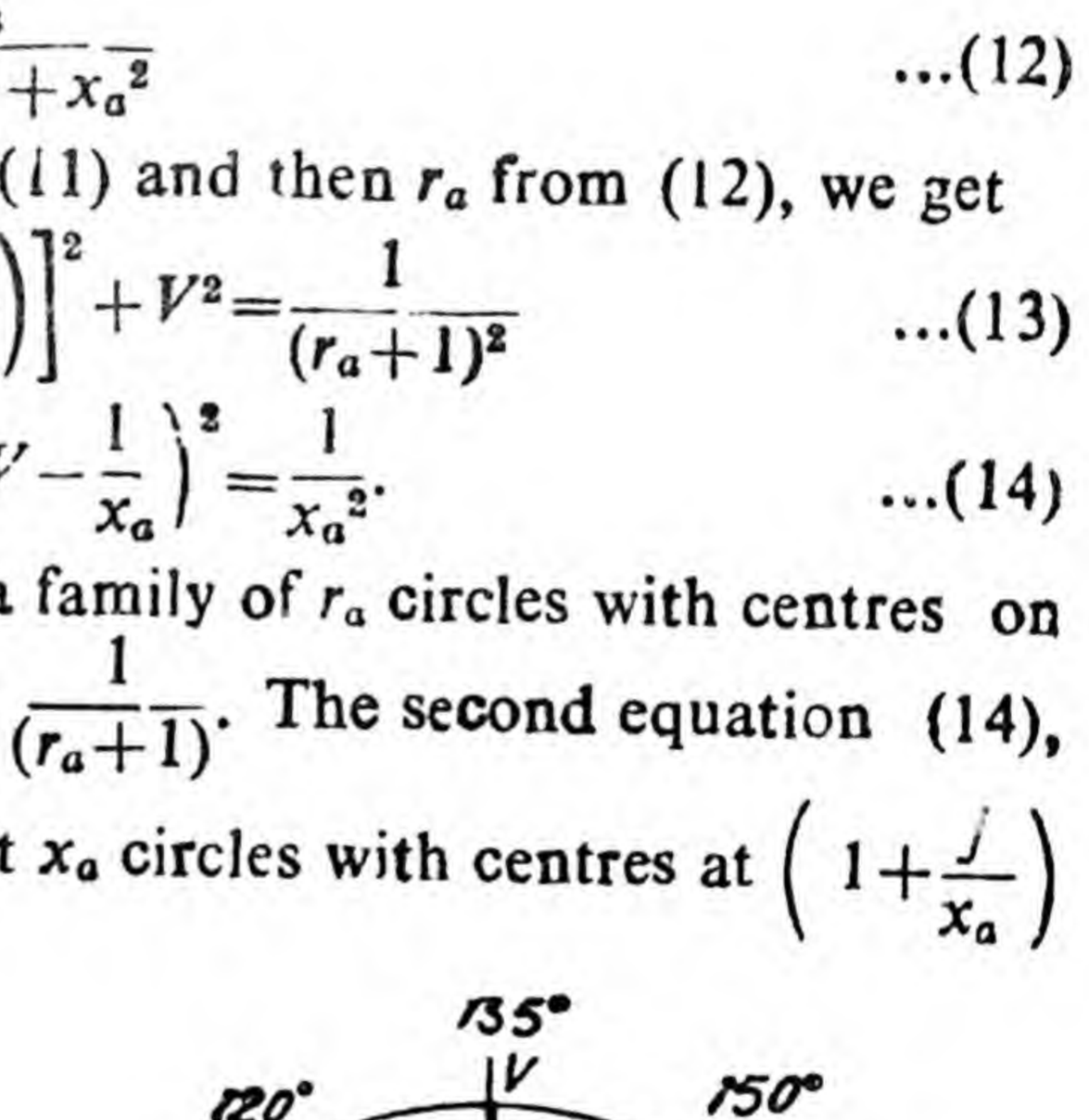


Fig. 26.

26·7. POWER : POWER CAPACITY AND REFLECTION LOSSES ON LOSSLESS LINES :

Of great importance in the application of transmission lines is the amount of power conveyed by a line from a source to a load.

For a dissipationless line ($\alpha=0$), the voltage and current equations are

$$E=\frac{I_R}{2}(Z_R+R_0)\,[e^{j\beta l}+|\rho|e^{-j\beta l}], \qquad ...(1)$$

$$I=\frac{I_R}{2}\left(\frac{Z_R+R_0}{R_0}\right)[e^{j\beta l}+|\rho|e^{-j\beta l}], \qquad ...(2)$$

The maximum of voltage along the line occurs at points at which the incident and reflected waves are in phase so that from equation (1),

$$E_{max}=\frac{I_R}{2}(Z_R+R_0)\,[1+|\rho|]. \qquad ...(3)$$

Likewise, voltage minimum occurs at points at which incident and reflected waves are out of phase so that

$$E_{min}=\frac{I_R}{2}(Z_R+R_0)\,[1-|\rho|]. \qquad ...(4)$$

The same is true for current maxima and minima at a point on the line so that

$$I_{max}=\frac{I_R}{2}\left(\frac{Z_R+R_0}{R_0}\right)[1+|\rho|] \qquad ...(5)$$

and

$$I_{min}=\frac{I_R}{2}\left(\frac{Z_R+R_0}{R_0}\right)[1-|\rho|]. \qquad ...(6)$$

From above equations, we find that

$$\frac{E_{max}}{I_{max}}=R_0 \text{ and } \frac{E_{min}}{I_{min}}=R_0,$$

predicting that ratio of extreme values of voltage and current on the line refers to a constant for the line given and not to input or load impedance.

Power and Power capacity :

From equations (1) and (2), we note that a voltage maximum and a current minimum occur at the same point on the lossless line. We also know the fact that at a point of maximum voltage on a lossless line, the total voltage and current are in phase and hence the power transmitted is

$$P=|E_{max}.||I_{min}.|.$$

But $|I_{min}.|=|I^+|-|I^-|$

$$=\frac{|E^+|}{R_0}-\frac{|E^-|}{R_0}$$

$$= \frac{|E^+| - |E^-|}{R_0}$$

$$= \frac{|E_{min.}|}{R_0}.$$

Therefore power transmitted on the line is

$$P = \frac{|E_{max.}| \cdot |E_{min.}|}{R_0}.$$

Putting the values of $|E_{max.}|$ and $|E_{min.}|$ interms of current and R_0 from equations (5) and (6) respectively, we find that

$$P = (|I_{max.}| \cdot |I_{min.}|).R_0.$$

These expressions for power permit easy measurements of power flow on a line of negligible loss.

As an application of power equation, we consider two examples :

(*i*) Consider a line terminated in an open circuit, a short circuit, or a reactance so that the standing wave ratio is infinite. Since for such cases reflection coefficient $|\rho|$ is one, power consumed by the load is zero. If the line is lossless, power delivered by the source and the transmitted power at each point of the line will be zero evaluated below.

Power delivered by tne source $= |E^+||I^+| \cos\theta$

$= 0,$

since phase between E^+ and I^+ is 90°.

Further $|E_{min.}| = |E^+| - |E^-| = 0$, power transmitted P is also zero.

(*ii*) Consider a line terminated in characteristic resistance. For such a line $|\rho| = 0$ and voltage and current at each point of the line are in phase and have the magnitudes $|E^+|$ and $|I^+|$. Therefore power transmitted

$$P = \frac{|E_{max.}||E_{min.}|}{R_0}$$

$$= |E^+||I^+|,$$

since $|E_{max.}| = |E^+| + |E^-| = |E^+|$, (as $E^- = \rho E^+ = 0$)

and $|E_{min.}| = |E^+| - |E^-| = |E^+|$,

so that $\frac{|E_{min}|}{R_0} = \frac{|E^+|}{R_0} = I^+.$

The maximum power that can be transmitted over a line without danger of insulation breakdown is called its power capacity. We have shown earlier (art. 26·3) that the high value of standing wave ratio is undesirable since it reduces the power capacity given by

$$P_{max.} = \frac{|E_{lim}|^2}{SR_0}.$$

Reflection losses on the unmatched line :

On a line, not terminated in charactristic impedance Z_0, reflected wave exists. These reflected waves, like to incident waves, cause losses on the line. Greater the difference between Z_R, the load impedance, and Z_0, the characteristic impedance, more stronger will be reflected wave and consequently, a larger loss due to reflection. The net effect of reflection being present on the line is the reduction in power delivered by the line to the load.

If we consider line at the radio frequencies $\omega L \gg R$ and $\omega C \gg G$, we can take $R=G=0$ approximately and hence $\alpha=0$. For such line characteristic impedance is purely resistive and we can denote it by R_0.

At a point of maximum voltage, reflected and incident voltages are in phase, so that

$$|E_{max.}| = |E^+| + |E^-|$$
$$= \frac{I_R(Z_R+R_0)}{2}(1+|\rho|),$$

and at a point of minimum voltage, reflected and incident voltages are in opposite phase so that

$$|E_{min.}| = |E^+| - |E^-|$$
$$= \frac{I_R(Z_R+Z_0)}{2}(1-|\rho|).$$

Hence the standing wave ratio is

$$S = \frac{|E_{max.}|}{|E_{min.}|} = \frac{|E^+|+|E^-|}{|E^+|-|E^-|}.$$

Now we know that the total power transmitted along the line and delivered to the toad is

$$P = \frac{|E_{max.}|\,|E_{min.}|}{R_0}$$
$$= \frac{(|E^+|+|E^-|).(|E^+|-|E^-|)}{R_0}$$
$$= \frac{|E^+|^2-|E^-|^2}{R_0}$$
$$= P^+ - P^-.$$

The ratio of the power P delivered to the load to the power transmitted by the incident wave is

$$\frac{P}{P^+} = \frac{P^+ - P^-}{P^+} = \frac{\frac{|E^+|^2}{R_0} - \frac{|E^-|^2}{R_0}}{\frac{|E^+|^2}{R_0}} = 1 - \frac{|E^-|^2}{|E^+|^2}$$
$$= 1-|\rho|^2$$
$$= 1-\left(\frac{S-1}{S+1}\right)^2$$

$$=\frac{4S}{(S+1)^2}.$$

Now when the line is terminated in characteristic resistance, reflection is absent, *i.e.*, $\rho=0$ and hence $S=\frac{1+|\rho|}{1-|\rho|}=1$. Therefore from above expression, we find that

$$\frac{P}{P^+}=\frac{4.1}{(1+1)^2}$$

or $$P=P^+,$$

i.e., whole of the power delivered by the source is absorbed by the load called an efficient terminating condition.

26·8. TRANSMISSION LINES AS RESONANT CIRCUITS :

Tranmission line can behave as series or parallel resonant circuit depending upon its way of operation. At high frequencies resonant lines are best substitute for resonant circuits because of superior performance.

At high frequencies, wavelength being sufficiently small, usual inductance coils and condensers cease to become usuful circuit elements. With the rise in frequency, the distributed capacitance existing between the turns of a coil offers a low impedance. Consequently, the coil acts more as a *complex transmission* line rather than as a lumped element with constant inductance. Further, at high frequencies losses are increased to an appreciable proportion due to skin effect, proximity effects and radiation of energy. Condensers also meet the same fate because of distributed inductance that arises due to the linkage of magnetic flux with the flow of current. Let us illustrate the point by taking an example. Suppose we could construct a coil of $1\mu H$, then capacitor required for resonance at Mc/s will have a value

$$C=\frac{1}{2\pi f X_C}=\frac{1}{2\pi f X_L} \text{ (since } X_L=X_C \text{ at resonance).}$$

But $$X_L=2\pi f L=6{\cdot}28\times 5\times 10^6\times 1\times 10^{-6}=3140\Omega$$

so that

$$C=\frac{1}{2\times 3{\cdot}14\times 5\times 10^6\times 3140}=\frac{10^4}{(314)^2}\cdot 10^{-12}$$
$$=0{\cdot}1\ pF.$$

But such a capacitance value is impossible. Its own stray capacitance would be more. In addition L/C ratio will be poor, the gain low and selectivity bad.

Due to the troublesome behaviour of the distributed effects of lumped elements, it is advisible to use distributed parameters in a simple geometry, *e.g.*, two wire line. The use of such short

line sections limits the losses to a minimum. Short line sections with reflecting termination *e g.*, short and open circuited lines, effectively serve as inductive and capacitive reactances and as resonant circuits at wavelengths below a few meters. The short circuit termination is most frequently used because it is easier to produce complete reflection in this way at smaller wavelengths. In open circuited lines losses of energy occurs due to radiation from open ends unless the spacing between conductors is quite small compared with a quarter wavelength.

Lines as parallel resonant circuits :

In parallel resonant circuit, impedance is maximum at resonant frequency and varies with frequency in the vicinity of resonance. Same behaviour is shown by a line *short circuited at the receiver and the length being an odd multiple of a quarter wavelength or open circuited at the receiver but even multiple of quarter wavelength long*. Such a line offers a high impedance at the sending end and therefore we say that at frequencies corresponding to $\beta s = n\pi/2$, where n is odd number in the case of short-circuited line or even number in open circuited line, as shown in fig. 27 (*a*) and (*b*), line

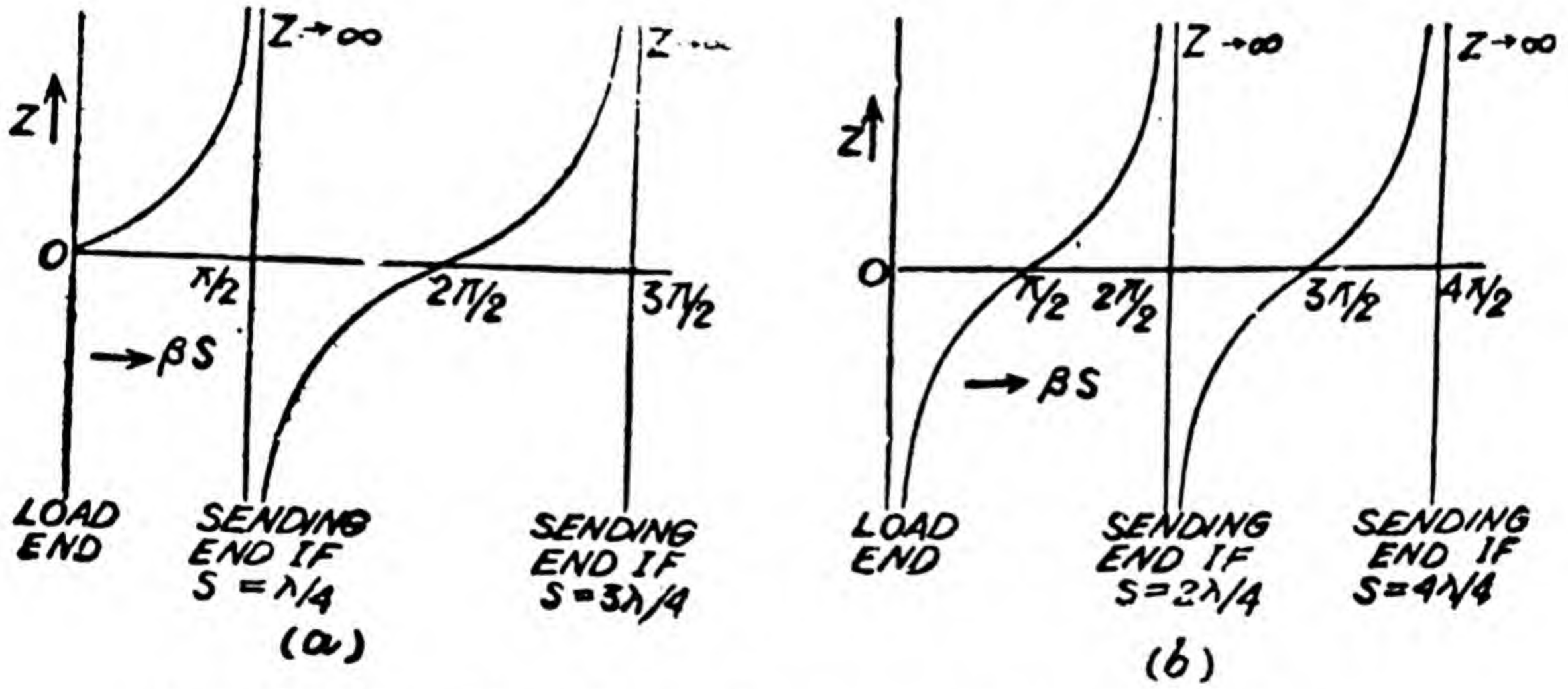


Fig. 27. Plot of impedance of open and short circuited line :
(*a*) for short-circuited line. (*b*) for open circuited line.

impedance is very high and these are termed as resonant frequencies This behaviour is similar to parallel resonant circuit. For both types of load conditions, the sending end impedance Z_S of the line at resonance is

$$Z_S = \frac{8Z_0^2 f_0}{Rnc}. \qquad ...(1)$$

where Z_0 = characteristic impedance of line,

f_0 = resonant frequency in cycles,

R = line resistance per unit length,

n = number of quarter wavelengths in line,

c = velocity of waves on the line taken to be equal to the velocity of light (unit of length as that of R).

The above expression can be derived as follows :

We know that input impedance of line can be expressed as

$$Z_S = \frac{Z_0 [Z_R \cosh \gamma s + Z_0 \sinh \gamma s]}{Z_0 \cosh \gamma s + Z_R \sinh \gamma s}$$

For short circuited line, Z_R is zero and there fore

$$Z_S = Z_0 \cdot \frac{Z_0 \sinh \gamma s}{Z_0 \cosh \gamma s} = Z_0 \frac{\sinh (\alpha + j\beta)\, s}{\cosh (\alpha + j\beta)\, s}$$

$$= Z_0 \cdot \frac{\sinh \alpha s \cos \beta s + j \cosh \alpha s \sin \beta s}{\cosh \alpha s \cos \beta s + j \sinh \alpha s \sin \beta s}$$

For line lengths that are an odd multiple of a quarter wavelength *i.e.*, $s = n\lambda/4$ or $\beta s = n\pi/2$, where n is an odd integer, we find

$$\sin \beta s = \pm 1,$$
$$\cos \beta s = 0.$$

so that

$$Z_S = Z_0 \frac{\cosh \alpha s}{\sinh \alpha s}.$$

For low loss lines αs is very small and we take $\cosh \alpha s \approx 1$ and $\sinh \alpha s = \sin \alpha s = \alpha s$. Thus

$$Z_S \approx Z_0 \cdot \frac{1}{\alpha s}. \qquad \ldots (A)$$

When $\omega L \gg R$ and $\omega C \gg G$, α is experessed as

$$\alpha = \tfrac{1}{2} \left\{ R \sqrt{\left(\frac{C}{L}\right)} + G \sqrt{\left(\frac{L}{C}\right)} \right\}$$

For lines having air as dielectric, losses due to the conductance G are negligible so that G can be neglected and

$$\alpha = \frac{R}{2} \sqrt{\left(\frac{C}{L}\right)} = \frac{R}{2Z_0}$$

Putting this value of α in equation (A), we get

$$Z_S = \frac{Z_0}{\alpha s} = \frac{2Z_0^2}{R_S}.$$

Therefore input impedance of a short circuited line, whose length is an odd multiple of a quarter wavelength, *i.e.*, $s = n\lambda/4$, is

$$Z_S = \frac{2Z_0^2}{R.n\lambda/4}.$$

Further substituting $\lambda = \frac{c}{f_0}$, where velocity of propagation of waves on the line is c, equal to the velocity of light, we get

$$Z_S = \frac{8Z_0^2 f_0}{Rnc},$$

The curve of resonant line impedance plotted against frequency in the vicinity of resonance resembles with resonance curve of paral-

lel resonant circuit and the sharpness is measured by an equivalent Q value given by

$$Q=\frac{2\pi Z_0 f_0}{R.c} \qquad ...(2)$$

Expression for Q can be derived as follows :

We have shown above that for short circuited line of length $\frac{n\lambda}{4}$, where n is an odd integer, input impedance is $Z_s=\frac{Z_0}{\alpha s}$.

When the frequency is shifted off resonance by a small amount δf, *i e.* when $f=f_0+\delta f$, then

$$\beta s=\frac{2\pi f}{c}.s=\frac{2\pi(f_0+\delta f)}{c}.s$$

$$=\frac{2\pi f_0 s}{c}+\frac{2\pi\cdot\delta f\cdot s}{c}$$

Since f_0 is resanant frequency for which $\beta s=n\pi/2$, we write

$$\beta s=\frac{2\pi s}{\lambda_0}+\frac{2\pi\ \delta f\cdot s}{c}$$

$$=(\beta s)_{res}+\frac{2\pi\cdot\delta f\cdot s}{c}$$

$$=\frac{n\pi}{2}+\frac{2\pi.\delta f.s}{c}$$

Under these conditions,

$$\cos\beta s=-\sin\left(\frac{2\pi.\delta f\cdot s}{c}\right)$$

$$\sin\beta s=\cos\left(\frac{2\pi.\delta f\cdot s}{c}\right).$$

Therefore input impedance is

$$Z_S=Z_0\frac{-\sinh\alpha s\sin\left(\frac{2\pi\delta f\cdot s}{c}\right)+j\cosh\alpha s.\cos\left(\frac{2\pi\delta f.s}{c}\right)}{-\cosh\alpha s\sin\left(\frac{2\pi\delta f\cdot s}{c}\right)+j\sinh\alpha s\cos\left(\frac{2\pi\delta f\cdot s}{c}\right)}.$$

For loss less lines, α is very small and hence $\sinh\alpha s\approx\alpha s$, $\cosh\alpha s\approx 1$; $\cos\left(\frac{2\pi\cdot\delta f\cdot s}{c}\right)\approx 1$ and $\sin\left(\frac{2\pi\delta f\cdot s}{e}\right)=\left(\frac{2\pi\delta f\cdot s}{c}\right)$

$$\text{Thus}\quad Z_S=Z_0\frac{-\alpha s.\frac{2\pi\delta f\cdot s}{c}+j\,1\cdot 1}{-1.\frac{2\pi.\delta f.s}{c}+j\alpha s.1}$$

$$=Z_0\frac{1+j.\alpha s.\frac{2\pi\delta f\cdot s}{c}}{\alpha s+j\frac{2\pi\delta f\cdot s}{c}}$$

For moderately high Q circuits, the second term in the numerator is the product of two small quantities and may be neglected in camparison with other terms, so that

$$Z_S = \frac{Z_0}{\alpha s + j.\frac{2\pi\delta f.s}{c}}$$

We know that for the parallel resonant circuit, when absolute magnitude of the impedance is 70·7 percent of the impedance at resonance $Z_0/\alpha s$, voltage across the parallel circuit is reduced to 70·7 percent of its value at resonance. This requires a frequency shift of $\Delta f/2$ on either side of resonant frequency f_0, where Δf is termed as bandwidth. Now, likewise, we observe from expression for Z_S that it will be $\frac{Z_0}{\alpha s\sqrt{2}}$ or 70·7 percent of its value for a resonant length† provided

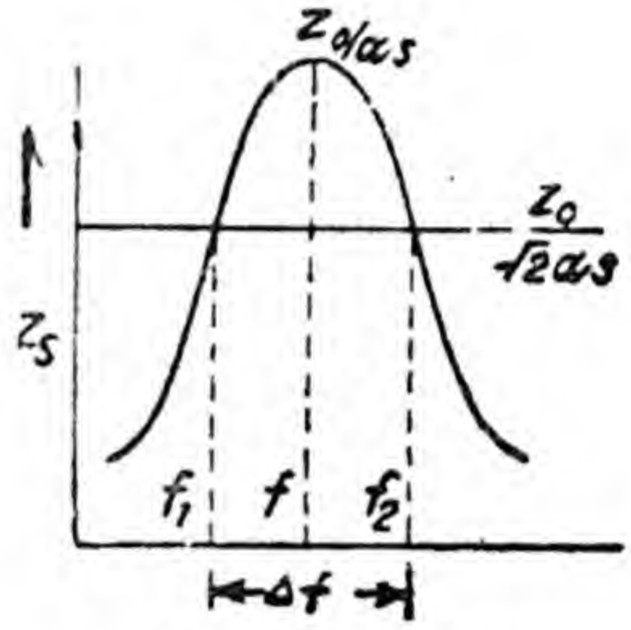


Fig. 27. (*c*)

$$\alpha s = \frac{2\pi\delta f\cdot s}{c}$$

but the frequency shift required to make this true will be $\Delta f/2$ and hence

$$\alpha s = \frac{2\pi\Delta f\cdot s}{2c}$$

or

$$\Delta f = \frac{\alpha c}{\pi} = \frac{2\alpha c}{2\pi\cdot}$$

The selectivity or ability of any resonant circuit to pass freely some frequencies while stopping all other is conveniently stated in terms of the ratio $\Delta f/f_0$, where f_0 is the resonant frequency and $\Delta f = f_2 - f_1$. the difference between half power frequencies. Therefore, Q of the resonant line section is

$$*Q = \frac{f_0}{\Delta f} = \frac{2\pi f_0}{2\alpha\, c}$$

---

† or resonant length

$$Z_S = \frac{Z_0}{\alpha s}$$

so that $\frac{Z_0}{\sqrt{2}.\alpha s} = 70{\cdot}7\%$ of its value at resonant length.

$$*Q = \frac{2\pi f_0}{2\alpha} = \frac{2\pi.\, f_0/c}{2x} = \frac{2\pi/\lambda}{2\alpha} = \frac{\beta}{2\alpha}$$

Putting, for lossless lines, $\alpha = \frac{R}{2Z_0}$, we get

$$Q = \frac{2\pi Z_0 f_0}{Rc}. \qquad \ldots(3)$$

Another important resemblance is that of current at resonance. In parallel resonant circuit current is minimum at resonant frequency and the same is true with a line short-circuited and $n\lambda/4$ long (n is odd number) or open circuited $n\lambda/4$ long (n is even number). From fig. 28 (a) and fig. 28 (b) respectively we see that in short circuited line, current is minimum at resonant frequencies $\beta s = \frac{n\pi}{2}$ $\left(\text{or } \frac{2\pi s}{\lambda} = \frac{n\pi}{2}, i.e \;\; s = \frac{n\lambda}{4}\right)$ where $n = 1,3,5,7\ldots\ldots$, and in circuited line, current is minimum at resonant frequencies $\beta s = n\pi/2$, where $n = 2, 4, 0$ etc.

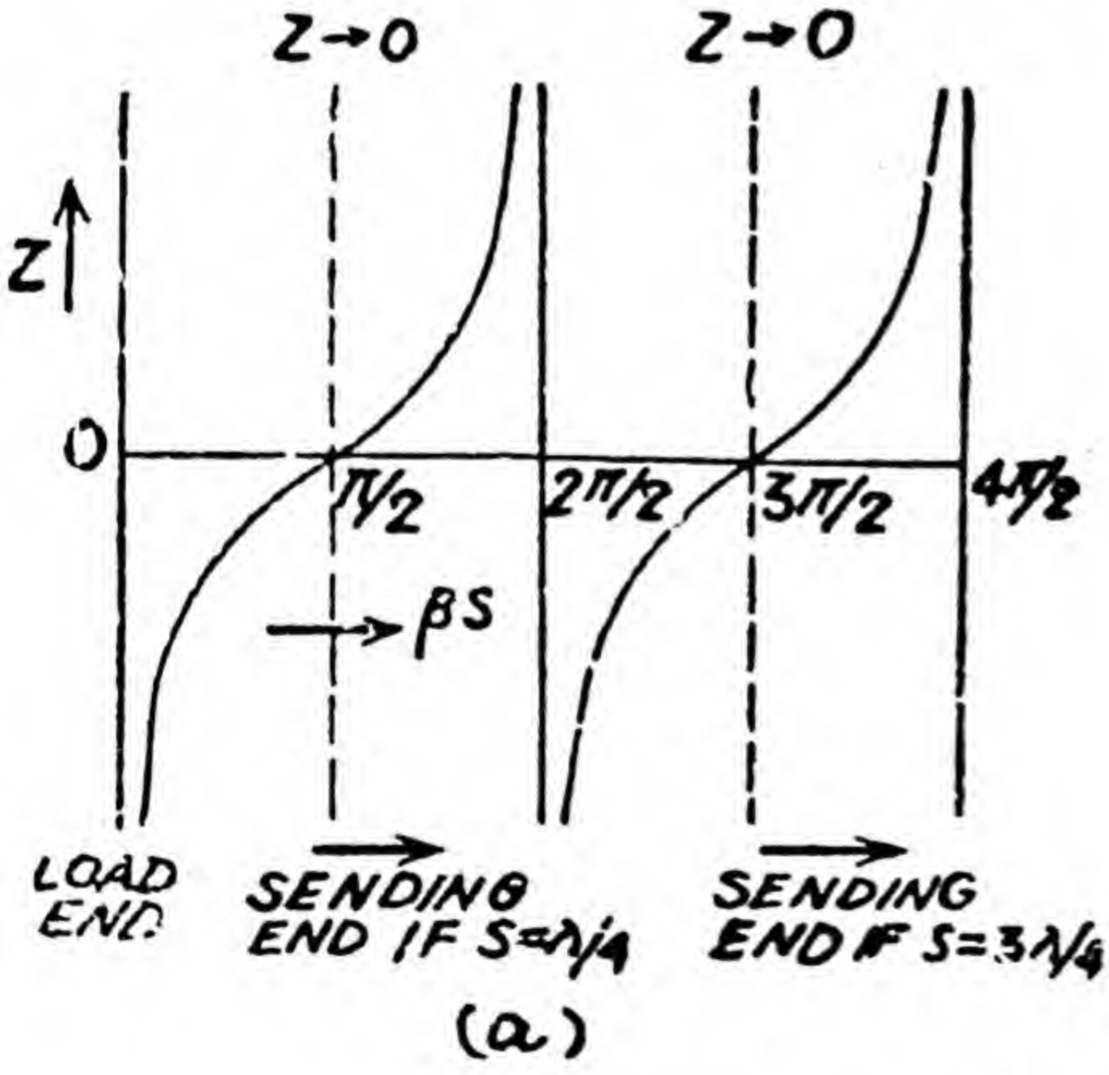


Fig 28. Plot of current for open and short-circuited lines.

Lines as series resonant circuits :

In series resonant circuits, impedance is minimum at the resonant frequency and current is maximum. The same behaviour is shown by a line *open-circuited at the receiver (load end) and odd multiple of quarter wavelength long.*

From figure 29 (a) it is obvious that impedance is minimum at frequencies corresponding to $\beta s = n\pi/2$ (or $s = n\lambda/4$) where n is



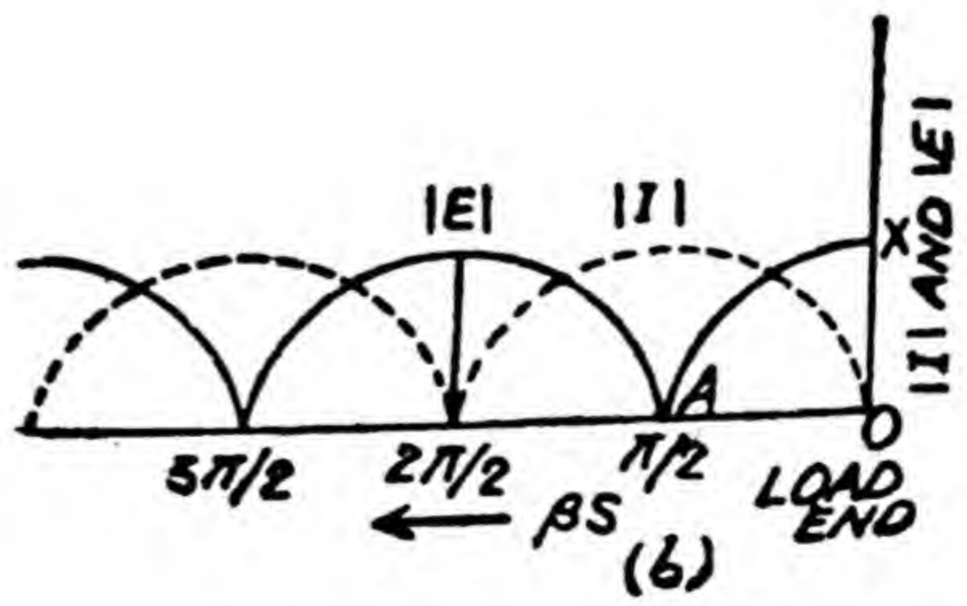


Fig. 29.

odd number. These are termed as resonant frequencies of resonant line.

Further from figure 29 (*b*), it is obvious that current is maximum at resonant frequencies $\beta s = n\pi/2$ ($n = 1, 3, 5, \ldots$).

Another important characteristic of such a line is that it acts as a voltage step-up device. Suppose line is $\lambda/4$ long, then sending end will be at *A*. It is obvious from figure 29 (*b*) that though voltage at *A* (generator) is zero yet at load end (receiver) voltage is of sufficient magnitude equal to *OX*. Thus a small voltage injected at the input will develop a sufficient voltage at the output *i.e.* an action like that of step-up device.

The further step up action can be explained as follows : We have assumed that line is open circuited at load end and $\lambda/4$ long, *i.e.* equal to *OA*. Since at *A* voltage is zero which on the other hand would mean that the line is short circuited at *A*. We can represent the line as shown in fig. 30 (*a*). The phase relation of voltage on such a line is shown in fig. 30 (*b*).

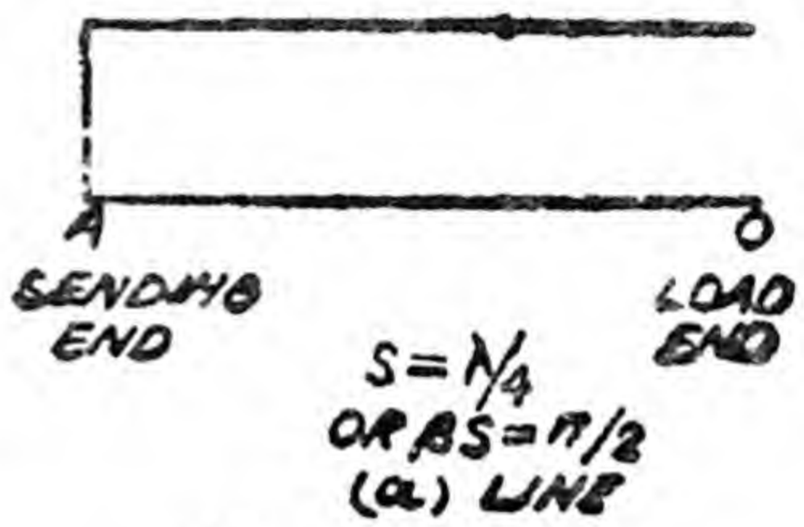


Fig. 30 (*a*).

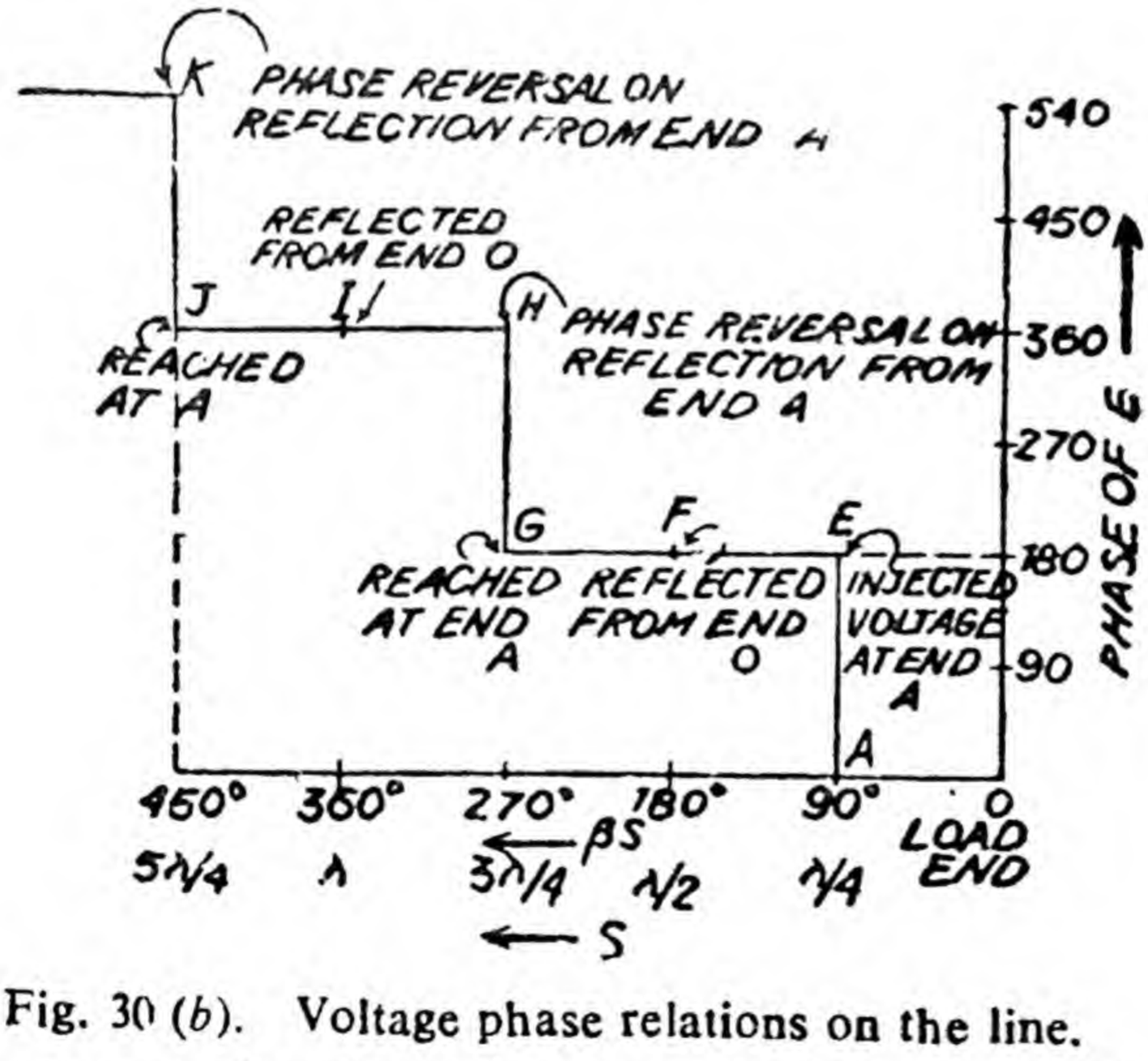


Fig. 30 (*b*). Voltage phase relations on the line.

Suppose a small voltage is injected at shorted end *A*. We also know that an incident wave is reflected from an open end with same magnitude and polarity. Thus at $\beta s = 90°$, we induce some voltage : it will travel to the end *O* ($\beta s = 180°$), and is reflected back without phase change to end *A* ($\beta s = 270°$). Total distance travelled is thus $\frac{1}{2}\lambda$. Now from end *A* ($\beta s = 270°$), it will be reflected back to *O* ($\beta s = 360°$) with a change of phase (since these points *G* and *H* are 180° out of phase). From open end *O* ($\beta s = 360°$) it will

be reflected to end A ($\beta s=450°$) in the same phase. Since the wave has travelled a total distance λ, the phase of this reflected wave at A ($\beta s=450°$) will be same as that of induced voltage (phase points K and E are at 450° and 180° and hence in same phase) and hence this reflected voltage is directly added to the incident voltage. This results in stepping up the voltage. This step up voltage after so many further reflections will again be step up. These additions continue to increase the voltage in the line until I^2R loss is equal to the power being put into the line. Voltage step up depends upon the Q value. The voltage step up ratio in the case of resonant line is $Q\times\frac{4}{\pi n}$. This expression can be derived as follows :

For quarter wave lossless line, we write

$$\frac{E_S}{E_R}=\frac{E_R\cosh j\beta s+I_RZ_0\sinh j\beta s}{E_R}$$

$$=\frac{E_R\cosh \pi/2+jI_RZ_0\sin \pi/2}{E_R}$$

$$=\frac{jI_RZ_0}{E_R}=\frac{jZ_0}{Z_R},$$

but for such a line, we know that the characteristic impedance Z_0 is the geometric mean of Z_R and Z_S, *i.e.*

$$Z_0=\sqrt{(Z_RZ_S)}.$$

Thus

$$\frac{E_S}{E_R}=j\sqrt{\left(\frac{Z_S}{Z_R}\right)}.$$

Therefore, voltage step up by the line is

$$\left|\frac{E_R}{E_S}\right|=\sqrt{\left(\frac{Z_R}{Z_S}\right)}$$

But for lossless quarter wave open circuited line Z_R is infinity and then above relation shows an infinite voltage step up. Such a conclusion is not surprising because, with $\alpha=0$, there is nothing to curb the step up. Therefore, a correct relation for voltage step up should include α-the attenuation factor. For correct relation in such a case, we write the equation for voltage E_S as

$$E_S=E_R\cosh(\alpha+j\beta)s+I_RZ_0\sinh(\alpha+j\beta)s$$

$$=E_R\cosh(\alpha s+j\pi/2)+I_RZ_0\sinh(\alpha s+j\pi/2)$$

$$=jE_R\sinh\alpha s+jI_RZ_0\cosh\alpha s.$$

In open circuit load current I_R is zero, so that

$$E_S=jE_R\sinh\alpha s.$$

Therefore, voltage step up

$$\left|\frac{E_R}{E_S}\right|=\frac{1}{\sinh\alpha s}\approx\frac{1}{\alpha s}$$ (as for low loss line α is very small and is $R/2Z_0$)

$$= \frac{2Z_0}{Rs}.$$

For resonant quarter wave section,

$$\beta s = n\pi/2, \text{ when } n \text{ is odd integer.}$$

$$\therefore \quad s = \frac{n\lambda}{4} = \frac{nc}{4f_0}.$$

Putting this value in above equation for voltage set up, we find

$$\left|\frac{E_R}{E_S}\right| = \frac{8Z_0 f_0}{\pi nc}$$

$$= \frac{2\pi Z_0 f_0}{Rc} \cdot \frac{4}{\pi n}$$

$$= Q \times \frac{4}{\pi n}. \qquad \ldots(4)$$

Example : What length of transmission line should be used and how should it be terminated for use as (a) tank circuit (b) series circuit at 500 mc/s.

(*a*) We have seen that for line to act as parallel circuit, we use a line short circuited at the load end of length $n\lambda/4$, where n is an odd integer ; or an open circuited line with even multiple of quarter wave long (*i.e.* n is even).

At 500 mc/s ,

$$\lambda = \frac{3.10^{10}}{500 \times 10^6} = 60 \text{ cm.}$$

Therefore for short-circuited line, length will be

$$1.\lambda/4 = \frac{60}{4} = 15 \text{ cm.}$$

or

$$3.\lambda/4 = \frac{3 \times 60}{4} = 45 \text{ cm.}$$

And for open circuited line, length will be

$$2.\lambda/4 = \frac{2 \times 60}{4} = 30 \text{ cm.}$$

or

$$4.\lambda/4 = \frac{4 \times 60}{4} = 60 \text{ cm.}$$

(*b*) For short circuited line to act as series resonant circuit it should have length equal to

$n\lambda/4$ where n is even

and for open circuited line to act as series resonant circuit it should have length equal to

$n\lambda/4$ where n is odd.

Therefore, length will be

30 cm.
or 60 cm. } for short circuited line.

15 cm
and 45 cm. } for open circuited line.

26·9. FIELD AND CIRCUIT CONCEPT CORRELATION :

(E and H about the parallel plane conductors)

Analysis of transmission can be had both in terms of voltage and current or electric and magnetic fields using Maxwell's field equations. The present discussion aims at establishing a correlation between the field analysis and voltage analysis. In this article, E would mean the electric field and not the voltage. For voltage, symbol V will be used.

Suppose there is a transmission line on which the waves are guided along z direction.

As shown in fig. 31 (a), line consists of two conducting plates with a separation a in X direction and b in Y direction. Line in Z direction is of infinite extent. There exists a current I in each plate flowing in Z direction. Linear surface density of such a current in Z direction will be

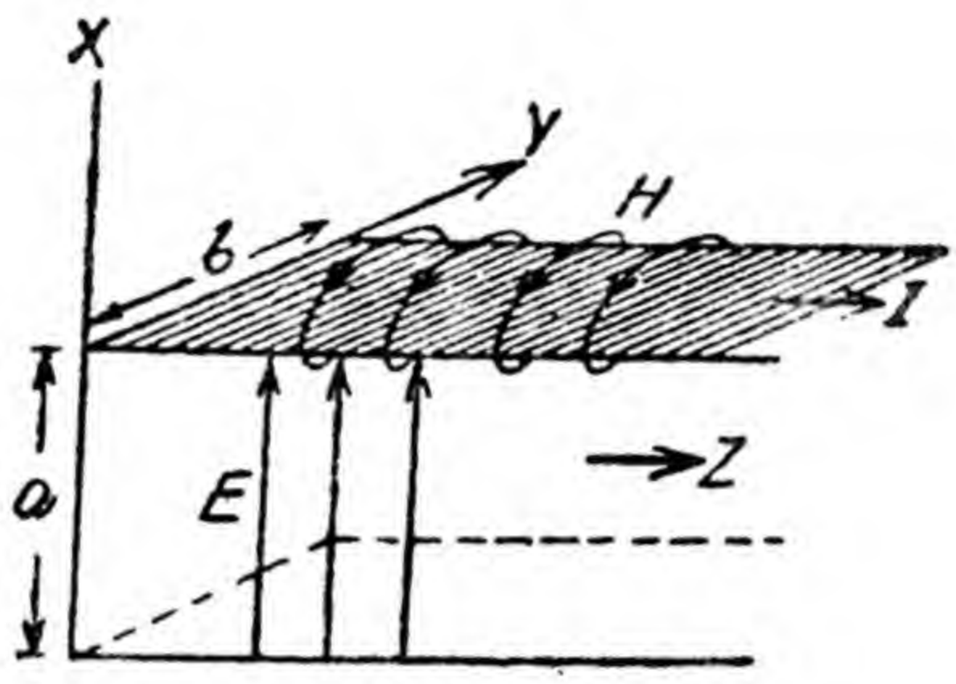


Fig. 31 (a).

$$J_{sz}=\frac{I}{b}. \qquad ...(1)$$

The magnetic field H will then surround the each plate forming closed loops. If the conductor has zero resistance, current will flow in the surface of the conductor and the magnetic field will not penetrate the conductor. Similarly, as there exists a potential difference between the two plates (since an alternating potential is applied at the input terminals ; if the voltage at any point in upper plate is +2 volts, the voltage at the corresponding point in the lower plate will −2 volts), an electric field will also exist purely along X-axis. If the conductor offers zero resistance, the electric field E will not develop any voltage difference between any two points on the same conductor because ($E\,R\,dz$) is zero, *i.e.* voltage will be shunted. Therefore, *we conclude that if the conductor is perfect,* **E** *and* **H** *will not penetrate the conductor and the losses due to conductor will be almost negligible.*

Let us consider a *TEM* wave travelling along the line stretched in Z direction so that

$$\mathbf{E}=\hat{\mathbf{x}}E_x \text{ and } \mathbf{H}=\hat{\mathbf{y}}H_y,$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ denote unit vectors along X and Y directions.

Now we shall establish relation between fields, and voltage and current. Consider the side view of the line, *i.e* , XZ plane [fig. 31 (b)].

Applying e.m.f. equation to the closed path *ABCDA*.

$$*\oint \mathbf{E}\cdot d\mathbf{l} = -j\omega \int_s \mathbf{B}\cdot d\mathbf{A}$$

or $$V_{AB}+V_{BC}+V_{CD}+V_{DA} = -j\omega B_n \, d\mathbf{A},$$

because field exists only along y direction therefore only $B_y=\mu H_y$ component of the magnetic flux density will exist. $\mu=\mu_r\mu_v$, where μ_r is the permeability of material, μ_v that of space between the plates but because we take that the fields do not penetrate the materials, μ would mean the permeability of space in between the plates.

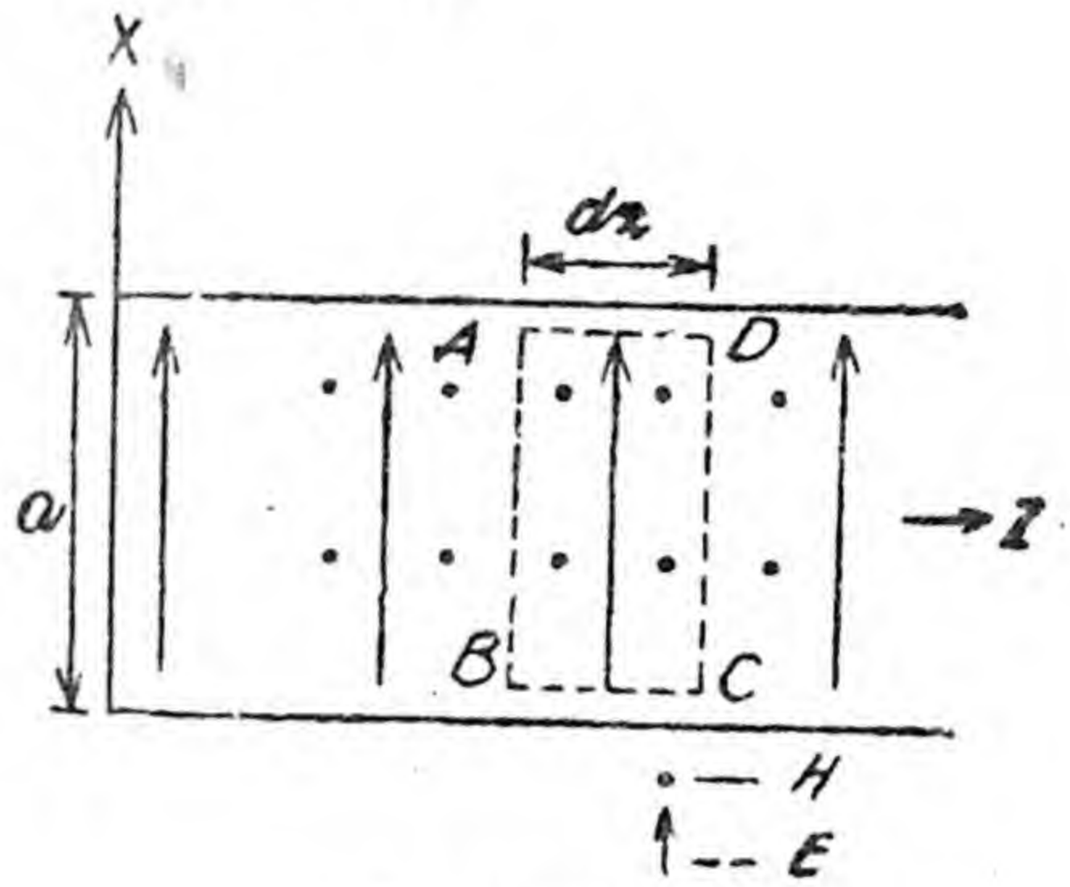


Fig. 31. (*b*)

Further, $$dA = AB.BC = a\,dz.$$

Also $$V_{AD}=V_{BC}=0,$$

since they are perpendicular to the direction of E. We take V_{AB} as negative and V_{CD} (work done in the direction of E) as positive.

Therefore, $$V_{CD}-V_{AB} = -j\omega\, B_y\, a\, dz$$

or $$dV = -j\omega\, B_y\, a\, dz,$$

$$\frac{dV}{dz} = -j\omega\, B_y\, a. \qquad \ldots(2)$$

Further, linear surface density in any of the plates will be

$$\mathbf{J}_s = \hat{\mathbf{x}}\times\mathbf{H}$$

$$= \hat{\mathbf{x}}\times\hat{\mathbf{y}}H_y \text{ (since } \mathbf{H}=\hat{\mathbf{y}}H_y : H_z \text{ and } H_x \text{ are zero)}$$

$$= \hat{\mathbf{z}}H_y$$

or $$\mathbf{J}_{sz} = H_y = \frac{I}{b} \text{ (from eq. 1)}$$

so that $$B_y = \mu H_y = \frac{\mu I}{b}.$$

---

*Maxwell's equation is

$$\nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t}.$$

Applying Stoke's theorem, we get its integral form. This is

$$\oint \mathbf{E}\cdot d\mathbf{l} = -\int \frac{\partial\mathbf{B}}{\partial t}.d\mathbf{A}.$$

But when the magnetic field is sinusoidal,

$$\mathbf{B} = \mathbf{B}_0\, e^{j\omega t}$$

so that $$\frac{\partial\mathbf{B}}{\partial t} = j\omega\, \mathbf{B}$$

Putting in equation (2), we get

$$\frac{dV}{dz}=-j\omega\mu\,\frac{Ia}{b} \qquad \ldots(3)$$

While dealing with the analysis of transmission line in voltage and current (art. 26·1), we have shown that

$$dV=-(R+j\omega L)\,dz.I,$$

minus sign has been imposed to indicate that dV represents fall in voltage on the line as the wave travels a distance dz.

If $R=0$, *then*

$$dV=-j\omega L\,dz.I. \qquad \ldots(4)$$

Comparing equations (3) and (4), we find that inductance per unit length of the line is

$$\mathbf{L}=\frac{\mathbf{a}\mu}{\mathbf{b}} \qquad \ldots(5)$$

Now we proceed to write mmf. equation. For this we shall consider the top view YZ planeof fig. 31 (*a*) shown in fig. 31 (*c*) Applying mmf. equation to the loop *FGHK*, *i.e.*

$$^{*}\oint \mathbf{H}\cdot d\mathbf{l}=\int (j\omega\epsilon\mathbf{E}+\mathbf{J})\cdot d\mathbf{A}.$$

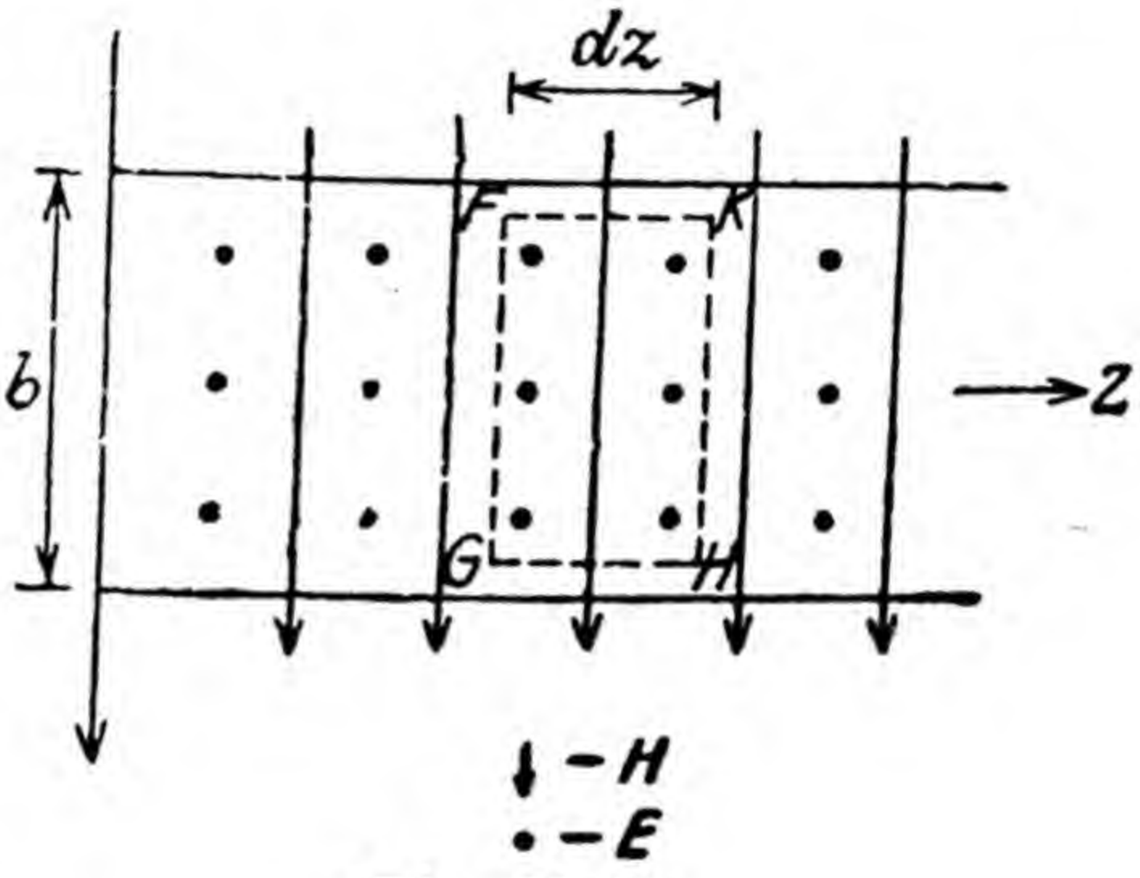


Fig. 31 (*c*)

ϵ refers to the permeability of the space in between the plates,

or $$(H_{FG}+H_{GH}+H_{HK}+H_{KF})\cdot d\mathbf{l}=j\omega\epsilon E_x\cdot b\,dz,$$

because $E=\hat{\mathbf{x}}E_x$ (E_y and E_z are zero for *TEM* wave)

$dA=b\,dz$

---

*Maxwell's equation is

$$\nabla\times\mathbf{H}=\frac{\partial\mathbf{D}}{\partial t}+\mathbf{J}.$$

Applying Stoke's theorem, equation is converted into integral form, that is,

$$\oint \mathbf{H}\cdot d\mathbf{l}=\int\left(\frac{\partial\mathbf{D}}{\partial t}+\mathbf{J}\right)\cdot d\mathbf{A}.$$

For sinusoidal electrostatic field

$$\mathbf{D}=\mathbf{D}_0\,e^{j\omega t}$$

so that

$$\frac{\partial\mathbf{D}}{\partial t}=j\omega\mathbf{D}$$

$$=j\omega\epsilon\mathbf{E}$$

and $\mathbf{J}.d\mathbf{A}=0$ as J is in YZ plane whereas dA points normal to this plane.

Further $\mathbf{H}_{GH}.d\mathbf{l}=\mathbf{H}_{KF}.d\mathbf{l}=0$ as path FK and GH are at right angles to field direction H_y.

and $$\mathbf{H}_{FG}.d\mathbf{l}+\mathbf{H}_{HK}\cdot d\mathbf{l}=(H_{FG}-H_{HK})\,b$$
$$=-b\,dH_y,$$

the negative sign is imposed because FG direction points in negative y direction.

Therefore mmf equation becomes

$$-b\,dH_y=-j\omega\epsilon E_x\,b\;dz$$

or $$\frac{d\,(bH_y)}{dz}=-j\omega\epsilon E_x b.$$

But $$bJ_y=bJ_{sz}=I,$$

so that $$\frac{dI}{dz}=-j\omega\epsilon E_x b. \qquad \ldots(6)$$

Again we refer to the current equation established while dealing with wave propagation on the line (art. 26·1) which is,

$$dI=-(G+j\omega C)\,dz.V,$$

minus sign indicates the fall in current as the wave travels a distance dz on the line.

With $G=0$, for the perfect dielectric

$$\frac{dI}{dz}=-j\omega CV. \qquad \ldots(7)$$

Comparing (6) and (7), we find that the capacitance per unit length of the line is

$$C=\frac{\epsilon E_x b}{V}.$$

But $V=E_x\,a$, so that

$$\mathbf{C}=\epsilon\,\frac{\mathbf{b}}{\mathbf{a}}. \qquad \ldots(8)$$

Phase and group velocities of the wave on the line :

The phase velocity of the wave on the *dissipationless* line ($R=G=0$) is :

$$v=\frac{\omega}{\beta}=\frac{\omega}{\omega\sqrt{(LC)}}=\frac{1}{\sqrt{(LC)}}\cdot$$
$$=\frac{1}{\sqrt{\left(\frac{\mu a}{b}\cdot\epsilon\frac{b}{a}\right)}}=\frac{1}{\sqrt{(\mu\epsilon)}}=c^*. \qquad \ldots(9)$$

---

*In case of perfect conductor, as stated previously, μ and ϵ refer to the space between the conductors. If there is free space then

$$v=\frac{1}{\sqrt{(\mu_0\epsilon_0)}}=3\cdot10^8 \text{ meters/sec.},$$

equal to the velocity of light.

Relation (9) shows that velocity of propagation of the wave on the line is independent of the geometry of the line (whether parallel wire, coaxial, etc.).

The group velocity, v_g, is given by

$$v_g = \frac{d\omega}{d\beta}$$

Since $\beta = \omega\sqrt{(LC)}$

$$d\beta = d\omega\ \sqrt{(LC)}$$

or $$\frac{d\omega}{d\beta} = \frac{1}{\sqrt{(LC)}} = \frac{1}{\sqrt{(LC)}} = c$$

Thus

$$v_g = c$$

we note that in case of dissipation less line ($R=G=0$), phase and group velocities are equal and

$$v_p\ v_g = c^2$$

Characteristic impedance :

If the line has unit width and unit separation, so that $a=b=1$, then $L=\mu$, and $C=\epsilon$

which implies that ϵ is the capacitance between conductors of 1 meter length of the parallel plate line, which is meter wide and has a separation of 1 meter. Similarly, μ is the inductance per meter length of the same line.

$$Z_0 = \sqrt{\left(\frac{L}{C}\right)}$$

$$= \sqrt{\left(\frac{\mu}{\epsilon}\right)} \cdot \frac{a}{b}$$

For the line having unit separation and unit width of each of the plates, the characteristic impedance is just the intrinsic impedance of the dielectric medium between the plates.

Maximum velocity : When R and G are negligible, the phase and group velocities are given by

$$v_p = v_g = \frac{1}{\sqrt{(LC)'}} \qquad \ldots(10)$$

where L is in henrys per unit length and C is in farads per unit length.

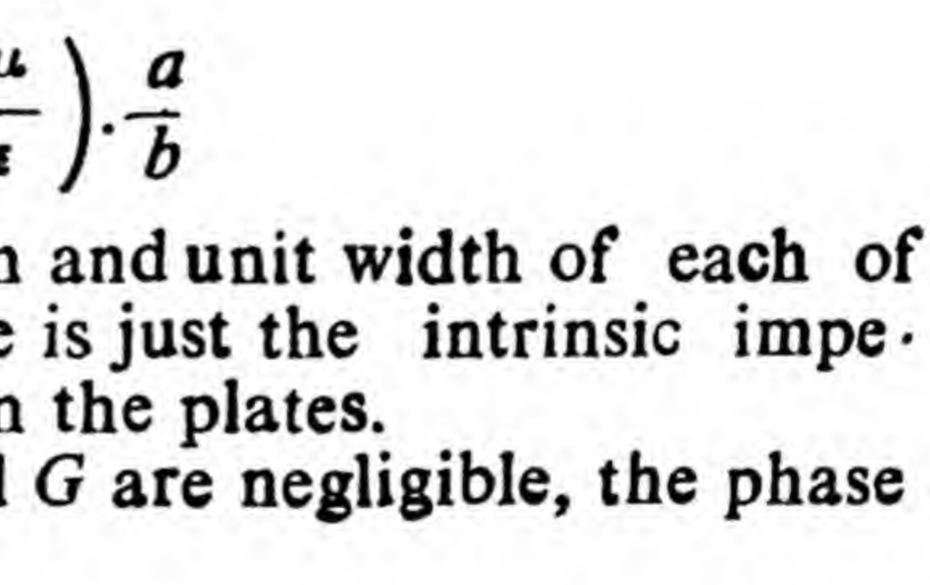
---

*To prove that in a lossless line, the voltage at any point is in magnitude equal to the product of surge impedance and current at a point quarter of wavelength away. The voltage and current relations as referred to load end are

$$E = E_R \cosh \gamma l + I_R Z_0 \sinh \gamma l,$$

$$I = I_R \cosh \gamma l + E_R/Z_0 \sinh \gamma l,$$

where $\gamma = j\beta l$.



(Continued on next page)

The inductance in millihenry per mile of line, when separation between its conductors is large compared with the radius of the wire, is

$$L=1\cdot481 \log \frac{s}{r}+L', \qquad \ldots(11)$$

where r is the radius of wires and s is the separation between them. L' is the inductance due to the linkage of flux and current within the conductor. At high frequencies when the current flows mainly in the surface of wires, this factor approaches zero.

Further, the capacitance between two wires for $s \gg r$, *i.e.*, separation being much greater than the radius of wires, in micro-faraps per mile of line is

$$C=\frac{0\cdot01\ 41}{\log (s/r)}. \qquad \ldots(12)$$

Therefore, putting values of L and C in henrys/mile and farads/mile, we get from eqs. (11) and (12), that

$$LC=(1\cdot481\times10^{-3})\ (0\cdot01941\times10^{-6})+\frac{0\cdot01941\times10^{-6}L'}{\log (s/r)} \qquad (13)$$

When the separation between the wires or frequency is increased, the second term becomes negligible as compared to first. The first term is a constant and represents the minimum possible value which LC can assume. Putting first term in eq. (10), we get maximum value of phase and group velocity, that is

$$(v_p)_{max.}=(v_g)_{max.}=\frac{1}{\sqrt{(1\cdot481\times0\ 01941\times10^{-9})}}$$
$$=186{,}300 \text{ miles/sec.}$$

Now voltage at BB ($x=l$) can be written in the form

$$E_l=E_R\cos\beta l+jI_RZ_0\sin\beta l$$

or $$|E_l|e^{i\psi}=E_R\cos\beta l+jI_RZ_0\sin\beta l,$$

where ψ represents the phase of E_l.

Further $$|E_l|\cos\psi=E_R\cos\beta l,$$
$$|E_l|\sin\psi=I_RZ_0\sin\beta l.$$
$$\therefore\ |E_l|=\sqrt{(E_R^2\cos^2\beta l+I_R^2Z_0^2\sin^2\beta l)}. \qquad \ldots(1)$$

Now current at AA ($x=l+\frac{1}{4}\lambda$),

$$I_{l+1/4}=I_R\cos\beta\ (l+\tfrac{1}{4}\lambda)+jE_R/Z_0\sin\beta\ (l+\tfrac{1}{4}\lambda)$$
$$=I_R\cos\left(\frac{\pi}{2}+\beta l\right)+j\ \frac{E_R}{Z_0}\sin\left(\frac{\pi}{2}+\beta l\right)$$
$$=-I_R\sin\beta l+jE_R/Z_0\cos\beta l.$$

So $$|I_{l+\lambda/4}|=\frac{\sqrt{(Z_0^2I_R^2\sin^2\beta l+E_R^2\cos^2\beta l)}}{Z_0}$$
$$=|E_l|/Z_0 \text{ from } 1).$$

Hence $$|E_1|=Z_0\cdot|I_{l+\lambda/4}|.$$

Thus maximum velocity of waves on such a transmission line is equal to the velocity of light. However, resistance R, conductance G and term L' reduce this value, but to a very small extent.

NUMERICALS

1. *A pair of high freque cy parallel wir transmission lines has distributed capacitance an1 inductance of ·04μF and 9·8 mH. per mile respectively. What is the surge impedance of the line ?*

$$Z_0=\sqrt{\left(\frac{L}{C}\right)}=\sqrt{\left(\frac{9\cdot8\times10^{-8}}{0\cdot04\times10^{-6}}\right)}=\sqrt{\left(\frac{9\,8\times10^{5}}{4}\right)}$$

$$=\frac{990}{2}$$

$$=\mathbf{495\ ohms.}$$

2. *The primary constants of an open wire transmission line at 1·2 kc/s. are $R=11$ ohms/mile, $L=4\cdot5$ mh./mile, $G=0\cdot6$ micro mho per mile and $C=0\cdot00825$ microfarad/mile. Calculate the characteristic constants of the line.*

$$\omega L=2\pi\times1\cdot2.10^{3}\times4\cdot5.10^{-3}$$
$$=33\cdot9,$$
$$\omega C=2\pi\times1\cdot2.10^{3}\times8\cdot25.10^{-9}$$
$$=62\cdot25.10^{-6}.$$

Now $^{*}Z=R+j\omega L=11+j.33\cdot9=35\cdot6\ \angle(72\cdot1)$

and $Y=G+j\omega C=10^{-6}\ (0\cdot6+j\ 62\cdot25)$
$=62\cdot25.10^{-6}\ \angle\ (90-\cdot55)$
$=62\cdot25.10^{-6}\ \angle\ (89\cdot45).$

Therefore,

characteristic impedance $Z_0=\sqrt{\left(\frac{Z}{Y}\right)}$

$$=\sqrt{\left(\frac{35\cdot6}{62\cdot25}\times10^{6}\right)}.\angle\ (-8\cdot7)$$

$$=\mathbf{755\ \angle\ (-8\ 7)\ ohms.}$$

propagation constant $\gamma=\sqrt{(ZY)}$

$$=\sqrt{(35\cdot6\times62\cdot25)}\times10^{-3}\ \angle\ (80\cdot8)$$
$$=4\ \cdot2\times10^{-3}\ \angle\ (80\cdot8)$$
$$=\mathbf{0\cdot00755+j.0\ 0466.}$$

Since $\gamma=\alpha+j\beta.$

$$\alpha=0\ 00755\ \text{neper/mile},$$
$$\beta=0\cdot0466\ \text{rad/mile}.$$

Also $\lambda=\frac{2\pi}{\beta}=135$ miles, $v=f\lambda=1,62,000$ miles/sec.

---

*Since Z will have amplitude and phase both, we can represent it as

$$Z=A\ \angle\ \theta=Ae^{j\theta}=A\cos\theta+j\ A\sin\theta=R+j\omega L.$$
$$\therefore\ A\cos\theta=R$$
$$A\sin\theta=\omega L$$

Hence amplitude of Z, *i.e.* $A=\sqrt{(R^2+\omega L^2)}$ and phase $\theta=\tan^{-1}\omega L/R.$

3. *The primary constants of a cable are $R=80$ ohms per mile, $L=2$mh/mile, $G=0{\cdot}3\times10^{-6}$ mho/mile $C=0\,07$ microfarad/mile. Calculate the characteristic constants of the line at 500 c/s.*

$$Z=R+j\omega L=80+j.6{\cdot}28$$
$$=80{\cdot}3\ \angle(4{\cdot}5),$$
$$Y=G+j\omega C=10^{-6}\ (0{\cdot}3+j.220)$$
$$=220{\cdot}10^{-6},\ \angle\ (90).$$

$\therefore$ Characteristic impedance $Z_0=\sqrt{\left(\frac{Z}{Y}\right)}$

$$*Z_0=\sqrt{\left(\frac{80{\cdot}3}{2\text{C}0}\times10^{6}\right)}/\underline{-42{\cdot}2}$$
$$=\mathbf{605}\ /\underline{\mathbf{42{\cdot}8}}$$

Propagation constant $\gamma=\sqrt{(ZY)}$

$$=(80{\cdot}3\times220\times10^{-6})^{1/2}\ /\underline{47{\cdot}3}\cdot$$
$$=133.10^{-3}\ /\underline{47{\cdot}3}$$
$$=\mathbf{0{\cdot}0903+j.0{\cdot}0977}$$
$$=\alpha+j\beta.$$

Hence $\alpha=\mathbf{0{\cdot}0903\ neper/mile},$

$\beta=\mathbf{0{\cdot}0976\ rad/mile}$

$$\lambda=\frac{2\pi}{\beta}=\frac{2\times3{\cdot}14}{0{\cdot}0977}=\mathbf{64{\cdot}4\ miles}$$
$$v=f\lambda=500\times64{\cdot}4$$
$$=\mathbf{32,2000\ mile/sec.}$$

4. *At* 1200 *cycles the characteristic impedance and propagation constants of an open wire transmission line are*

$$Z_0=650-j\cdot150,$$
$$\gamma=0{\cdot}\text{l}\,05+j\,0{\cdot}007.$$

Calculate the distributed parameters of the line.

The characteristic impedance $Z_0=\left(\frac{Z}{Y}\right)^{1/2}$

and propagation constant $\gamma=\sqrt{(ZY)}.$

Therefore $Z_0.\gamma=Z=R+j\omega L$

and $\frac{\gamma}{Z_0}=Y=G+j\omega C.$

Now $R+j\omega L=Z_0.\gamma$

$$=(650-j.1\text{5}0)(0{\cdot}005+j0{\cdot}007)$$
$$=666\ /\underline{-13}\times8{\cdot}5\times10^{-3}\ /\underline{45}$$
$$=0{\cdot}666\times8\ 5\ /\underline{32}$$
$$=4{\cdot}8+j\ 3{\cdot}0.$$

$$*Z_0=\sqrt{\left(\frac{Z}{Y}\right)}=\sqrt{\left(\frac{80\ 3.e^{j\cdot4\ 5}}{220.10^{-6}.e^{j\cdot90}}\right)}=\sqrt{\left(\frac{80{\cdot}3}{2\text{2}0.10^{-6}}\right)}\ e^{-j(85\cdot5)/2}$$
$$=\sqrt{\left(\frac{80{\cdot}3}{220}.10^{6}\right)}\ e^{-j\cdot42\cdot8}=605\ /\underline{-42{\cdot}8}$$

Equating real and imaginary parts,

$$\mathbf{R=4{\cdot}8\ ohms/mile.}$$

$$\omega L=3.$$

$$\therefore \quad L=\frac{3}{2\pi.1200}=\mathbf{0{\cdot}398\ mH./mile.}$$

Further, $G+j\omega C=\dfrac{\gamma}{Z_0}=\dfrac{8{\cdot}6.10^{-3}/\underline{45}}{666/\underline{-13}}$

$$=(6{\cdot}75+j.10{\cdot}8)\ 10^{-6}.$$

Hence $\mathbf{G=6{\cdot}75\ micro\ mhos/mile}$

and $$G=\frac{10{\cdot}8.10^{-6}}{2\pi\times 1200}=\mathbf{0{\cdot}00143\ microfarad/mile.}$$

5. *The short-circuited and open circuited impedances of a transmission line of length* 100 *miles are* $(Z_s)_{short}=700\ /\underline{30}$ *and* $(Z_s)_{open}=500\ /\underline{-40}$. *Find the characteristic constant of the line.*

We have seen in theory that

$$(Z_s)_{Z_{R=0}}=Z_0 \tanh \gamma l=(Z_s)_{short}$$

and $$(Z_s)_{Z_{R=\infty}}=Z_0 \coth \gamma l=(Z_s)_{open}.$$

$$\therefore \quad \frac{(Z_s)_{short}}{(Z_s)_{open}}=\tanh^2 \gamma l.$$

Further, $\tanh \gamma l=\dfrac{e^{2\gamma l}-1}{e^{2\gamma l}+1}$,

$$\therefore \quad e^{2\gamma l}=\frac{1+\tanh \gamma l}{1-\tanh \gamma l},$$

$$e^{2\gamma l}=e^{2(\alpha+j\beta)l}=e^{2\alpha l}./\underline{2\beta l}.$$

Now from problem,

$$\tanh \gamma l=\sqrt{\left\{\frac{(Z_s)_{short}}{(Z_s)_{open}}\right\}}=\sqrt{\left(\frac{700\ /\underline{30}}{500\ /\underline{-40}}\right)}$$

$$=1{\cdot}18\ /\underline{35}$$

$$=0{\cdot}968+j.0{\cdot}678$$

Therefore, $$e^{2\gamma l}=\frac{1+0{\cdot}968+j.0{\cdot}678}{1-0{\cdot}968-j\,0\,678}$$

$$=\frac{1{\cdot}968+j.0{\cdot}678}{0{\cdot}032-j.0{\cdot}678}$$

$$=3{\cdot}07\ /\underline{106.4}$$

This is $e^{2\alpha l}./\underline{2\beta l}=3{\cdot}07\ /\underline{106{\cdot}4}$

or $$e^{2\alpha l}=3{\cdot}07$$

or $$2\alpha l=\log_e 3{\cdot}07=2{\cdot}3026\times{\cdot}4871=1{\cdot}20$$

or $$\alpha=\frac{1{\cdot}20}{2\times 100}=\mathbf{0{\cdot}0056\ neper/mile\ :}$$

Also $2\beta l = 106\ 4 \text{ deg.} = 105{\cdot}4 \times \frac{\pi}{100}$ rad.

$$\therefore\ \beta = \frac{106{\cdot}4 \times 3{\cdot}14}{2 \times 100 \times 184} = \mathbf{0{\cdot}00928\ rad/mile.}$$

Characteristic impedance

$$Z_0 = \sqrt{\{(Z_s)_{short}(Z_s)_{open}\}}$$
$$= \sqrt{(600\ \underline{/30} \times 500\ \underline{/-40})}$$
$$= \mathbf{592\ \underline{/-5}}.$$

6. *Consider a two wire line with following constants :*

$l = 100$ *miles.*
$f = 1000$ *c/s.,*
$Z_0 = (685 - j.92)$ *ohms.*
$\gamma = (0{\cdot}000497 + j0{\cdot}0352)$ *per mile,*
$Z_R = (200 + j0)$ *ohms,*
$E_g = 10$ *volts (generator voltage at the sending end).*
$Z_g = (700 + j.0)$ *ohms (generator impedance).*

Find sending end impedance Z_S, sending end current I_S and the receiving end voltage E_R, using hyperbolic form of the line equations.

Sending end impedance is expressed as

$$Z_s = \frac{Z_0\,[Z_R \cosh \gamma l - Z_0 \sinh l\gamma]}{Z_0 \cosh \gamma l + Z_R \sinh \gamma l}$$
$$= \frac{Z_0\,[Z_R + Z_0 \tanh \gamma l]}{Z_0 + Z_R \tanh \gamma l}. \qquad \ldots(1)$$

We note that

$$\coth \gamma l = \frac{\cosh \gamma l}{\sinh \gamma l}$$
$$= \frac{\cosh \alpha l \cos \beta l + j \sinh \alpha l \sin \beta l}{\sinh \alpha l \cos \beta l + j \cosh \alpha l \sin \beta l}$$
$$= \frac{\cosh 0{\cdot}497 \cos 3{\cdot}52 + j \sinh 0{\cdot}497 \sin 3{\cdot}52}{\sinh 0{\cdot}497 \cos 3{\cdot}52 + j \cosh 0{\cdot}497 \sin 3{\cdot}52}$$
$$= \frac{-1{\cdot}126 \times 0{\cdot}927 - j\,0{\cdot}518 \times 0{\cdot}375}{-0{\cdot}518 \times 0{\cdot}927 - j.1{\cdot}126 \times 0{\cdot}375}$$
$$= \frac{-1{\cdot}042 - j.0{\cdot}194}{-0{\cdot}480 - j\,0{\cdot}422}$$

or $$\tanh \gamma l = \frac{0{\cdot}480 + j.0{\cdot}422}{1{\cdot}042 + j.0{\cdot}194}$$
$$= 0{\cdot}523 + j.0{\cdot}311.$$

Therefore from (1),

$$Z_s = (685 - j.92)\left[\frac{2000 + (685 - j.92)\ (0{\cdot}532 + j.0{\cdot}311)}{(685 - j.92) + (2000)\ (0{\cdot}532 + j.0{\cdot}311)}\right]$$
$$= \mathbf{(859 - j.325)\ ohms.}$$

Sending end current will be

$$|I_s| = \left|\frac{E_g}{E_g + Z_s}\right| \qquad \ldots(2)$$

$$= \left|\frac{10}{(700 + 859 - j\,325)}\right|$$

$$= \mathbf{6{\cdot}26 \times 10^{-3}\ amp.}$$

Receiving end voltage is

$$|E_R| = |I_R|\,|Z_R|, \qquad \ldots(3)$$

for which we have to calculate first $|I_R|$ given by

$$\frac{|I_s|}{|I_R|} = \left|\frac{I_R \cosh \gamma l + E_R/Z_0 \sinh \gamma l}{I_R}\right|$$

$$= \left|\cosh \gamma l + \frac{Z_R}{Z_0} \sinh \gamma l\right|$$

$$\therefore \quad |I_R| = \frac{|I_s|}{\left|\cosh \gamma l + \dfrac{Z_R}{Z_0} \sinh \gamma l\right|}$$

$$= \frac{6{\cdot}26 \times 10^{-3}}{\left|(-1{\cdot}042 - j.0{\cdot}194) + \left(\dfrac{2000}{685 - j.92}\right)(-0{\cdot}480 - j.0{\cdot}422)\right|}$$

$$= \frac{6{\cdot}26 \times 10^{-3}}{2{\cdot}76}$$

$$= 2\ 27 \times 10^{-3}\ \text{amp.}$$

From equation (3), we get

$$|E_R| = 2{\cdot}27 \times 10^{-3} \times 2000 = \mathbf{4{\cdot}54\ volts}.$$

7. *A parallel wire line has*

$$Z_0 = 70\ \textit{ohms}.$$
$$Z_R = (115 - j.80)\ \textit{ohms}.$$
$$S = 2{\cdot}7$$

If the line is to transmit a power of 50 watts, find the magnitude of maximum and minimum voltage and current, also the magnitude of the receiving end voltage.

The maximum and minimum impedances are purely resistive and therefore, we can write for the maximum point that

$$P = \frac{|E_{max.}|^2}{Z_{max.}} = \frac{|E_{max.}|^2}{SZ_0}$$

or
$$50 = \frac{|E_{max.}|^2}{2{\cdot}7 \times 70}$$

from which

$$|E_{max.}| = \mathbf{97{\cdot}3\ volts.}$$

Further,
$$|E_{min.}| = \frac{|E_{max}|}{S}$$

$$= \frac{97{\cdot}3}{2{\cdot}7}$$

$$= \mathbf{36{\cdot}0\ volts.}$$

The minimum current is

$$|I_{min.}| = \frac{|E_{min.}|}{Z_0}$$

$$= \frac{36{\cdot}0}{70}$$

$$= \mathbf{0{\cdot}514\ amp.}$$

The maximum current is

$$|I_{max.}| = \frac{|E_{max.}|}{Z_0}$$

$$= \frac{97{\cdot}3}{70}$$

$$= \mathbf{1{\cdot}39\ amp.}$$

The receiving end current can be found from the relation

$$I_R = \sqrt{\left(\frac{P}{R_R}\right)},$$

where R_R is the resistive part of Z_R.

$$|I_R| = \sqrt{\left(\frac{50}{115}\right)}$$

$$= 0{\cdot}660\ \text{amp.}$$

The receiving end voltage will be

$$|E_R| = |I_R|\,|Z_R|$$

$$= 0{\cdot}660 \times \sqrt{\{(115)^2 + (80)^2\}}$$

$$= \mathbf{92{\cdot}5\ volts}.$$

8. *Calculate the air insulated line composed of two parallel copper tubes* 0·15 *inches radius and spaced* 1 *inch between centres. The line is short circuited at the one end and is driven with* 200 *mc/s. generator at the other. Neglecting leakage conductance and radiation, the attenuation constant is found to be*

$$\alpha = 6{\cdot}78 \times 10^{-4}\ \textit{neper/meter}.$$

Show that resonance is sharp.

To show that resonance is sharp, we shall calculate bandwidth Δf between half points given by

$$\Delta f = \frac{f}{Q}.$$

But

$$Q = \frac{\beta}{2\alpha},$$

so that

$$\Delta f = \frac{2\alpha f}{\beta}$$

$$= \frac{2\alpha f v}{\omega},$$

on putting

$$\beta = \frac{\omega}{v}.$$

Assuming a velocity equal to that of light, we find

$$\Delta f = \frac{2\alpha f v}{2\pi f}$$

$$= \frac{\alpha v}{\pi}$$

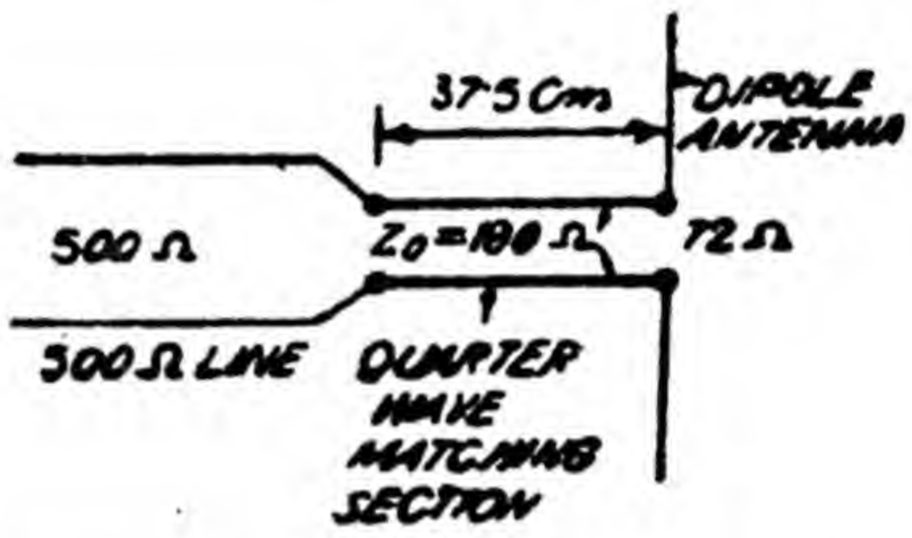
$$= \frac{6{\cdot}78 \times 10^{-4} \times 3 \times 10^{8}}{3{\cdot}14} = \frac{20{\cdot}34 \times 10^{4}}{3{\cdot}14}$$

$$= 64700 \text{ c/s.} = 0{\cdot}6 \text{ mc/s.},$$

which means, for a line operating at 200 mc/s, resonance is extremely sharp.

9. *Show by a diagram and calculations how a quarter wave section can be used to match a 500 ohm line to a 72 ohm dipole antenna at 200 mc/s.*

The diagram shows a quarter wave matching section between a 500 ohm line and 72 ohm dipole antenna ($\lambda/2$ antenna).



We know that a quarter wave line matches the two impedances —input impedance Z_s ($=500\Omega$) and output impedance Z_L ($=72\ \Omega$) provided its characteristic impedance Z_0 is the geometric mean of the two. That is

$$Z_0 = \sqrt{(Z_s\, Z_L)}$$

$$= \sqrt{(500 \times 72)} = 190\ \Omega$$

Length ($\lambda/4$) of this line will be equal to

$$= \frac{3.10^{10}}{4 \times 200.10^{6}} = \frac{150}{4}$$

$$= 37{\cdot}5 \text{ cm.}$$

Therefore, we should use a line of length 37·5 cm. of characteristic impedance 190 ohms.

EXERCISES AND PROBLEMS

1. Derive an expression for the input impedance of a dissipationless transmission line terminated by any impedance and hence deduce the above expression :
   (*i*) when it is short-circuited,
   (*ii*) when it is open-circuited.
2. What do you understand by surge impedance of a transmission line ? Show the effect of terminating such lines with their surge impedance.
3. Show that a quarter wave line matches two impedances connected to its extremities, if the characteristic impedance of the line is equal to the geometric mean of the impedances.

4. Discuss the propagation of high frequency currents along a transmission line and obtain expression for (*a*) attenuation constant (*b*) phase constant.
5. Explain (*a*) attenuation constant, (*b*) wavelength constant, and (*c*) surge impedance of a transmission line. Obtain expression for them interms of inductance, capacitance, resistance and leakance for unit length of the line. What will be their values for negligible resistance and leakance ? Describe the main types of high frequency feeder lines.
6. A low loss transmission line of characteristic impedance 85 ohms is terminated by (*a*) short-circuited load (*b*) open circuited load (*c*) a resistive load of 50 ohms and (*d*) a resistive load of 100 ohms. Calculate the positions of maxima and minima in each case and draw figures to show them.
7. Show that there exists similarity between the relationship of plane-wave propagation of electromagnetic waves and those for waves on transmission line.
8. Derive a formula that gives the standing wave ratio at the receiving end of the transmission line in terms of the reflection coefficient of the load.
9. Explain what is meant by standing wave ratio. Prove that resistive load of R_1 and R_2 produce the same standing wave ratio, provided that $R_1 R_2 = Z_0^2$ where Z_0 is the surge impedance of the transmission line.
10. Explain how transmission line stubs can be used for eliminating harmonics generated from transmitters.
11. Calculate the characteristic impedance of a uniform transmission line having $L = 0{\cdot}5\ mH/km$ and $C = 0{\cdot}08\ \mu F/km$.
12. In a transmission line having a characteristic impedance of 500 ohms, the load impedance is a resistance of 100 ohms shunted by capacitance of 100 μF. Calculate the magnitude and base of the reflection coefficient at 200 kc/s.
13. What do you understand by the characteristic impedance of a transmission line. Discuss the important properties and applications of a half wave and quarter wave line.
14. Describe the use of a transmission line for the impedance matching. How will you measure the frequency of a microwave oscillator with the help of a transmission line.
15. What do you understand by Q of a transmission line. A length of loss-less transmission line is short circuited at one end and open circuited at the other end. The lowest frequency at which the line exhibits resonance is 10 mc/s. When 5 *pf* capacitance is connected across the previously open circuited end of the line the resonant frequency is 9·5 mc/s. What is the characteristic impedance of the line ?
16. Define the terms 'reflection coefficient' and 'voltage standing wave ratio' and find the relation between the two.
17. Define propagation constant, attenuation constant and phase constants. Show that the input impedance of a transmission line terminated by

Z_R is given by

$$Z_{in}=Z_0\left[\frac{Z_R\cosh\gamma l+Z_0\sinh\gamma l}{Z_0\cosh\gamma l+Z_R\sinh\gamma l}\right]$$

18. Derive the equation of propagation of travelling wave in transmission lines and obtain the solution interms of the characteristic impedance and propagation constant. Discuss the interpretation of the transmission line equations of travelling waves and obtain the expression of the reflection coefficient.
19. Write an essay on (i) transmission line analogy for waveguide (ii) stub line matching.
20. Establish a relation between the transmission line parameters and the attenuation and phase constants. Discuss the performance of a perfect transmission line for which $R=G=0$.
    What happens when the line is not perfect ? Derive a condition for distortionless propagation on an imperfect line.
21. Discuss the formation of standing waves on a loss less *UHF* transmission line and derive relationship between the load impedance and the nature and position of standing wave. Describe how dielectric constant of materials at *UHF* can be measured by this method.
22. Explain the term 'phase velocity' and 'group velocity'. Obtain expressions for these velocities in respect of a transmission line and show that $v_p\, v_g=c^2$.
23. Comment briefly but critically on quarter wave transformer.

27 WAVEGUIDES

Till now we have discussed that a two wire transmission line provides a means of guiding radio frequency power from one point to another. It should be noted, however, that two conductors are not always necessary for the transference of radio frequency power between two points but only a single conductor, *e.g.* a hollow metal pipe, is sufficient if the frequency is high enough. Similar to a coaxial line in which an electromagnetic wave, travelling along it, is confined to the space between the conductors, in these metal pipes or tubes too, wave is confined to move in the direction of the tubes. Such a wave is called a guided wave and the single conductor as *waveguide*. They are employed at and above 3000 mc/s.

In waveguide, voltage and current are distributed in a very complicated manner. Therefore, like to transmission line, it is not convenient for these guides to be analysed interms of ordinary electric circuit quantities of voltage, current and impedances. For their analysis, it is convenient to consider the electric and magnetic fields within the guide. These fields are perpendicular to each other and together form an e.m. wave which travels through the waveguide. As the wave moves in the guide, infinite variety of pattern are possible. These patterns are termed as *modes*. In the articles on waveguide, we shall use the letter **E** to denote the *electric field*, instead of voltage.

27·1. BASIC CONCEPTS OF GUIDED WAVES :

As a preliminary to studying propagation through waveguide, we shall investigate the problem of propagation between two parallel perfectly conducting planes of infinite extent in y and z directions fig. 1 and separated by a distance 'a'.

All electromagnetic fields, in the space between conductors, must obey certain physical laws. We have mentioned earlier that electric field is always perpendi-

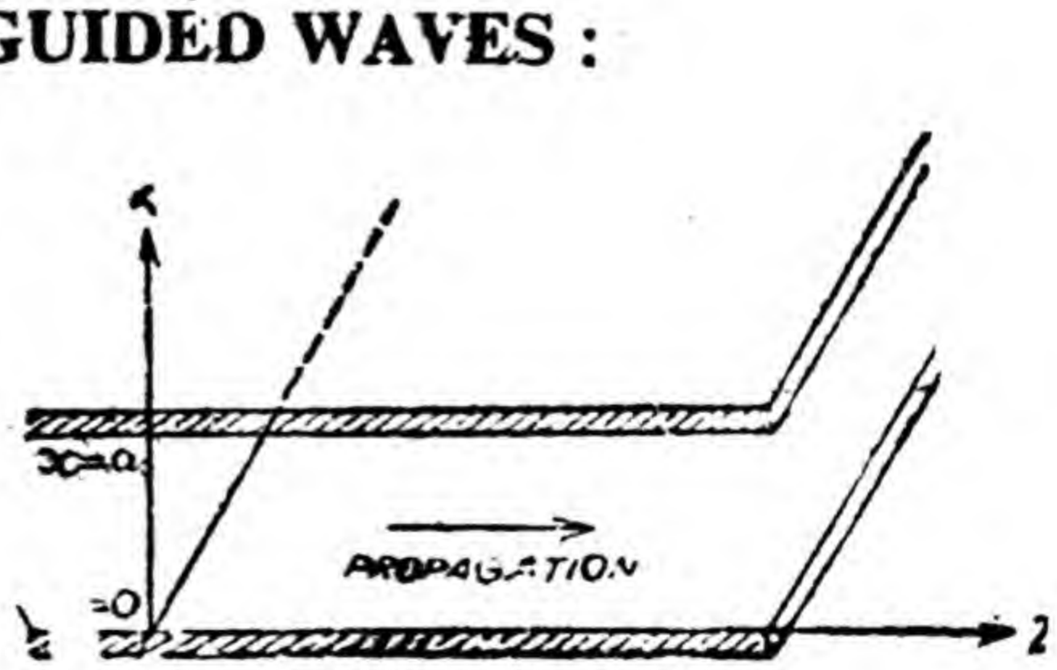


Fig 1 Two parallel conducting planes.

cular to the magnetic field. Now at the surface of a conductor, any component of electric field parallel to the surface must be zero which means the *electric field must always be perpendicular to the surface of a conductor.* The magnetic field, being perpendicular to electric field, must now be parallel to the surface of a conductor. Thus

$$E_{tangential}=0$$
$$H_{normal}=0$$

Maxwell's equation subject to these boundary conditions at the surface of conductor, can determine the configuration of electromagnetic field in the space between the planes.

In rectangular co-ordinates and for non-conducting region between planes. Maxwell's equations are†

$$\nabla\times\mathbf{H}=j\omega\epsilon\mathbf{E} \qquad \nabla\times\mathbf{E}=-j\omega\mu\mathbf{H}.$$

Writing the above equations in terms of x, y and z components, we have :

$$\left.\begin{aligned}
\frac{\partial H_z}{\partial y}-\frac{\partial H_y}{\partial z}&=j\omega\epsilon E_x, & \frac{\partial E_z}{\partial y}-\frac{\partial E_y}{\partial z}&=-j\omega\mu H_x\\
\frac{\partial H_x}{\partial z}-\frac{\partial H_z}{\partial x}&=j\omega\epsilon E_y, & \frac{\partial E_x}{\partial z}-\frac{\partial E_z}{\partial x}&=-j\omega\mu H_y\\
\frac{\partial H_y}{\partial x}-\frac{\partial H_x}{\partial y}&=j\omega\epsilon E_z, & \frac{\partial E_y}{\partial x}-\frac{\partial E_x}{\partial y}&=-j\omega\mu H_z
\end{aligned}\right\} \quad ...(1)$$

Also $*\nabla^2\mathbf{E}=-\omega^2\mu\epsilon\mathbf{E}$

$$\nabla^2\mathbf{H}=-\omega^2\mu\epsilon\mathbf{H}.$$

or

$$\left.\begin{aligned}
\frac{\partial^2\mathbf{E}}{\partial x^2}+\frac{\partial^2\mathbf{E}}{\partial y}+\frac{\partial^2\mathbf{E}}{\partial z}&=-\omega^2\mu\epsilon\mathbf{E}\\
\frac{\partial^2\mathbf{H}}{\partial x^2}+\frac{\partial^2\mathbf{H}}{\partial y^2}+\frac{\partial^2\mathbf{H}}{\partial z^2}&=-\omega^2\mu\epsilon\mathbf{H}
\end{aligned}\right\} \quad ...(2)$$

where ϵ is the permittivity and μ is the permeability of the dielectric medium between the plates, ω is the angular frequency of the wave.

$\dagger\vec{\nabla}\times\mathbf{H}=\left(\frac{\partial\mathbf{D}}{\partial t}+\mathbf{J}\right)$ $\qquad \vec{\nabla}\times\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t}$

with $\mathbf{J}=0$, $\qquad =-j\omega\mathbf{B}=-j\omega\mu\mathbf{H}$

we get $\vec{\nabla}\times\mathbf{H}=\frac{\partial\mathbf{D}}{\partial t}=j\omega\epsilon\mathbf{E}$ for sinusoidal varying fields.

Because for sinusoidal varying field in isotropic media $\mathbf{D}=\mathbf{D}_m e^{j\omega t}$,

$*\nabla\times(\nabla\times\mathbf{E})=\nabla(j\omega\mu\mathbf{H})$

$\nabla(\nabla.\mathbf{E})-\nabla^2\mathbf{E}=-j\omega\mu(\nabla\times\mathbf{H})$

$-\nabla^2\mathbf{E}=-j\omega\mu(j\omega\epsilon\mathbf{E})=\omega^2\mu\epsilon\mathbf{E}$

Assume that the wave is propagating in z direction and since the field components will vary in this direction, we can further suppose that variation of all field components in this direction is expressed in the form $e^{-\gamma z}$, where γ is a complex constant given by

$$\gamma = \alpha + j\beta$$

where α and β are called attenuation and phase constants. When we include time variation factor to this z variation factor, then $e^{j\omega t}\, e^{-\gamma z}$, *i e.*, $e^{(j\omega t - \gamma z)}$ represents a wave propagating in the z-direction.

While considering x direction, we have to impose certain boundary conditions stated previously but for y-direction, however, there are no boundary conditions and space between the planes in this direction is of infinite extent. We, therefore, can assume that field in y-direction is uniform or constant and hence all the derivatives with respect to y in equation (1) should be zero. Furthermore, representing the field component H_y, varying in z-direction by $H_y = H_y{}^0 e^{-\gamma z}$, we can easily find

$$\frac{\partial H_y}{\partial z} = -\gamma H_y{}^0 e^{-\gamma z} = -\gamma H_y,$$

Similarly, for other z-derivatives, expressions can be written.

On putting $\frac{\partial}{\partial y} = 0$ and for $\frac{\partial}{\partial z} = -\gamma$, equations (1) and (2) become

$$\left.\begin{aligned} \gamma H_y &= j\omega\epsilon E_x, & \gamma E_y &= -j\omega\mu H_x. \\ -\gamma H_x - \frac{\partial H_z}{\partial x} &= j\omega\epsilon E_y, & -\gamma E_x - \frac{\partial E_z}{\partial x} &= -j\omega\mu H_y. \\ \frac{\partial H_y}{\partial x} &= j\omega\epsilon E_z, & \frac{\partial E_y}{\partial x} &= -j\omega\mu H_z, \end{aligned}\right\} \quad \ldots(2)$$

$$\left.\begin{aligned} \frac{\partial^2 \mathbf{E}}{\partial x^2} + \gamma^2 \mathbf{E} &= -\omega^2\mu\epsilon\mathbf{E} \\ \frac{\partial^2 \mathbf{H}}{\partial x^2} + \gamma^2 \mathbf{H} &= -\omega^2\mu\epsilon\mathbf{H} \end{aligned}\right\} \quad \ldots(4)$$

Equations (3) can be solved to give the following equations :

$$\left.\begin{aligned} H_x &= -\frac{\gamma}{h^2}\frac{\partial H_z}{\partial x}, & E_x &= -\frac{\gamma}{h^2}\frac{\partial E_z}{\partial x} \\ H_y &= -\frac{j\omega\epsilon}{h^2}\frac{\partial E_z}{\partial x}, & \dagger E_y &= +\frac{j\omega\mu}{h^2}\frac{\partial H_z}{\partial x} \end{aligned}\right\} \quad \ldots(5)$$

where $h^2 = \gamma^2 + \omega^2\mu\epsilon$.

†From equation (3), $H_x = -\frac{\gamma}{j\omega\mu} E_y$. Putting this values of H_x in

$$j\omega\epsilon E_y = -\gamma H_x - \frac{\partial H_z}{\partial x} = +\frac{\gamma^2}{j\omega\mu} E_y - \frac{\partial H_z}{\partial x},$$

so that

$$-\frac{(\gamma^2 + \omega^2\mu\epsilon)\, E_y}{j\omega\mu} = -\frac{\partial H_z}{\partial x} \text{ or } E_y = \frac{j\omega\mu}{h^2}\frac{\partial H_z}{\partial x}.$$

Similarly, other equations can be obtained.

We note from equation (5) that E_x, E_y, H_x and H_y depends on E_z and H_z, the components of electric and magnetic field in the direction of propagation. If, however, we assume that there is no component of both the fields in the direction of propagation*, there would be no fields at all in the space considered. Therefore either E_z must be present or H_z, or both. We, however, leaving general case in which E_z and H_z both may be present, divide the solutions of above equations in two sets, first when E_z is present and not H_z, *i e.* there is component of **E** in the direction of propagation but no component of **H**. Such waves are called *transverse magnetic waves* (*TM* mode) since **H** is purely transverse : secondly when H_z is present and E_z is zero. Such waves are termed as *transverse electric waves* (*TE* mode).

27·2. GUIDED WAVES BETWEEN PARALLEL PLANES :

In this article we shall discuss *TE* and *TM* waves guided along the two parallel planes, the basis of which has already been discussed in the previous article

(A) TRANSVERSE ELECTRIC WAVES : ($E_z=0$) : From eq. (4) of the previous article, the equation of motion of electric vector **E** is given by

$$\frac{\partial^2 \mathbf{E}}{\partial x^2} + \gamma^2 \mathbf{E} = -\omega^2 \mu \epsilon E$$

or
$$\frac{\partial^2 \mathbf{E}}{\partial x^2} + h^2 \mathbf{E} = 0$$

where
$$h^2 = \gamma^2 + \omega^2 \mu \epsilon.$$

Writing the above wave equation for the component E_y and recalling that $E_y = E_y^0 e^{-\gamma z}$ we arrive at

$$\frac{\partial^2 E_y^0}{\partial x^2} + h^2 E_y^0 = 0$$

Solution of above equation is

$$E_y^0 = C_1 \sin hx + C_2 \cos hx,$$

where C_1 and C_2 are arbitrary constants, and will be determined by the application of the boundary condition which is

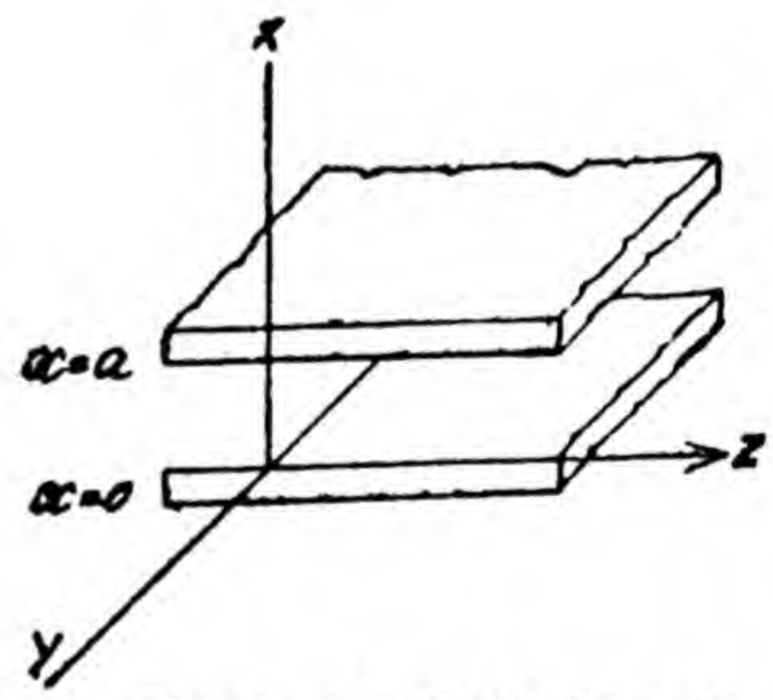


Fig. 2. Two parallel planes; z is the direction of propagation.

$$E_{tangential} = 0,$$

implying that

$$\left.\begin{aligned} E_y &= 0 \text{ at } x = 0 \\ E_y &= 0 \text{ at } x = a \end{aligned}\right\} \text{ for all values of } z.$$

---

*We are not to mention here *TEM* mode with one possibility (TM_{00}) for such a case, of course.

From first of these we get $C_2=0$ so that solution is reduced to

$$E_y^0=C_1 \sin hx.$$

From second boundary condition, we get

$$0=C_1 \sin ha$$

giving $$ha=m\pi$$

or $$h=\frac{m\pi}{a},\ m=\pm 1,\ \pm 2,\ldots$$

The solution is thus

$$E_y^0=C_1 \sin\left(\frac{m\pi x}{a}\right) \qquad \ldots(1)$$

The other field components can be calculated on applying equation (5) of previous article. We have

$$E_x^0=-\frac{\gamma}{h^2}\frac{\partial E_z^0}{\partial x}=0$$

$$E_y^0=C_1 \sin\left(\frac{m\pi x}{a}\right)$$

$$H_z^0=\frac{h^2}{j\omega\mu}\int E_y^0\, dx$$

$$=-\frac{h^2 a}{j\omega\mu . m\pi} C_1 \cos\left(\frac{m\pi x}{a}\right)$$

$$=-\frac{m\pi}{j\omega\mu a} C_1 \cos\left(\frac{m\pi x}{a}\right)$$

$$H_z^0=-\frac{\gamma}{h^2}\frac{\partial H_z^0}{\partial x}$$

$$=-\frac{\gamma}{j\omega\mu} C_1 \sin\left(\frac{m\pi x}{a}\right)$$

$$H_y^0=-\frac{j\omega\epsilon}{h^2}\frac{\partial E_z}{\partial x}=0.$$

Therefore field components of TE waves between parallel planes are

$$\left.\begin{aligned} E_y&=E_y^0\, e^{-\gamma z}\\ &=C_1 \sin\left(\frac{m\pi x}{a}\right) e^{-\gamma z}\\ H_z&=-\frac{m\pi}{j\omega\mu a} C_1 \cos\left(\frac{m\pi x}{a}\right) e^{-\gamma z}\\ H_x&=-\frac{\gamma}{j\omega\mu} C_1 \sin\left(\frac{m\pi x}{a}\right) e^{-\gamma z}. \end{aligned}\right\} \qquad \ldots(2)$$

Each value of m refers to a particular field configuration or mode. It denotes the number of half period variations of electric

field along the x direction (wide dimension) of the parallel planes. A particular mode is written like $TE_{m,0}$ in which subscript 0 refers to y direction and will have some value in waveguide where planes in y direction will also be considered.

The propagation constant is given by

$$\gamma^2 = h^2 - \omega^2\mu\epsilon$$

$$= \sqrt{\left[\left(\frac{m\pi}{a}\right)^2 - \omega^2\mu\epsilon\right]} \qquad \ldots(3)$$

We know that if γ is purely real then $\beta=0$ which implies that there is no phase change along the line *i.e.* no propagation. Therefore propagation will be possible when

$$\left(\frac{m\pi}{a}\right)^2 < \omega^2\mu\epsilon. \qquad \ldots(4)$$

The frequency above which the propagation is possible is given by

$$\omega_0^2\mu\epsilon = \left(\frac{m\pi}{a}\right)^2$$

or
$$\omega_c = \frac{v_0\, m\pi}{a}$$

or
$$f_c = \frac{v_0}{2}\cdot\frac{m}{a}$$

or
$$\lambda_c = \frac{v_0}{f_0} = \frac{2a}{m}. \qquad \ldots(5)$$

That means, higher is the separation between the planes, lower is the cut-off frequency. It may be noted that propagation is possible when

$$\omega_0^2\mu\epsilon < \omega^2\mu\epsilon$$

or
$$\omega > \omega_0$$

or
$$\lambda < \lambda_0 \quad \text{or} \quad \lambda_0 > \lambda$$

i.e., when cut off wavelengths is longer than the free space wavelength. With $\gamma = j\beta$, field equations become

$$E_y = C_1 \sin\left(\frac{m\pi x}{a}\right) e^{-j\beta z}$$

$$H_z = -\frac{m\pi}{j\omega\mu a} C_1 \cos\left(\frac{m\pi x}{a}\right) e^{-j\beta z} \qquad \ldots(6)$$

$$H_x = -\frac{\beta}{\omega\mu} C_1 \sin\left(\frac{m\pi x}{a}\right) e^{-j\beta z}$$

Since $m=0$ reduces the whole field to zero, the lowest value of m is 1. Therefore lowest order mode that can exist in this case is $TE_{1,0}$ mode.

Phase and Group Velocities :

The propagation constant is given by

$$\gamma=\sqrt{(\omega_c^2\mu\epsilon-\omega^2\mu\epsilon)}$$

or
$$\beta=\sqrt{(\omega^2\mu\epsilon-\omega_c^2\mu\epsilon)}.$$

Phase velocity with which the wave travels within the guide is

$$(v_g)_{ph}=\frac{\omega}{\beta}=\frac{\omega}{\sqrt{(\omega^2\mu\epsilon-\omega_c^2\mu\epsilon)}}$$

$$=\frac{1}{\sqrt{(\mu\epsilon)}}\,\frac{1}{\sqrt{(1-\omega_c^2/\omega^2)}}$$

$$=\frac{v_0}{\sqrt{(1-\lambda_0^2/\lambda_c^2)}} \qquad ...(7)$$

The group velocity with which the energy travels within the guide is

$$(v_g)_{gr.}=\frac{C^2}{(v_g)_{ph.}}=\frac{v_0^2}{(v_g)_{ph.}}$$

$$=\frac{(v_g)^2_{ph.}\sqrt{(1-\lambda_0^2/\lambda_c^2)}}{(v_g)_{ph.}}$$

$$=(v_g)_{ph.}\sqrt{(1-\lambda_0^2/\lambda_c^2)} \qquad ...(8)$$

Further wavelength of the wave within the guide is given by

$$\lambda_g=\frac{(v_g)_{ph.}}{f}=\frac{v_0}{f}\,\frac{1}{\sqrt{(1-\lambda_0^2/\lambda_c^2)}}=\frac{\lambda_0}{\sqrt{(1-\lambda_0^2/\lambda_c^2)}} \qquad ..(9)$$

or
$$\sqrt{\left(1-\frac{\lambda_0^2}{\lambda_c^2}\right)}=\frac{\lambda_0}{\lambda_g}$$

or
$$1-\frac{\lambda_0^2}{\lambda_c^2}=\frac{\lambda_0^2}{\lambda_g^2}$$

or
$$\frac{1}{\lambda_0^2}=\frac{1}{\lambda_c^2}+\frac{1}{\lambda_g^2} \qquad ...(10)$$

Ex. 1. *A wave is propagated in a parallel plane guide operating in TE mode at a frequency of 6Gc/s. The separation between planes is 5 cm. Find*

(*i*) *The cut off wavelength for the dominant mode.*

(*ii*) *Wavelength with the guide operating in dominant mode.*

(*iii*) *Corresponding group and phase velccities.*

(**a**) For dominant mode, $m=1$ and therefore from eq. (5), the corresponding cut off wavelength will be

$$\lambda_0=\frac{2a}{m}=2a=2\times 5=10 \text{ cm}.$$

(b) From eq. (9), the wavelength within the guide for dominant mode ($\lambda_0 = 10$ cm.) will be

$$\lambda_g = \frac{\lambda_0}{\sqrt{(1-\lambda_0^2/\lambda_c^2)}}$$

where

$$\lambda_0 = \frac{v}{f} = \frac{c}{f} = \frac{3\times10^{10}}{6\times10^9} = 5 \text{ cm.}$$

Therefore

$$\lambda_g = \frac{5}{\sqrt{\left\{1-\left(\frac{5}{10}\right)^2\right\}}} = \frac{5}{\sqrt{(1-0{\cdot}25)}} = \frac{5}{\sqrt{(0{\cdot}75)}}$$

$$= \frac{5}{0{\cdot}866} = 5{\cdot}77 \text{ cm.}$$

Note that guide wavelength is greater than free space wavelength. Further as cut off wavelength is much greater than free space wavelength propagation is possible.

(c) From eq. (7), phase velocity is

$$(v_g)_{ph.} = \frac{v_0}{\sqrt{(1-\lambda_0{}^2/\lambda_c{}^2)}} = \frac{3\times10^{10}}{0{\cdot}866} = 3{\cdot}46\times10^8 \ m/s.$$

and group velocity is

$$(v_g)_{gr.} = (v_g)_{ph.}\sqrt{(1-\lambda_0{}^2/\lambda_c{}^2)}$$
$$= 3{\cdot}46\times10^8\times0{\cdot}866$$
$$= 2{\cdot}99\times10^8 \ m/s.$$

Ex. 2. *Find the greatest number of half waves of electric intensity with which it may be possible to propagate a signal of* 10Gc/s *in a waveguide whose wall separation is* 5 *cm. Calculate the guide wavelength for this mode of propagation.*

In this problem we are to find the largest value of m for which the cut-off wavelength, λ_c, is larger than free space wavelength, λ_0, $-a$ condition for the wave to propagation down the waveguide.

The free space wavelength is

$$\lambda_0 = \frac{c}{f} = \frac{3\times10^{10}}{10\times10^9} = 3 \text{ cm.}$$

Cut-off wavelength

for $m=1$, $\quad \lambda_c = \frac{2a}{m} = \frac{2\times5}{1} = 10$ cm.

(this mode will propagate)

for $m=2$, $\quad \lambda_0 = \frac{2a}{m} = \frac{2\times5}{2} = 5$ cm.

(this mode will propagate)

for $m=3$, $\lambda_c=\frac{2a}{m}=\frac{2\times5}{3}=3{\cdot}33$ cm.

(this mode will propagate)

for $m=4$, $\lambda_c=\frac{2a}{m}=\frac{2\times5}{4}=2{\cdot}5$ cm.

The last mode will not propagate because for this $\lambda_c<\lambda_0$. Thus greatest number of half waves of electric intensity that can be established between the walls is three. Cut-off wavelength for this mode is 3·33 cm. Therefore guide wavelength will be

$$\lambda_g=\frac{\lambda_0}{\sqrt{(1-\lambda_0^2/\lambda_c^2)}}=\frac{3}{\sqrt{\left\{1-\left(\frac{3}{3{\cdot}33}\right)^2\right\}}}$$

$=6\,91$ cm.

Ex. 3. *What should be the separation between the planes of a plane parallel guide so that a wave of* 10Gc/s *may propagate in a mode having three as the number of half waves of electric intensity* ($m=3$) *and guide wavelength equal to* 6·91 *cm* ($\lambda_g=6{\cdot}91$ *cm.*)

Free space wavelength is

$$\lambda_0=\frac{c}{f}=\frac{3\times10^{10}}{10\times10^9}=3 \text{ cm.}$$

We know that

$$\lambda_c=\frac{2a}{m}$$

or $$a=\frac{m}{2}\ \lambda_c=\frac{3}{2}\lambda_c \qquad \ldots(1)$$

To calculate cut-off wavelength, we have the relation

$$\lambda_g=\frac{\lambda_0}{\sqrt{(1-\lambda_0^2/\lambda_c^2)}}$$

or $$1-\frac{\lambda_0^2}{\lambda_c^2}=\frac{\lambda_0^2}{\lambda_g^2}$$

or $$\lambda_c=\frac{\lambda_0}{\sqrt{(1-\lambda_0^2/\lambda_g^2)}}=\frac{3}{\sqrt{\left\{1-\left(\frac{3}{6{\cdot}91}\right)^2\right\}}}$$

$$=\frac{3}{0\,9}=3{\cdot}33 \text{ cm.}$$

Therefore putting in eq. (1), we get the separation as

$$a=\frac{3}{2}\ \lambda_0=\frac{3\times3{\cdot}33}{2}$$

$=4{\cdot}99$ cm.

(B) TRANSVERSE MAGNETIC WAVES ; ($H_z=0$)

The wave equation for the y component of magnetic field can be written as

$$\frac{\partial^2 H_y^0}{\partial x^2}+\gamma^2 H_y^0=-\omega^2\mu\epsilon H_y^0$$

or $$\frac{\partial^2 H_y}{\partial x^2} + h^2 H_y^0 = 0,$$

solution of which is

$$H_y^0 = C_3 \sin hx + C_4 \cos hx$$

In order to apply boundary conditions, we may put H_y^0 in terms of electric field component. We know that

$$H_y^0 = -\frac{j\omega\epsilon}{h^2}\frac{\partial E_z^0}{\partial x}$$

so that

$$E_z^0 = -\frac{h^2}{j\omega\epsilon}\int H_y^0\, dx$$

$$= \frac{h^2}{j\omega\epsilon}\,[C_3 \cos hx - C_4 \sin hx].$$

The boundary conditions are then

$$\left.\begin{aligned} E_z^0 &= 0 \quad \text{at} \quad x = 0 \\ E_z^0 &= 0 \quad \text{at} \quad x = a \end{aligned}\right\}$$

Applying these boundary conditions we come at the following two results :

$$C_3 = 0 \quad \text{and} \quad h = \frac{m\pi}{a}.$$

Solution is then

$$E_z^0 = -\frac{h}{j\omega\epsilon} C_4 \sin\left(\frac{m\pi x}{a}\right)$$

$$= -\frac{m\pi}{j\omega\epsilon a} C_4 \sin\left(\frac{m\pi x}{a}\right)$$

and $$H_y^0 = C_4 \cos\left(\frac{m\pi x}{a}\right).$$

The other field components are

$$E_x^0 = \frac{\gamma C_4}{j\omega\epsilon}\cos\left(\frac{m\pi x}{a}\right)$$

and $$E_y^0 = 0$$

Therefore, with $\gamma = j\beta$, the field components of TM waves between parallel perfectly conducting planes are

$$\left.\begin{aligned} H_y &= C_4 \cos\left(\frac{m\pi x}{a}\right) e^{-j\beta z} \\ E_x &= \frac{\beta}{\omega\epsilon} C_4 \cos\left(\frac{m\pi x}{a}\right) e^{-j\beta z} \\ E_z &= \frac{jm\pi}{\omega\epsilon a} C_4 \sin\left(\frac{m\pi x}{a}\right) e^{-j\beta z} \end{aligned}\right\}$$

It is obvious that with $m=0$, all field components are not reduced to zero. That means, in this case, lowest mode will be $TM_{0,0}$. For propagation to be possible we must have

$$\gamma = j\beta = j\sqrt{\left[\omega^2\mu\epsilon - \left(\frac{m\pi}{a}\right)^2\right]}$$

or cut-off frequency is $\omega_c^2\mu\epsilon = \left(\frac{m\pi}{a}\right)^2$

$$f_c = \frac{v_0}{2}\cdot\frac{m}{a}.$$

Thus for a given value of m, there is a corresponding cut-off frequency below which wave propagation cannot occur. The wave or phase velocity with which the wave will propagate between the planes is given by

$$v = \frac{\omega}{\beta} = \frac{\omega}{\sqrt{\left[\omega^2\mu\epsilon - \left(\frac{m\pi}{a}\right)^2\right]}},$$

which predicts that near cut-off, velocity approaches infinity and decreases as the frequency is raised above cut-off, approaching the lowest limit

$$v \to v_0 = \frac{1}{\sqrt{(\mu\epsilon)}}$$

when the frequency is high enough so that term $\left(\frac{m\pi}{a}\right)^2$ is negligible compared with $\omega^2\mu\epsilon$. When the dielectric medium between the planes is air, μ and ϵ have their free space values so that lowest limit of velocity is

$$v_0 = \frac{1}{\sqrt{(\mu_v\epsilon_v)}} = 3\times 10^8 \text{ meters/sec}$$
$$= c, \text{ the velocity of light.}$$

This wave velocity or phase velocity is different from the velocity with which the energy propagates. The latter is called group velocity, v_g, and is related to phase velocity, v, by the relation

$$v\, v_g = c^2.$$

TEM WAVES IN PARALLEL PLANE GUIDES :

In case of *TM* wave, if $m=0$, then field components are

$$\left.\begin{aligned} H_y &= C_4\, e^{-j\beta z} \\ E_x &= \frac{\beta}{\omega\epsilon} C_4\, e^{-j\beta z} \end{aligned}\right\}$$
$$E_z = 0.$$

Thus we find in TM_{00} mode, the component of electric field as well as magnetic field in the direction of propagation is zero, that is, electric and magnetic field both are transverse to the direction of propagation. Therefore, it is *TEM* wave. Further discussion about *TEM* wave will be takenup after discussing waveguide (art. 27·6).

27·3 ATTENUATION IN PARALLEL PLANE GUIDES :

While discussing the propagation along parallel plane guides,

it has been assumed that walls are of infinite conductivity so that losses are zero and $\gamma = j\beta$ (taking $\alpha = 0$). In actual waveguides conductivity is rarely infinite and therefore there will always be some losses. The problem is now to calculate attenuation constant, α.

The voltage and current relations on a two wire transmission lines are given by

$$V = V_0\, e^{-(\alpha + j\beta)z}$$
$$I = I_0\, e^{-(\alpha + j\beta)z}$$

and the average power transmitted is

$$P_{av} = \tfrac{1}{2}\, Re\,[VI^*]$$
$$= \tfrac{1}{2}\, Re\,[V_0\, I_0^*]\, e^{-2\alpha z}$$

The rate of decrease of transmitted power along the line will be

$$-\frac{\partial P_{av}}{\partial z} = 2\alpha P_{av}.$$

The decrease of transmitted power per unit length of line is

$$-\Delta P_{av} = 2\alpha\, P_{av}$$

and this must be equal to the power lost or dissipated per unit length. Therefore

$$\frac{\text{Power lost per unit length}}{\text{Power transmitted}} = \frac{2\alpha\, P_{av}}{P_{av}} = 2\alpha$$

so that

$$\alpha = \frac{\text{Power lost per unit length}}{2 \times \text{Power transmitted}} = \frac{P_{lost}}{2 \times P_{trans.}}$$

We shall now proceed to calculate attenuation constant α in the following three cases :

(A) ATTENUATION FACTOR FOR TEM WAVES :

We have already shown that field components for *TEM* wave are

$$H_y = C_4\, e^{-j\beta z}$$

$$E_x = \frac{\beta}{\omega\epsilon} C_4\, e^{-j\beta z} = \eta\, C_4\, e^{-j\beta z} = \eta\, H_y.$$

Since H_y is wholly tangential, it is continuous across the boundary and represents the magnetic field of the wave entering the metal. The linear current density in each of the conducting planes will be

$$\mathbf{J}_S = \hat{\mathbf{x}} \times \mathbf{H} = \hat{\mathbf{x}} \times \hat{\mathbf{y}}\, H_y = \hat{\mathbf{z}}\, H_y$$

where $\hat{\mathbf{x}}$ is the unit vector in x-direction. Therefore

$$|\, J_S \,| = C_4.$$

The loss per square meter in each conducting plate is

$$= \tfrac{1}{2} J_S^2\, R_S = \tfrac{1}{2} C_4^2\, R_S,$$

where R_S is the resistive component of the surface impedance

$$Z_S=\sqrt{\left(\frac{j\omega\mu_m}{\sigma_m}\right)}$$

where

$$Z_S=R_S+jX_S=\sqrt{\left(\frac{\omega\mu_m}{2\sigma_m}\right)}+j\sqrt{\left(\frac{\omega\mu_m}{2\sigma_m}\right)}$$

or

$$R_S=\sqrt{\left(\frac{\omega\mu_m}{2\sigma_m}\right)}$$

μ_m and σ_m refer to the values in the metallic conductor. The total loss in the upper and lower conducting surfaces per meter length for a width of one meter of the guide is

$$P_{lost}=2\,(\tfrac{1}{2}\,R_S\,J_S^2)$$
$$=R_S\,C_4^2.$$

The power transmitted down the guide per unit cross sectional area is

$$=\tfrac{1}{2}\,R_e\,(\mathbf{E}\times\mathbf{H}^*)_z$$
$$=\tfrac{1}{2}\,(E_xH_y)$$
$$=\tfrac{1}{2}\,\eta C_4^2.$$

Suppose the spacing between the plates is a, then cross-sectional area with width unit meter of plates will be a so that power transmitted through this area will be

$$P_{trans.}=\tfrac{1}{2}\,\eta C_4^2 a.$$

Therefore attenuation factor is

$$\alpha_{TEM}=\frac{P_{lost}}{2\times P_{trans.}}=\frac{R_SC_4^2}{2\times\frac{1}{2}\eta C_4^2 a}=\frac{R_S}{\eta a}$$

$$=\frac{1}{\eta a}\sqrt{\left(\frac{\omega\mu_m}{2\sigma_m}\right)}$$

$$=\frac{1}{\eta a}\sqrt{\left(\frac{\pi f\mu_m}{2\sigma_m}\right)}$$

nepers/meter.

A curve of attenuation versus frequency for TEM wave between parallel planes is drawn in fig. 3.

(B) ATIENUATION FOR TE WAVES :

We have shawn in TE case that field components are

$$\left.\begin{aligned}E_y&=C_1\sin\left(\frac{m\pi x}{a}\right)e^{-j\beta z}\\ H_x&=-\frac{\beta}{\omega\mu}C_1\sin\left(\frac{m\pi x}{a}\right)e^{-j\beta z}\\ H_z&=j\frac{m\pi}{\omega\mu a}C_1\cos\left(\frac{m\pi x}{a}\right)e^{-j\beta z}\end{aligned}\right\}$$

The amplitude of linear current density in the conducting plane will be equal to the tangential component of H *i.e.* H_z at $x=0$ and $x=a$.

$$|J_{Sy}|=|H_z|\ (x=0,\ x=a$$
$$=\frac{m\pi C_1}{\omega\mu a}$$

For this mode, as is obvious, there is no flow of current in the direction of wave propagation. The loss in each plate is

$$P'_{lost}=\tfrac{1}{2}\,R_S\,|\,J_{Sy}\,|^2$$
$$=\frac{m^2\pi^2C_1^{\,2}}{2\omega^2\mu^2a^2}\cdot\sqrt{\left(\frac{\omega\mu_m}{\sigma_m}\right)}$$

Therefore total power lost will be $P_{lost}=2P'_{lost}$.

Power transmitted per unit area $=\tfrac{1}{2}\,R_e\,(\mathbf{E}\times\mathbf{H}^*)_z$

$$=-\tfrac{1}{2}\,(E_y\,H_x)$$
$$=\frac{\beta C_1^{\,2}}{2\omega\mu}\sin^2\left(\frac{m\pi x}{a}\right)$$

The power transmitted in the z-direction through an element of area of width dx and breadth unity, $da=dx$, is

$$=\frac{\beta C_1^{\,2}}{2\omega\mu}\sin^2\left(\frac{m\pi x}{a}\right)dx.$$

Then the power transmitted in the z-direction for a guide *one meter wide* with a spacing between conductors of a meter is

$$\int_{x=0}^{x=a}\frac{\beta C_1^{\,2}}{2\omega\mu}\sin^2\left(\frac{m\pi x}{a}\right)dx=\frac{\beta C_1^{\,2}a}{4\omega\mu}.$$

Therefore,

$$P_{trans.}=\frac{\beta C_1^{\,2}a}{4\omega\mu}.$$

We can now write for attenuation factor

$$\alpha=\frac{P_{lost}}{2\times P_{trans.}}=\frac{2P'_{lost}}{2\times\dfrac{\beta C_1^{\,2}a}{4\omega\mu}}$$
$$=\frac{2\times\dfrac{m^2\pi^2C_1^{\,2}}{2\omega^2\mu^2a^2}\sqrt{\left(\dfrac{\omega\mu_m}{2\sigma_m}\right)}}{2\times\dfrac{\beta C_1^{\,2}a}{4\omega\mu}}$$
$$=\frac{2m^2\pi^2\sqrt{(\omega\mu_m/2\sigma_m)}}{\beta a\,\mu a^3}$$

since

$$\beta=\sqrt{\left\{\omega^2\mu\epsilon-\left(\frac{m\pi}{a}\right)^2\right\}}$$
$$=\sqrt{(\omega^2\mu\epsilon-\omega_c^{\,2}\mu\epsilon)}=\omega\sqrt{(\mu\epsilon)}\sqrt{(1-f_c^{\,2}/f^2)}$$
$$=\frac{\omega}{v}\sqrt{(1-f_c^{\,2}/f^2)}.$$

Therefore,

$$\alpha_{TE}=\frac{2m^2\pi^2\sqrt{(\omega\mu_m/2\sigma_m)}}{\dfrac{\omega}{v}\cdot\sqrt{(1-f_c^{\,2}/f^2)}\cdot\omega\mu a^2}$$

Further

$$\omega_c^{\,2}\mu\epsilon=\left(\frac{m\pi}{a}\right)^2$$

so that

$$f_c=\frac{mv}{2a}$$

Then we can write α in another form. That is,

$$\alpha_{TE}=\frac{2\left(\frac{mv}{2a}\right)^{3/2}.\left(\frac{2a}{mv}\right)^{3/2}.m^2\pi^2.\sqrt{(\omega\mu_m/2\sigma_m)}}{\omega^2a^3\sqrt{(1-f_c^2/f^2)}.\frac{\mu}{v}}$$

$$=\frac{(2)^{5/2}\,(f_c)^{3/2}\,(m)^{1/2}\,\pi^2\sqrt{(\omega\mu_m/2\sigma_m)}}{4\pi^2f^2a^{3/2}\sqrt{(1-f_c^2/f^2)}.v^{1/2}.\mu}$$

$$=\frac{\sqrt{m}.(f_c)^{3/2}\,(f_c)^{1/2}\,\sqrt{(2\pi\mu_m/\sigma_m)}}{f^2.a^{3/2}.\sqrt{(1-f_c^2/f^2)}.\mu^{1/2}.(\sqrt{\mu/\epsilon})^{1/2}}$$

$$=\frac{\sqrt{m}}{a^{3/2}}\cdot\sqrt{\left(\frac{2\pi\mu_m}{\sigma_m\mu\eta}\right)}\,\frac{(f_c/f)^{3/2}}{(1-f_c^2/f^2)^{1/2}}\text{ nepers/m,}$$

where $\eta=\sqrt{(\mu/\epsilon)}$, the intrinsic impedance of the space. The variation of attenuation factor with frequency, f, is depicted in fig. 3 The attenuation factor, as is obvious from the above expression, depends upon (*i*) the function of (f/f_c), (*ii*), on m, and (*iii*) the metal plate parameters. It decreases with increasing frequency, becomes zero for infinite frequency.

(C) ATTENUATION FOR TM WAVES :

We have shown in *TM* case that the field components are

$$\left.\begin{aligned}H_y&=C_4\cos\left(\frac{m\pi x}{a}\right)e^{-j\beta z}\\E_x&=\frac{\beta}{\omega\epsilon}C_4\cos\left(\frac{m\pi x}{a}\right)e^{-j\beta z}\\E_z&=\frac{jm\pi}{\omega\epsilon a}C_4\sin\left(\frac{m\pi x}{a}\right)e^{-j\beta z}\end{aligned}\right\}$$

The amplitude of linear current density in the conducting plane will be equal to the tangential component of H, that is,

$$J_s=|\,H_y\,|=C_4.$$

Then the loss in each plate one meter wide is

$$P'_{lost}=\tfrac{1}{4}\,R_s\,|\,J_s\,|^2=\tfrac{1}{2}\,R_sC_4^2$$

$$=\tfrac{1}{2}\,C_4^2\left(\frac{\omega\mu_m}{2\sigma_m}\right),$$

so that the total power lost per unit length of both the plates, each being one meter wide is

$$P_{lost}=2\times\tfrac{1}{2}\,C_4^2\,\sqrt{\left(\frac{\omega\mu_m}{2\sigma_m}\right)}$$

$$=C_4^2\,\sqrt{\left(\frac{\omega\mu_m}{2\sigma_m}\right)}.$$

The power transmitted in z-direction per unit area of cross section is

$$=\tfrac{1}{2}\,R_e\,(\mathbf{E}\times\mathbf{H}^*)_z$$

$$= \tfrac{1}{2} (E_x H_y)$$

$$= \tfrac{1}{2} \frac{\beta C_4^2}{\omega \epsilon} \cos^2 \left(\frac{m\pi x}{a} \right)$$

Power transmitted in the z-direction for a guide ***one meter wide*** in y direction with a spacing between conductors of a meters is

$$P_{trans} = \int_0^a \tfrac{1}{2} \frac{\beta C_4^2}{\omega \epsilon} \cos^2 \left(\frac{m\pi x}{a} \right) dx$$

$$= \frac{\beta C_4^2 a}{4\omega \epsilon}.$$

Therefore attenuation factor is

$$\alpha_{TM} = \frac{P_{lost}}{2 \times P_{trans.}} = \frac{C_4^2 \sqrt{(\omega \mu_m / 2\sigma_m)}}{2 \times \frac{\beta C_4^2 a}{4\omega \epsilon}}$$

$$= \frac{2\omega \epsilon}{a\beta} \sqrt{\left(\frac{\omega \mu_m}{2\sigma_m} \right)}.$$

This expression can be further modified by putting the value of β and manipulating in the way adopted already for *TE* case. We can put it like :

$$\alpha_{TM} = \frac{\sqrt{m}}{a^{3/2}} \sqrt{\left(\frac{2\pi \mu_m}{\sigma_m \mu \eta} \right)} \cdot \frac{(f_c/f)^{3/2}}{\sqrt{(1 - f_0^2/f^2)}} \text{ nepers/meter.}$$

This is also an involved function of (f/f_c) as well as being dependent on m, and on the metal plane parameters. Obviously,

$$\alpha_{TE} = \left(\frac{f_0}{f} \right)^2 \alpha_{TM}.$$

The variation of attenuation factor with frequency for *TM* wave is shown in fig. 3. The minimum in the attenuation curve for this mode may be shown to occur at

$$f = \sqrt{3} f_0.$$

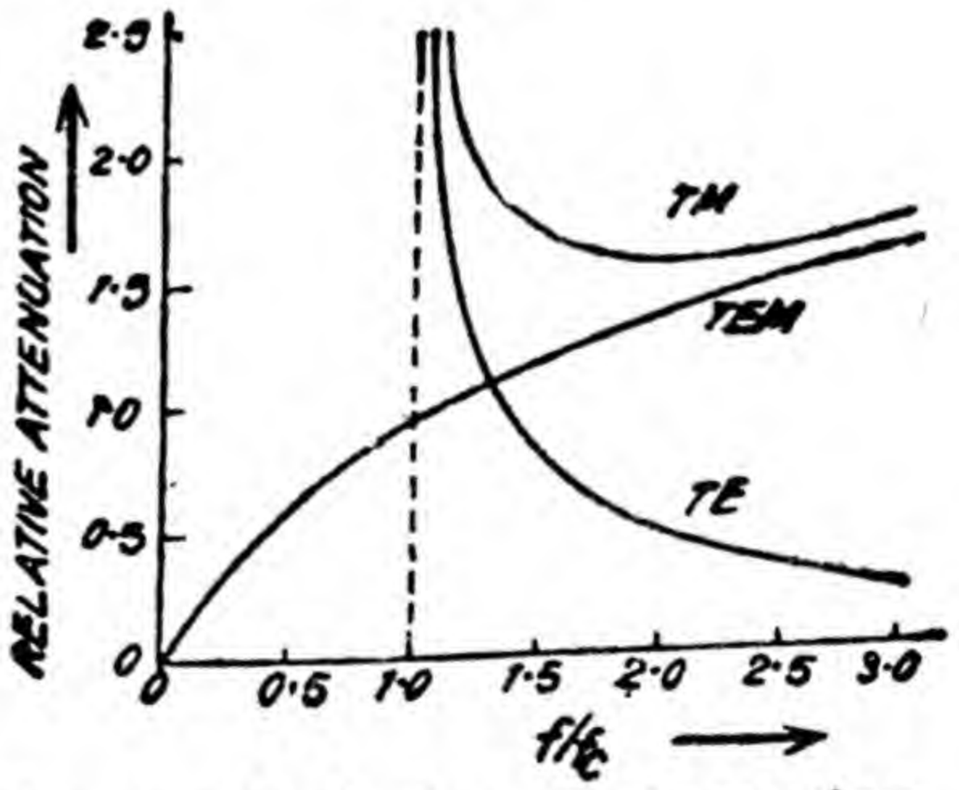


Fig. 3. Variation of attenuation of *TM*, *TE*, and *TEM* modes between parallel planes ($m=1$).

27·4. RECTANGULAR WAVEGUIDE :

When the cross-section of the metal tube is rectangular, it is called rectangular waveguide and if circular, then it is called circular waveguide.

It will have two pairs of conducting plates. The solution of Maxewell's equations to determine the electromagnetic field configuration in the guide, will be obtained by applying the boundary conditions at the walls of the guide. Maxwell's equations

for the loss free region within the guide (the direction of propagation is z and since guide is finite in both x and y directions $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ will not be zero), on putting $\frac{\partial}{\partial z}=-\gamma$ in eq. (1) art. 27·1 are

$$\left.\begin{aligned}
&\frac{\partial H_z}{\partial y}+\gamma H_y=j\omega\epsilon E_x, \qquad \frac{\partial E_z}{\partial y}+\gamma E_y=-j\omega\mu H_x,\\
&\frac{\partial H_z}{\partial x}+\gamma H_x=-j\omega\epsilon E_y, \quad \frac{\partial E_z}{\partial x}+\gamma E_x=j\omega\mu H_y,\\
&\frac{\partial H_y}{\partial x}-\frac{\partial H_x}{\partial y}=j\omega\epsilon E_z, \qquad \frac{\partial E_y}{\partial x}-\frac{\partial E_x}{\partial y}=-j\omega\mu H_z,
\end{aligned}\right\} \quad \ldots(1)$$

and wave equations for E_z and H_z are

$$\left.\begin{aligned}
&\frac{\partial^2 E_z}{\partial x^2}+\frac{\partial^2 E_z}{\partial y^2}+\gamma^2 E_z=-\omega^2\mu\epsilon E_z,\\
&\frac{\partial^2 H_z}{\partial x^2}+\frac{\partial^2 H_z}{\partial y^2}+\gamma^2 H_z=-\omega^2\mu\epsilon H_z.
\end{aligned}\right\} \quad \ldots(2)$$

Equation (1) can be combined to give

$$\left.\begin{aligned}
\mathbf{H}_x&=-\frac{\gamma}{\mathbf{h}^2}\frac{\partial \mathbf{H}_z}{\partial \mathbf{x}}+\frac{\mathbf{j}\omega\epsilon}{\mathbf{h}^2}\frac{\partial \mathbf{E}_z}{\partial \mathbf{y}},\\
\mathbf{H}_y&=-\frac{\gamma}{\mathbf{h}^2}\frac{\partial \mathbf{H}_z}{\partial \mathbf{y}}-\frac{\mathbf{j}\omega\epsilon}{\mathbf{h}^2}\frac{\partial \mathbf{E}_z}{\partial \mathbf{x}},\\
*\mathbf{E}_x&=-\frac{\gamma}{\mathbf{h}^2}\frac{\partial \mathbf{E}_z}{\partial \mathbf{x}}-\frac{\mathbf{j}\omega\mu}{\mathbf{h}^2}\frac{\partial \mathbf{H}_z}{\partial \mathbf{y}},\\
\mathbf{E}_y&=-\frac{\gamma}{\mathbf{h}^2}\frac{\partial \mathbf{E}_z}{\partial \mathbf{y}}+\frac{\mathbf{j}\omega\mu}{\mathbf{h}^2}\frac{\partial \mathbf{H}_z}{\partial \mathbf{x}},
\end{aligned}\right\}\dagger \quad \ldots(3)$$

where $h^2=\gamma^2+\omega^2\mu\epsilon$.

These equations give relationship among the fields within the guide. Now we shall again here divide the possible field confi-

---

*From equation (1),

$$E_x=-\frac{j}{\omega\epsilon}\frac{\partial H_z}{\partial y}-\frac{j\gamma}{\omega\epsilon}H_y$$

$$=-\frac{j}{\omega\epsilon}\frac{\partial H_z}{\partial y}-\frac{j\gamma}{\omega\epsilon}\left[\frac{1}{j\omega\mu}\left(\frac{\partial E_z}{\partial x}+\gamma E_x\right)\right]$$

$$=-\frac{j}{\omega\epsilon}\frac{\partial H_z}{\partial y}-\frac{\gamma}{\omega^2\mu\epsilon}\frac{\partial E_z}{\partial x}-\frac{\gamma^2 E_x}{\omega^2\mu\epsilon}$$

or $$\frac{(\gamma^2+\omega^2\mu\epsilon)\,E_x}{\omega^2\mu\epsilon}=-\frac{j}{\omega\epsilon}\frac{\partial H_z}{\partial y}-\frac{\gamma}{\omega^2\mu\epsilon}\frac{\partial E_z}{\partial x},$$

$$E_x=-\frac{j\omega\mu}{h^2}\frac{\partial H_z}{\partial y}-\frac{\gamma}{h^2}\frac{\partial E_z}{\partial x}.$$

Similarly, other equations can be obtained.

guration within the guide into two sets :

(1) Transverse magnetic waves (*TM* mode), for which $H_z=0$.

(2) Transverse electric wave (*TE* mode); for which $E_z=0$.

For rectangular waveguide (fig. 4), the boundary conditions are

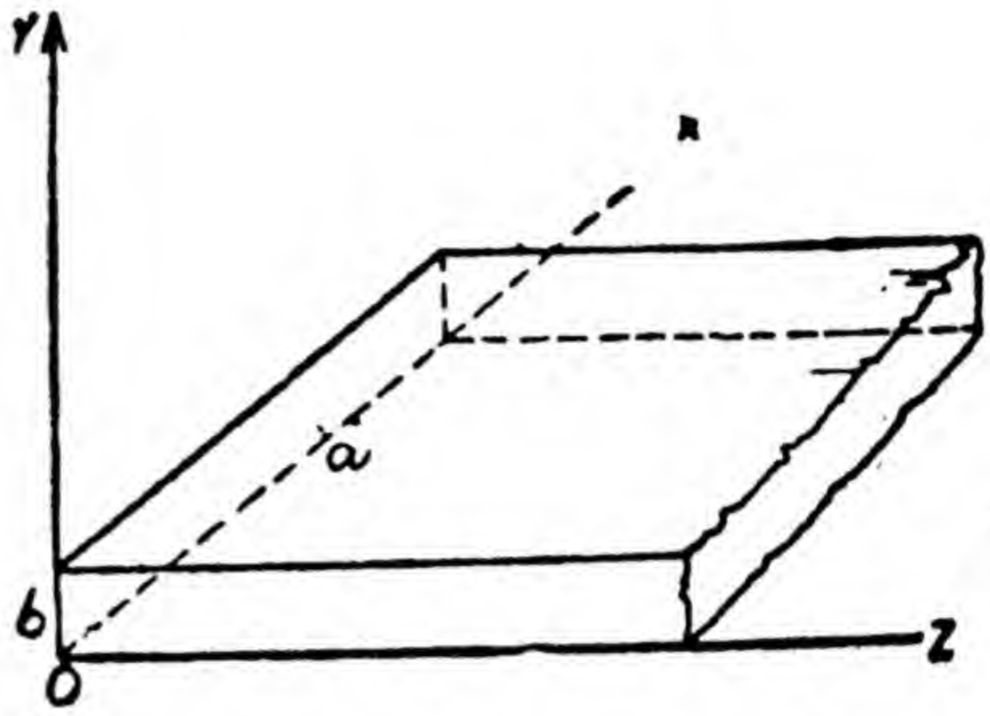


Fig 4. Rectangular waveguide.

$E_{tangential}=0$, *i.e.* $E_x=E_z=0$ at $y=0$ and $y=b$,

$E_{tangential}=0$, *i.e.* $E_y=E_z=0$ at $x=0$ and $x=a$.

27·4·1. TRANSVERSE MAGNETIC WAVES :

TM mode in rectangular guide : (*a*) Field components : We shall solve the partial differential equation (2) to obtain the values of field components corresponding to this case.

For *TM* mode, H_z is zero while E_z will be a function of x and y coordinates. Let $E_z(x, y, z)=E_z^\circ(x, y)\, e^{-\gamma z}$, with

$$E_z^\circ=X(x),\ Y(y). \qquad \ldots(4)$$

Putting eq. (4) in eq. (2), we get

$$Y\frac{d^2X}{dx^2}+X\frac{d^2Y}{dy^2}+\gamma^2XY=-\omega^2\mu\epsilon XY. \qquad \ldots(5)$$

Substituting the value of γ^2 from

$$h^2=\gamma^2+\omega^2\mu\epsilon, \qquad \ldots(6)$$

$$Y\frac{d^2X}{dx^2}+X\frac{d^2Y}{dy^2}+h^2XY=0$$

or

$$\frac{1}{X}\frac{d^2X}{dx^2}+\frac{1}{Y}\frac{d^2Y}{dy^2}+h^2=0$$

or

$$\frac{1}{X}\frac{d^2X}{dx^2}+h^2=-\frac{1}{Y}\frac{d^2Y}{dy^2}. \qquad \ldots(7)$$

In any equation, a function, purely dependent on x, cannot be equated to some other function, which on the other hand, purely depends upon another independent co-ordinate y unless each of them is equal to a constant, *i.e.*

$$\frac{1}{X}\frac{d^2X}{dx^2}+h^2=-\frac{1}{Y}\frac{d^2Y}{dy^2}=A^2. \qquad \ldots(8)$$

Thus we should solve two equations ;

first

$$\frac{1}{X}\frac{d^2X}{dx^2}+h^2=A^2, \qquad \ldots(9)$$

solution of which is

$$X=C_1\cos Bx+C_2\sin Bx,$$

where $B^2=h^2-A^2$.

And *second* $$\frac{1}{Y}\frac{d^2Y}{dy^2}=-A^2, \qquad \ldots(10)$$

whose solution is $Y=C_3\cos Ay+C_4\sin Ay$.

Thus, $E_z^0=XY$

$$=C_1C_3\cos Bx.\cos Ay+C_1C_4\cos Bx.\sin Ay$$
$$+C_1C_3\sin Bx.\cos Ay+C_2C_4\sin Bx\sin Ay. \qquad \ldots(11)$$

To calculate the values of constant involved in above equation, boundary conditions will be applied. We know that

$$E_z^0=0,$$

when $x=0$, $x=a$ and $y=0$, $y=b$.

Applying $x=0$, equation (11) takes the form

$$E_z^0=C_1C_3\cos Ay+C_1C_4\sin Ay=0,$$

which shows that for all values of y, E_z will vanish if $C_1=0$.

Thus equation (11) is reduced to

$$E_z^0=C_2C_3\sin Bx\cos Ay+C_2C_4\sin Bx\sin Ay$$

which, when $y=0$, reduces to

$$E_z^0=C_2C_3\sin Bx=0.$$

Assuming $B\neq 0$, C_2 or C_3 should be zero to satisfy above equation. But if C_2 is zero, whole expression for E_z^0 will become zero.

Thus taking $C_3=0$, equation (11) will be reduced to

$$E_z^0=C_2C_4\sin Bx\sin Ay$$
$$=C\sin Bx.\sin Ay$$

Now if $x=a$, then

$$E_z^0=C\sin Ba.\sin Aj=0.$$

This expression to vanish, for all values of y, A cannot be taken as zero since it would make E_z^0 identically zero and the only alternative is that B should assume the value

$$B=\frac{m\pi}{a}, \qquad \text{where } m=1, 2, 3\ldots$$

Again if $y=b$, then

$$E_z^0=C\sin\frac{m\pi x}{a}.\sin Ab.$$

For this to vanish for values of x, A must have the value

$$A=\frac{n\pi}{b}. \qquad \text{where } n=1, 2, 3,\ldots$$

Therefore, the final solution for E_z^0 is

$$E_z^0=C\sin\frac{m\pi x}{a}.\sin\frac{n\pi y}{b}. \qquad \ldots(12)$$

We shall see from eq. (14) that for the waves guided by *perfectly conducting walls*, γ $(=\alpha+j\beta)$ is either purely real or purely imaginary. In the range of frequencies, where γ is real, β is zero,

i.e., no wave motion; but in the range of frequencies, where γ is imaginary, α is zero but β has some value and therefore, there is propagation by wave motion without attenuation. The latter range is of chief interest in waveguide propagation. With each type of wave, *a cut off frequency* is associated below which, for a guide of given size, propagation becomes impossible. In general, value of cut off frequency for each type of wave is lowered with increasing size of the guide.

Now from equations (3) and (12), on putting $H_z=0$ and $\gamma=j\beta$ for frequencies above the cut off frequency, we can get

$$\left.\begin{aligned} E_x^0 &= \frac{-j\beta C}{h^2} B \cos Bx \sin Ay, \\ E_y^0 &= \frac{-j\beta C}{h^2} A \sin Bx \cos Ay, \\ H_x^0 &= \frac{j\omega\epsilon C}{h^2} A \sin Bx \cos Ay, \\ H_y^0 &= -\frac{j\omega\epsilon C}{h^2} B \cos Bx \sin Ay, \end{aligned}\right\} \quad \ldots(13)$$

where $B=\frac{m\pi}{a}$ and $A=\frac{n\pi}{b}$.

In above values of B and A, m and n are integers, while a and b are the width and height of the waveguide.

(*b*) Cutoff wavelength :

We know that

$$B^2=h^2-A^2 \quad \text{or} \quad h^2=A^2+B^2.$$

Also, $h^2=\gamma^2+\omega^2\mu\epsilon.$

Therefore, $\gamma=\sqrt{(h^2-\omega^2\mu\epsilon)}$

$$=\sqrt{(A^2+B^2-\omega^2\mu\epsilon)}$$

$$=\sqrt{\left\{\left(\frac{m\pi}{a}\right)^2+\left(\frac{n\pi}{a}\right)^2-\omega^2\mu\epsilon\right\}}. \quad \ldots(14)$$

Equation (14) represents the value of propagation constant for *TM* waves in a rectangular guide. For low frequencies, $\omega^2\mu\epsilon$ is small and γ is a real number for which β must be zero. When β is zero, there can be no phase shift along the tube, *i.e.* no wave motion along the guide for low frequencies.

On increasing the frequency further, a stage will come when $\omega^2\mu\epsilon=h^2$, and $\gamma=0$. This value of ω is called cut-off angular frequency and is denoted by ω_c. Now for frequencies greater than ω_c, (*i.e.* $\omega > \omega_c$), γ will be imaginary and will have the from γ $j\beta$. For perfectly conducting walls attenuation constant α is zero for all frequencies when $\omega > \omega_c$. For the frequencies above cutoff value, we have

$$\gamma=j\beta=\sqrt{j\left\{\left(\frac{m\pi}{a}\right)^2+\left(\frac{n\pi}{b}\right)^2-\omega^2\mu\epsilon\right\}},$$

$$*\beta = \sqrt{\left\{\omega^2\mu\epsilon - \left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2\right\}}. \qquad \text{...(15)}$$

The cutoff frequency is given by

$$\omega_0{}^2\mu\epsilon = h^2.$$

$$\therefore \quad \omega_0 = \frac{1}{\sqrt{(\mu\epsilon)}}\sqrt{\left\{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2\right\}},$$

or
$$f_c = \frac{1}{2\pi\sqrt{(\mu\epsilon)}}\sqrt{\left\{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2\right\}}$$

$$= \frac{\dagger v_0}{2}\sqrt{\left\{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2\right\}}$$

or
$$\frac{v_0}{f_c} = \frac{2}{\sqrt{\left\{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2\right\}}}$$

or
$$\lambda_c = \frac{2}{\sqrt{\left\{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2\right\}}} \qquad \text{...(16)}$$

In expression (16) λ_c is called cut-off wavelength, v_0 in above expression is the velocity of electromagnetic wave in a dielectric having constants μ and ϵ. Equation shows that cut-off wavelength depends upon the size (a, b) of the guide and the mode of propagation (m, n). Propagation is possible only when the wavelength of the wave is less than the cut-off wavelength $(\lambda_0 < \lambda_c)$.

(*c*) **Guide wavelength :**

Now the velocity of wave propagation in the guide will be

$$v_g = \frac{\omega}{\beta} = \frac{\omega}{\sqrt{\left\{\omega^2\mu\epsilon - \left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2\right\}}}$$

$$= \frac{\omega}{\sqrt{(\omega^2\mu\epsilon - \omega_c{}^2\mu\epsilon)}} = \frac{v_0}{\sqrt{\left(1 - \frac{\omega_0{}^2}{\omega^2}\right)}} \qquad \text{...(17)}$$

This expression shows that velocity of wave propagation in the guide, v_g, is greater than the phase velocity in free space. The phase velocity decreases from an infinitely large value when the frequency is increased above cut-off and for frequencies much above cut-off, it approaches v_0 the velocity in free space.

---

*Since $\beta = \sqrt{(\omega^2\mu\epsilon - \omega_c{}^2\mu\epsilon)}$, wave motion is possible only when frequency of the wave ω is greater than ω_c, the cut-off frequency.

$\dagger v_0 = \dfrac{1}{\sqrt{(\mu\epsilon)}}$ velocity of the wave in a dielectric of permeability μ and permittivity ϵ. Since here dielectric is air, these are referred to free space and v_0 is called velocity in free space.

The guide wavelength λ_g is given by v_g/f. Since v_g is greater than phase velocity in free space, λ_g will be longer than the corresponding free space wavelength. Thus

$$\lambda_g = \frac{\omega}{f\sqrt{\left\{\omega^2\mu\epsilon - \left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2\right\}}}$$

$$= \frac{2\pi}{\sqrt{\left\{\omega^2\mu\epsilon - \left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2\right\}}}. \quad \text{...(18)}$$

We thus find that

$$\lambda_g = \frac{v_0}{f} = \frac{v_0}{f\sqrt{\left(1 - \frac{\omega_c^2}{\omega^2}\right\}}}$$

$$= \frac{\lambda_0}{\sqrt{\left(1 - \frac{\omega_c^2}{\omega^2}\right)}}$$

$$= \frac{\lambda_0}{\sqrt{\left(1 - \frac{\lambda_0^2}{\lambda_c^2}\right)}}$$

or

$$\frac{1}{\lambda_0^2} = \frac{1}{\lambda_c^2} + \frac{1}{\lambda_g^2}$$

which is called guide equation.

Two subscripts m and n are used to designate a particular mode. The first subscript m indicates the number of half period variations of electric field (even in TM mode) along x-axis which is taken *along the wide dimension of the waveguide* and the second subscript n indicates the number of half period variations of electric field along y-axis taken along *narrow dimension.*

We see from equation (13) that if $m=0$ or $n=0$, all the field components become zero. Thus lowest value for m or n can be unity. For $m=n=1$, the lowest cut-off frequency will occur. This particular wave is called TM_{11} wave. For the propagation of TM waves with larger value of m and n, through a guide of given dimensions, higher frequencies are required.

27·4-2. TRANSVERSE ELECTRIC WAVE :

TE mode in rectangular guide : We shall solve the partial differential equation (2) to obtain the values of field components corresponding to this case.

(a) Field Components : For TE mode E_z is zero, while H_z will be a function of x and y coordinates. Let

$$H_z = H^0_z\, e^{-\gamma z}$$

with

$$H_z^0 = X(x)\, Y(y).$$

Putting above equation in eq. (2), we get

$$Y\frac{d^2X}{dx^2}+X\frac{d^2Y}{dy^2}+\gamma^2XY=-\omega^2\mu\epsilon XY.$$

Substituting the value of γ^2 from

$$h^2=\gamma^2+\omega^2\mu\epsilon,$$

we get, $$Y\frac{d^2X}{dx^2}+X\frac{d^2Y}{dy^2}+h^2XY=0$$

or $$\frac{1}{X}\frac{d^2X}{dx^2}+h^2=-\frac{1}{Y}\frac{d^2Y}{dy^2}$$

Now proceeding exactly in a parallel way as in transverse magnetic waves, we shall finally obtain

$$H_z^0=XY=C_1C_3\cos Bx\cos Ay+C_1C_4\cos Bx\sin Ay$$
$$+C_2C_3\sin Bx\cos Ay+C_2C_4\sin Bx\sin Ay. \quad ...(19)$$

Differentiating above equation with respect to x,

$$\frac{\partial H_z^0}{\partial x}=-C_1C_3B\cos Ay\sin Bx-C_1C_4B\sin Ay.\sin Bx$$
$$+C_2C_3B\cos Ay\cos Bx+C_2C_4B\cos Bx\sin Ay.$$

Now from equation (9), since $E_z^0=0$,

$$\frac{\partial H_z^0}{\partial x}=\frac{h^2}{j\omega\mu}E_y^0$$

$$\therefore \quad \frac{h^2}{j\omega\mu}E_y^0=-C_1C_3B\cos Ay\sin Bx-C_1C_4B\sin Ay.\sin Bx$$
$$+C_2C_3B\cos Ay.\cos Bx+C_2C_4B\sin Ay.\cos Bx. \quad ...(20)$$

Now we shall apply the *boundary condition*,

$$E_y^0=0 \text{ at } x=0$$

At $x=0$

$$\frac{h^2}{j\omega\mu}E_y^0=C_2C_3B\cos Ay+C_2C_4B\sin Ay=0,$$

which shows that for all values of y, E_y^0 will vanish if $C_2=0$. Thus equation (20) reduces to

$$\frac{h^2}{j\omega\mu}E_y^0=-C_1C_3B\cos Ay\sin Bx-C_1C_4B\sin Ay\sin Bx. \quad ...(21)$$

Again differentiating eq. (19) with respect to y.

$$\frac{\partial H_z^0}{\partial y}=-C_1C_3A\cos Bx\sin Ay+C_1C_4A\cos Bx\cos Ay$$
$$-C_2C_3A\sin Bx.\sin Ay+C_2C_4A\sin Bx\cos Ay.$$

But from equation (9), since $E_z^0=0$, we get

$$\frac{\partial H_z^0}{\partial y}=-\frac{h^2}{j\omega\mu}E_x^0.$$

Therefore,

$$-\frac{h^2}{j\omega\mu}E_x^0=-C_1C_3A\cos Bx.\sin Ay+C_1C_4A\cos Bx\cos Ay$$
$$-C_2C_3A\sin Bx.\sin Ay+C_2C_4A\sin Bx\cos Ay \quad ..(22)$$

Again the *boundary condition is*

$$E_x^0=0 \text{ at } y=0$$

At $y=0$,

$$-\frac{h^2}{j\omega\mu} E_x^0=C_1C_4A\cos Bx+C_2C_4A\sin Bx=0.$$

which shows that for all values of x, E_x^0 will vanish if $C_4=0$.

Thus equation (22) reduces to

$$-\frac{h^2}{j\omega\mu} E_x^0=-C_1C_3A\cos Bx\ \sin Ay-C_2C_3A\sin Bx\sin Ay \quad ...(23)$$

Since constants $C_2=C_4=0$, equations (21) and (23) can be written as

$$\frac{h^2}{j\omega\mu} E_y^0=-C_1C_3B\cos Ay\sin Bx. \quad ...(24)$$

and

$$-\frac{h^2}{j\omega\mu} E_x^0=-C_1C_3A\cos Bx\sin Ay. \quad ...(25)$$

Now at *boundary condition* $x=a$, $E_y^0=0$, equation (24) will be satisfied for all values of y if

$$Ba=m\pi$$

or

$$B=\frac{m\pi}{a}, \qquad \text{where } m=0, 1, 2, 3..$$

Similarly, at *boundary condition* $y=b$, $E_x^0=0$, equation (25) will be satisfied for all values of x if

$$Ab=n\pi$$

or

$$A=\frac{n\pi}{b}, \qquad \text{where } n=0, 1, 2, ,...$$

Putting $C_1C_3=C$, equations (24) and (25) become

$$\mathbf{E_y^0=-\frac{j\omega\mu}{h^2}.CB\ \cos Ay\sin Bx} \quad ...(26)$$

$$\mathbf{E_x^0=\frac{j\omega\mu}{h^2}\ CA\cos Bx\sin Ay} \quad ...(27)$$

and for H_z^0, put $C_2=C_4=0$ and $C_1C_3=C$ eq. (19), so that

$$\mathbf{H_z^0=C\cos Ay\cos Bx.} \quad ...(28)$$

To calculate the values of other field components, we see that from equation (13),

$$H_x^0=-\frac{\gamma}{h^2}\cdot\frac{\partial H_z^0}{\partial x}$$

$$=-\frac{\gamma}{h^2}\ \frac{h^2E_y^0}{j\omega\mu}.$$

Putting $\gamma=j\beta$, and value of E_y^0,

$$\mathbf{H_x^0}=-\frac{j\beta}{h^2}\ \frac{h^2}{j\omega\mu}\times-\frac{j\omega\mu}{h^2}\ CB\cos Ay\sin Bx$$

$$=\frac{j\beta}{\mathbf{h^2}}\ \mathbf{CB\cos Ay\sin Bx}. \quad ...(29)$$

Similarly, $\mathbf{H_y^0 = \frac{j\beta}{h^2}\ CA \cos Bx \sin Ay}$ ...(30)

Field equations (26), (27), (28), (29) and (30) gives the field component values in the case of *TE* mode.

(b) Cutoff wavelength :

Quite similar to *TM* case, for frequencies above cut off

$$\gamma = j\beta = \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2 - \omega^2\mu\epsilon}$$

or $$\beta = \sqrt{\left\{\omega^2\mu\epsilon - \left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2\right\}}$$

The cutoff frequency is given by

$$\omega_c^2\mu\epsilon = \left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2$$

or $$f_c = \frac{v_0}{2}\left[\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2\right]^{1/2} \quad \text{...(31)}$$

or $$\lambda_c = \frac{v_0}{f_c}$$

$$= \frac{2}{\sqrt{\left\{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2\right\}}} \quad \text{...(32)}$$

(c) Guide wavelength :

Velocity of propagation within the guide

$$v_g = \frac{\omega}{\beta} = \frac{\omega}{\sqrt{\omega^2\mu\epsilon - \omega_c^2\mu\epsilon}}$$

$$= \frac{1/\sqrt{\mu\epsilon}}{\sqrt{1 - \omega_c^2/\omega^2}} = \frac{v_0}{\sqrt{1 - \omega_c^2/\omega^2}} \quad \text{...(33)}$$

Gu de wavelength is

$$\lambda_g = \frac{v_g}{f}$$

$$= \frac{v_0}{f} \cdot \frac{1}{\sqrt{1 - \omega_c^2/\omega^2}}$$

$$\lambda_g = \frac{\lambda_0}{\sqrt{1 - \lambda_0^2/\lambda_c^2}} \quad \text{. (34)}$$

or $$1 - \frac{\lambda_0^2}{\lambda_c^2} = \frac{\lambda_0^2}{\lambda_c^2}$$

or $$\frac{1}{\lambda_0^2} = \frac{1}{\lambda_c^2} + \frac{1}{\lambda_g^2} \quad \text{...(35)}$$

Above relation is called guide equation. It is also obvious that

$$v_g > v_0$$

i.e. velccity in the guide is greater than wave velocity in free space.

$TE_{1,0}$ mode :

In the case of *TE* wave, we see from equations (26), (27), (28), (29) and (30) that all field components do not vanish if either *m* or *n* (but not both) is given a value zero, —quite contrary to *TM* case. Thus in the present case $TE_{1,0}$ and $TE_{0,1}$ modes are possible; *m* subscript is always taken across wide dimensions of the guide and hence $TE_{1,0}$ *wave has a lower cut-off frequency than the* $TE_{0,1}$ wave as shown below :

We know cut-off frequency is

$$\omega_c^2\mu\epsilon=\left(\frac{m\pi}{a}\right)^2+\left(\frac{n\pi}{b}\right)^2$$

or
$$f_c=\frac{v_0}{2}\cdot\sqrt{\left[\left(\frac{m}{a}\right)^2+\left(\frac{n}{b}\right)^2\right]}$$

For $TE_{1,0}$ mode $m=1$; $n=0$.

So
$$(f_c)_{1,0}=\frac{v_0}{2}\cdot\frac{1}{a}.$$

Also for $TE_{0,1}$ mode $m=0$, $n=1$.

So
$$(f_c)_{0,1}=\frac{v_0}{2}\cdot\frac{1}{b}$$

Since $b < a$,

$$(f_c)_{1,0} < (f_c)_{0,1}.$$

The wave, which has the lowest cut-off frequency, is called *dominant wave*. For $TE_{1,0}$ wave used in most experimental work, t he field components can be written by putting $m=1$ and $n=0$ so that $B=\pi/a$, $A=0$ and $h=\pi/a$, in field equations *i.e.*,

$$\left.\begin{aligned}
\mathbf{E}_y^0&=\frac{-\mathbf{j}\omega\mu}{\mathbf{h}^2}\,\mathbf{C.\,B.\,\sin}\,\frac{\pi\mathbf{x}}{\mathbf{a}}\\
&=\frac{-j\omega\mu\mathbf{a}}{\pi}\,\mathbf{C.sin}\,\frac{\pi x}{a}\\
\mathbf{E}_x^0&=\mathbf{H}_y^0=0,\\
\mathbf{H}_x^0&=\frac{j\beta\mathbf{a}\mathbf{C}}{\pi}.\mathbf{sin}\,\frac{\pi\mathbf{x}}{\mathbf{a}}\\
\mathbf{H}_z^0&=\mathbf{C}\,\mathbf{cos}\,\frac{\pi\mathbf{x}}{\mathbf{a}}.
\end{aligned}\right\}\quad\ldots(36)$$

$$\beta=\sqrt{\left\{\ \omega^2\mu\epsilon-\left(\frac{\pi}{\mathbf{a}}\right)^2\right\}},$$

$$\mathbf{f}_c=\frac{v_0}{2\mathbf{a}}\ ;\ \lambda_c=2\mathbf{a}\ ;\ \mathbf{h}=\frac{\pi}{\mathbf{a}}$$

Thus for $TE_{1,0}$ wave the cut-off frequency is that frequency for which the corresponding free space half wavelength is equal to the width of the guide. *The waves with wavelengths longer than λ_c cannot be transmitted through the guide.* Thus the waveguide is a form of high pass filter capable of transmitting a wave of frequency greater than the cut-off frequency $f_o = \frac{v_0}{\lambda_c}$. Variation of $\frac{\lambda_g}{\lambda_c}$ with $\frac{\lambda_0}{\lambda_o}$ is shown in fig. 5.

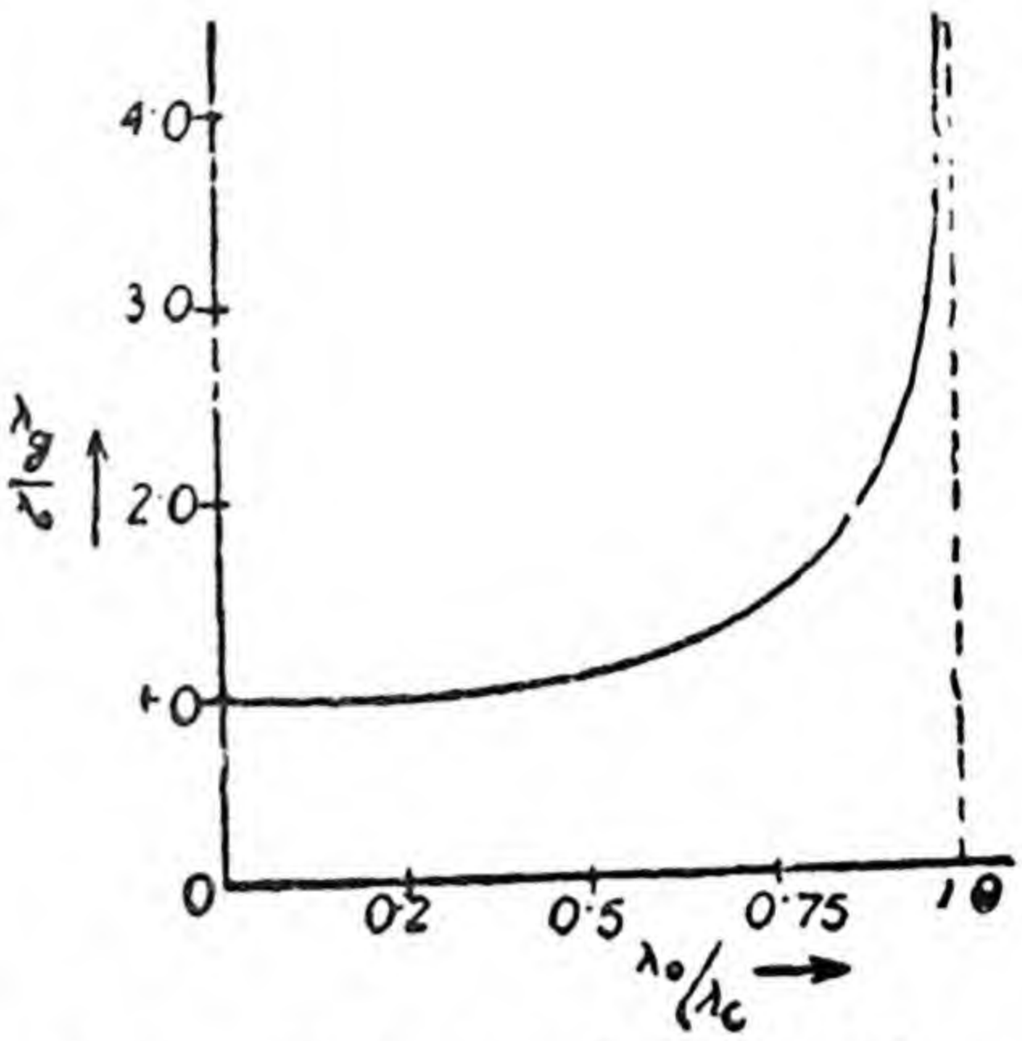


*Fig. 5. Variation of λ_g/λ_0 with λ_0/λ_o.

For $TE_{1,0}$ mode the guide equation will take the form

$$\lambda_g = \frac{\lambda_0}{\sqrt{\{1-(\lambda_0/2a)^2\}}}.$$

We see from the curve that

†(*a*) at very high frequencies (for which $\lambda_0/\lambda_c \to 0$), $\lambda_g \approx \lambda_0$, *i.e.*, guide wavelength nearly equals the free space wavelength.

‡(*b*) as the frequency is decreased, λ_0 and ratio λ_g/λ_0 both increases, and hence guide wavelength, λ_g, increases.

(*c*) if $\lambda_0/\lambda_c < 0{\cdot}5$, change in λ_g/λ_0 is not remarkable but for the value of λ_0 approaching λ_o, λ_c/λ_0 and the guide wavelength. λ_g, increase rapidly and approaches infinity.

As is obvious from end view, there is voltage between top and bottom which is maximum at the centre. To show this maximum, more solid arrows with tail (represented by cross, ×) above and point head (represented by dot, •) below, are drawn in the centre (fig. 6*d*). At the side walls, the voltage is short cut by them and hence assumes zero value there. After quarter of a cycle voltage is zero, but again appears with reversed polarity after further quarter

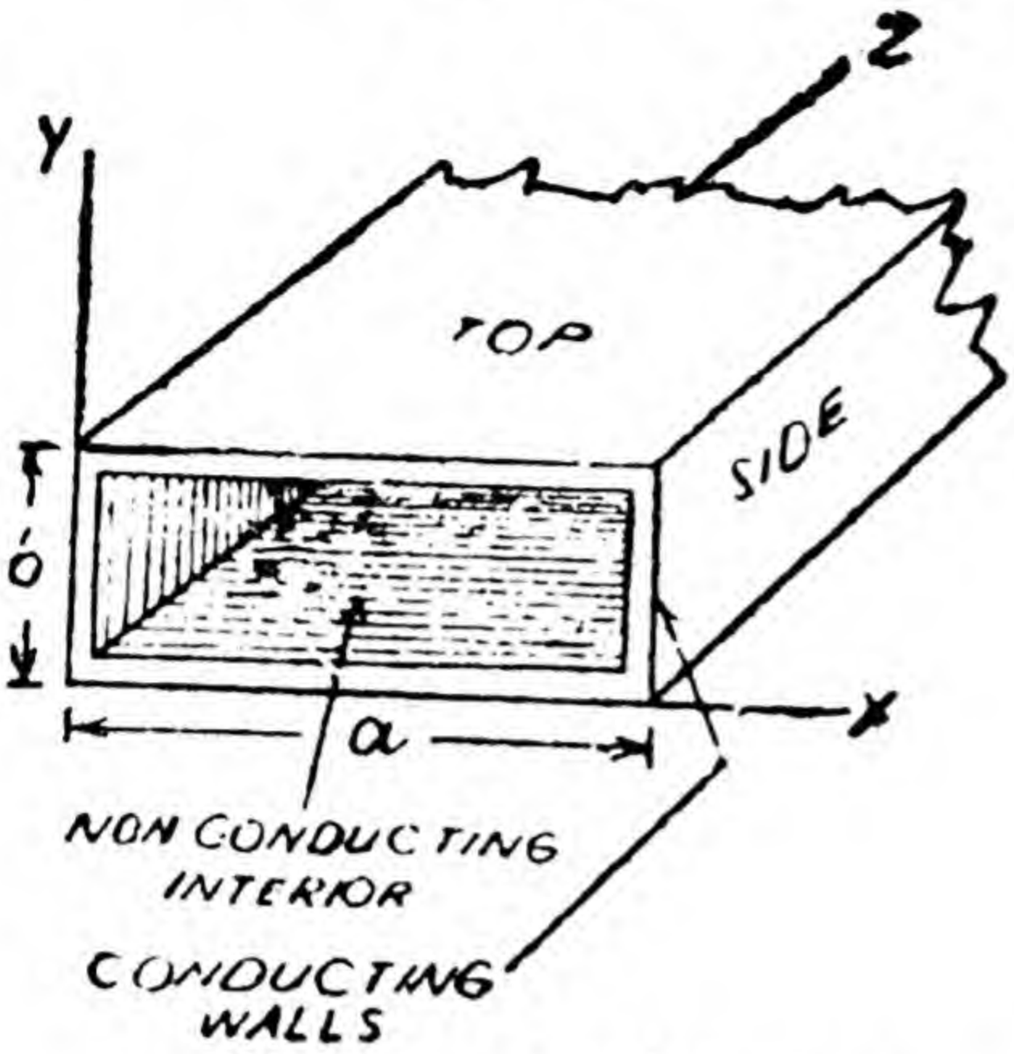
---

*The plot easily follows from $\lambda_0^2/\lambda_g^2 = 1 - \lambda_0^2/\lambda_o^2$.

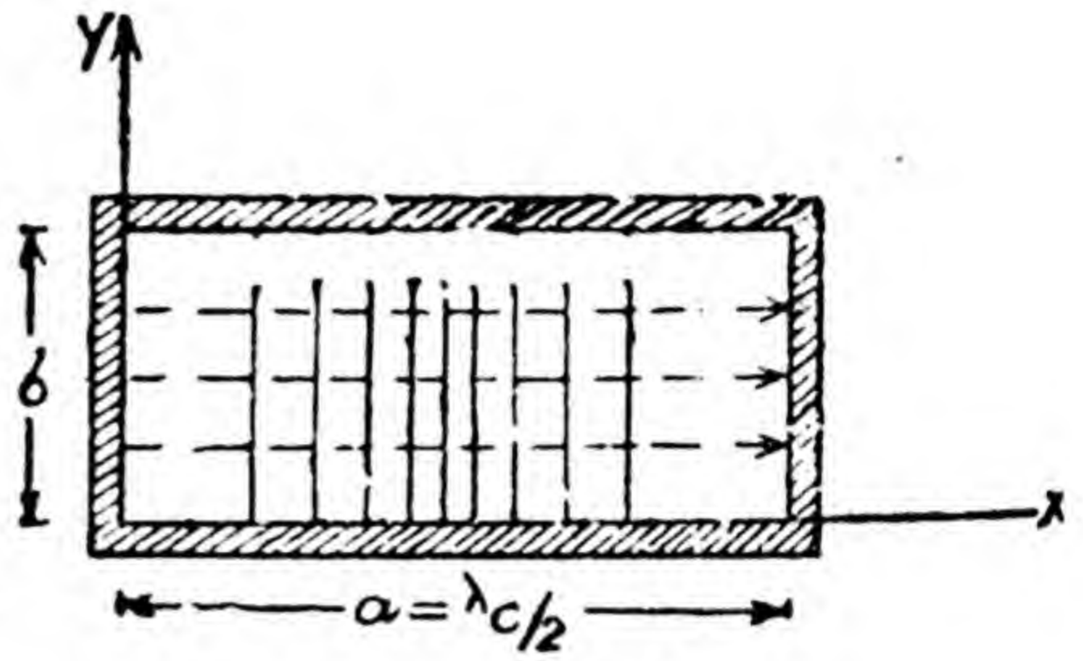
†We know $\omega_o^2/\omega^2 = \lambda_0^2/\lambda_o^2$. When frequency ω is high, we can assume that $\omega \to \infty$ so that $\omega_o^2/\omega^2 \to 0$ or $\lambda_0/\lambda_c \to 0$. But when frequency ω is low, it can assume as lowest a value as ω_c; because below ω_c propagation is not possible, so $\omega = \omega_c$ *i.e.* $\lambda_0/\lambda_c = 1$.

‡$\frac{\lambda_g}{\lambda_0} = \frac{1}{\sqrt{(1-\omega^2/\omega_c^2)}}$ so that as ω decreases, λ_g/λ_0 increases.

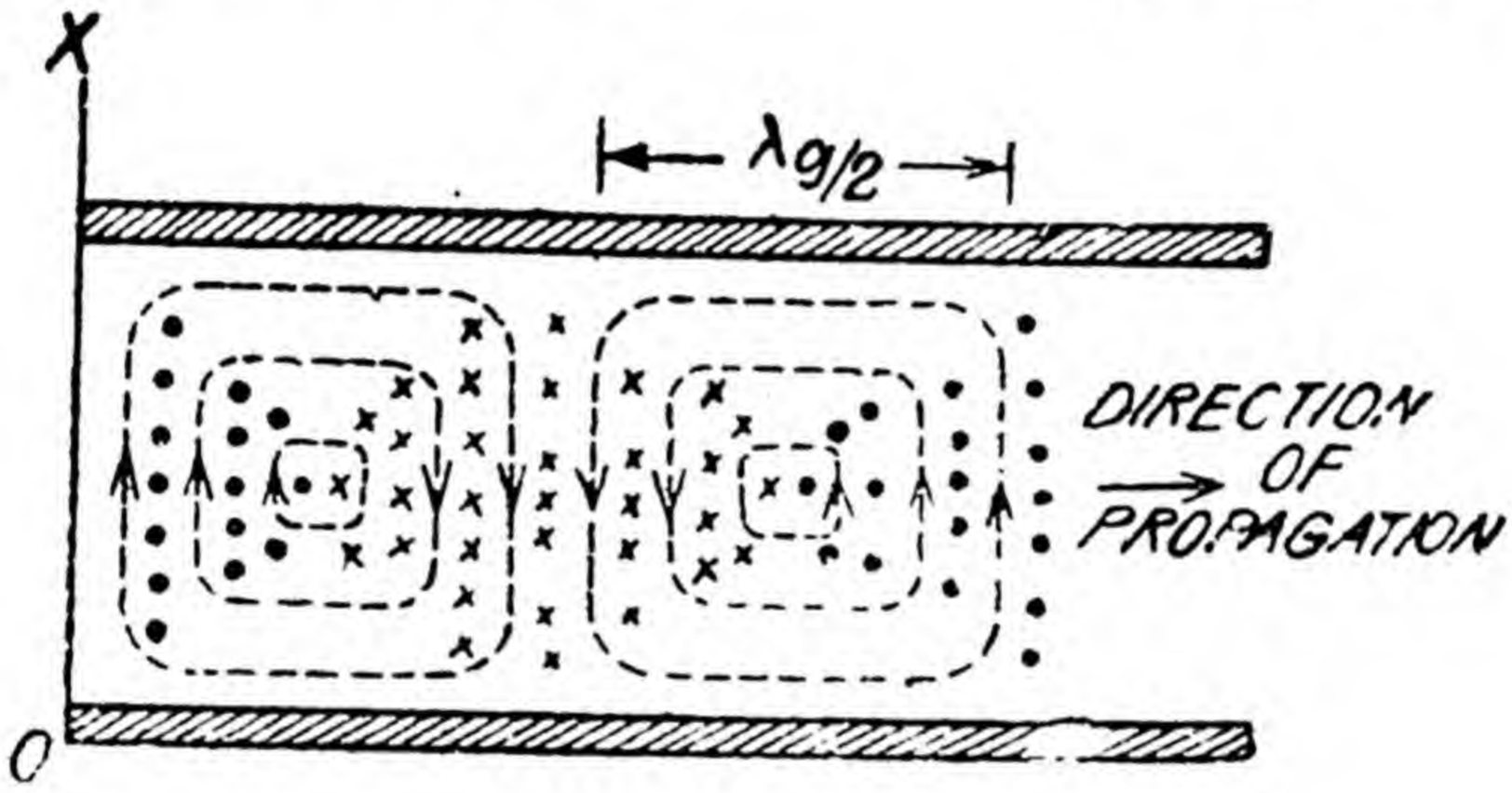
Field configuration for $TE_{1,0}$ mode in rectangular waveguide :



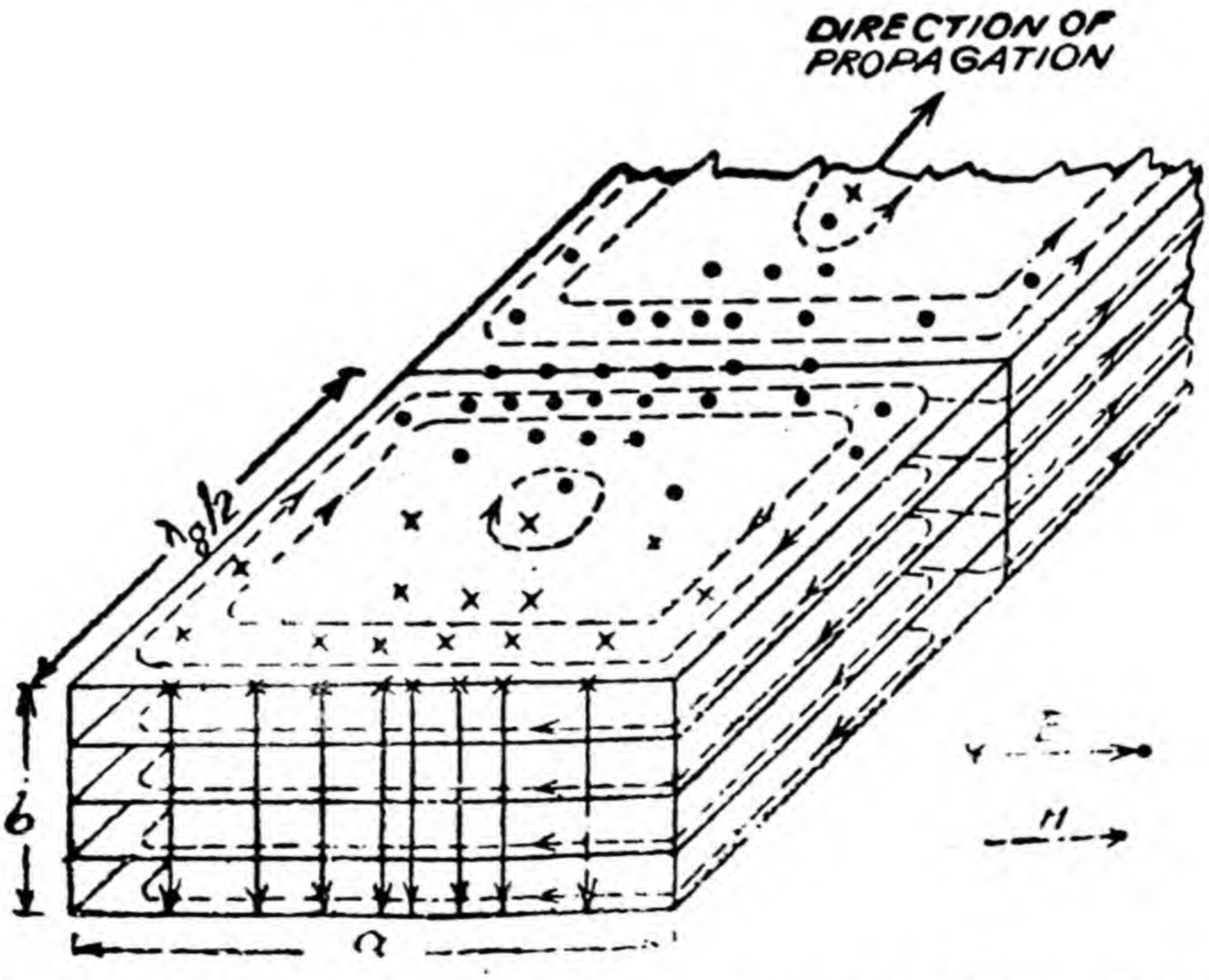
(*a*) Rectangular wave guide.



(*b*) End view



(*c*) Bottom view

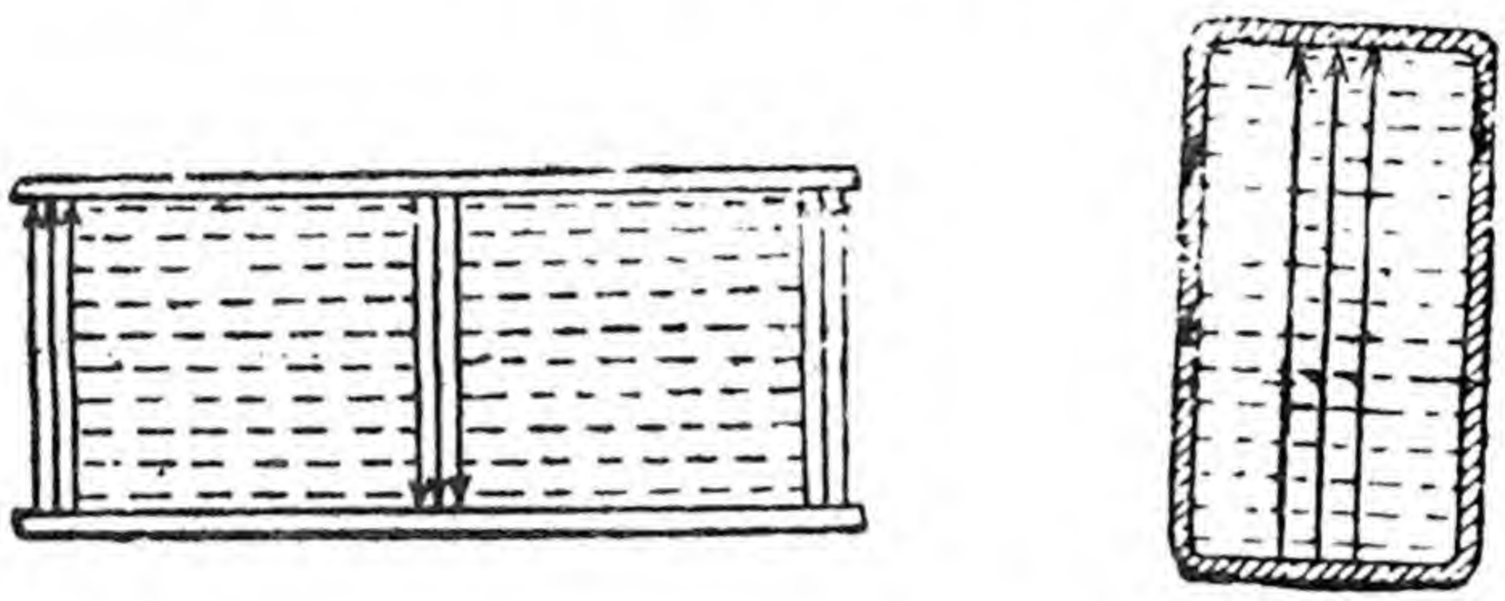


(*d*) Pictorial representation of field pattern of $TE_{1,0}$ wave.

Fig. 6. Electric and magnetic fields of $TE_{1,0}$ wave in rectangular waveguide. Refer to eq. (36). For $TE_{1,0}$ we have only E_y, H_x and H_z : rest other

of a cycle as shown with tail below and point head above. The magnetic field forms the closed loops.

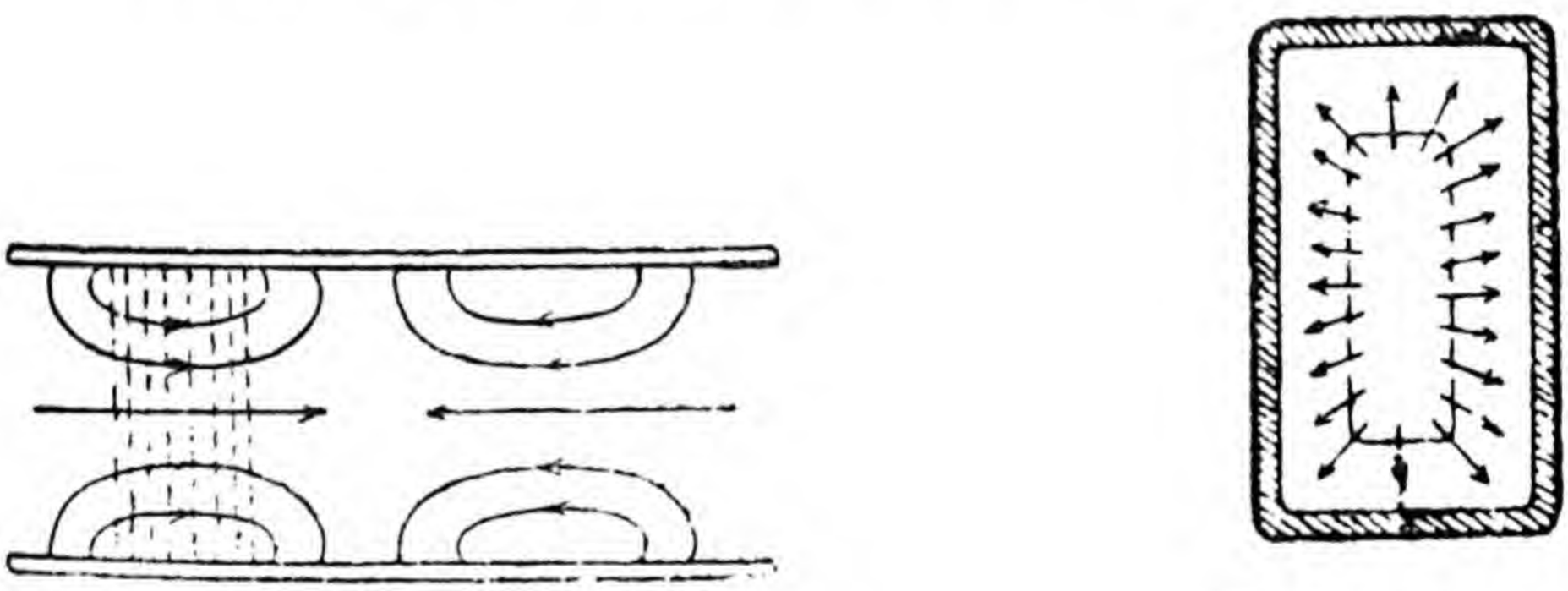
Field configuration for $TE_{0,1}$ mode in rectangular waveguide :



Top view. End view.

Fig. 7.

Field configuration for $TM_{1,1}$ mode :



Top view End view.

Fig. 8.

Example : *A hollow rectangular waveguide has waveguide dimensions 6 cm × 4 cm. The frequency of impressed signal is 3GHz. Compute, for the $TE_{1,0}$ mode :*

(a) cut-off wevelength
(b) the guide wavelength
(c) the phase constant
(d) the intrinsic wave impedance.

Given $a = 6$ cm. $= \cdot 06$ m.
$b = 4$ cm. $= \cdot 04$ m.

For $TE_{1,0}$ mode :

(*a*) Cutoff wavelength

$$\lambda_0 = \frac{v_0}{(f_c)_{1,0}} = 2a$$

$$= 2 \times \cdot 06 = 0 \cdot 12 \text{ m.}$$

(*b*) Phase constant

$$\beta = \sqrt{\left\{ \omega^2 \mu\epsilon - \left(\frac{\pi}{a}\right)^2 \right\}}$$

$$=\sqrt{\frac{4\pi^2 f^2}{c^2}-\left(\frac{\pi}{a}\right)^2}$$

$$=\pi\left[\frac{4\times(3\times10^9)^2}{(3\times10^8)^2}-\left(\frac{1}{\cdot06}\right)^2\right]^{1/2}$$

$$=\pi\left[400-\frac{10000}{36}\right]^{1/2}$$

$$=34\cdot73 \text{ rad/met.}$$

(*c*) Guide wavelength :

$$\lambda_g=\frac{2\pi}{\sqrt{\omega^2\mu\epsilon-\left(\frac{\pi}{a}\right)^2}}=\frac{2\pi}{\beta}$$

$$=\frac{2\pi}{34\cdot73}=0\cdot181 \text{ met.}$$

(*d*) Wave impedance :

$$Z_z=-\frac{E_y}{H_x}=\frac{\omega\mu}{\beta}=\frac{2\pi f\times4\pi\times10^{-7}}{34\cdot73}$$

$$=\frac{2\pi\times3\times10^9\times4\pi\times10^{-7}}{34\cdot73}$$

$$=682\cdot034 \text{ ohm.}$$

27·5. CIRCULAR WAVEGUIDES :

The method to solve the Maxwell's equation, in order to determine the field configuration in the circular guide, is quite similar to that adopted in rectangular guides. Wave equations and field equations for simplicity, are expressed in cylindrical co-ordinates and then boundary conditions are applied to obtain the final expression for field components.

TM and TE Waves in Circular Guides :

In order to determine the conditions for propagation of waves inside a hollow, perfectly conducting cylinder as shown in fig. 9, it is convenient to employ cylindrical co ordinates r, ϕ and z instead of rectangular co ordinates. The Maxwell's equations may be converted into cylindrical co ordinates.

In cylinderical co-ordinates in a non-conducting region, assuming like art. 27·1, that variation of field components in z direction is represented by a factor $e^{-\gamma z}$, Maxwell's equations are

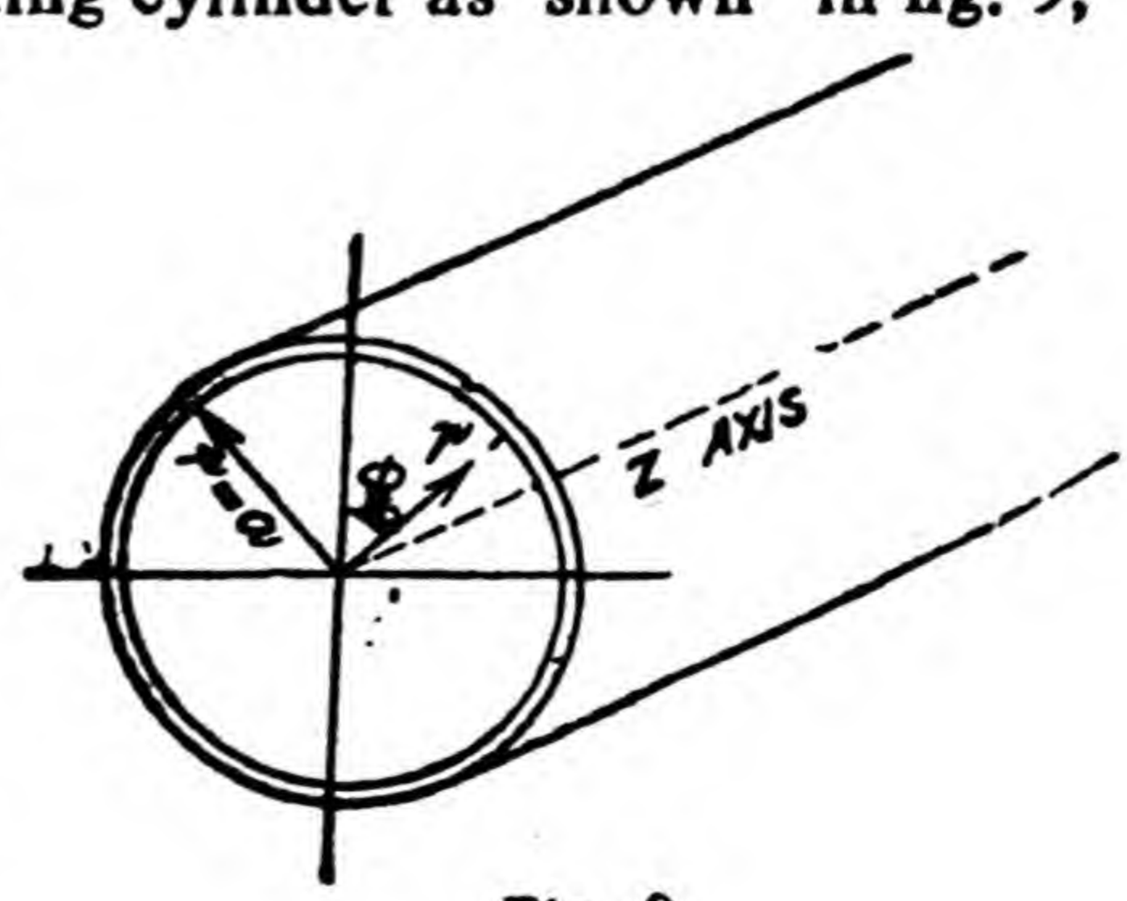


Fig. 9.

$$\left.\begin{aligned}
&\frac{1}{r}\cdot\frac{\partial H_z}{\partial\phi}+\gamma H_\phi=j\omega\epsilon E_r\\
&-\gamma H_r-\frac{\partial H_z}{\partial r}=j\omega\epsilon E_\phi\\
&\frac{1}{r}\left[\frac{\partial(rH_\phi)}{\partial r}-\frac{\partial H_r}{\partial\phi}\right]=j\omega\epsilon E_z,
\end{aligned}\right\} \quad \ldots(1)$$

$$\left.\begin{aligned}
&\frac{1}{r}\cdot\frac{\partial E_z}{\partial\phi}+\gamma E_\phi=-j\omega\mu H_r\\
&-\gamma E_r-\frac{\partial E_z}{\partial r}=-j\omega\mu H_\phi,\\
&\frac{1}{r}\left[\frac{\partial(rE_\phi)}{\partial r}-\frac{\partial E_r}{\partial\phi}\right]=-j\omega\mu H_z,
\end{aligned}\right\}$$

The equations can be combined to give

$$\left.\begin{aligned}
h^2H_r&=j\frac{j\omega\epsilon}{r}\frac{\partial E_z}{\partial\phi}-\gamma\frac{\partial H_z}{\partial r},\\
h^2H_\phi&=-j\omega\epsilon\frac{\partial E_z}{\partial r}-\frac{\gamma}{r}\frac{\partial H}{\partial\phi},\\
h^2E_r&=-\gamma\frac{\partial E_z}{\partial r}-\frac{j\omega\mu}{r}\frac{\partial H_z}{\partial\phi}\\
h^2E_\phi&=-\frac{\gamma}{r}\frac{\partial E_z}{\partial\phi}+j\omega\mu\frac{\partial H_z}{\partial r}
\end{aligned}\right\} \quad \ldots(2)$$

where $$h^2=\gamma^2+\omega^2\mu\epsilon.$$

The wave equation in cylindrical co-ordinates for E_z is

$$\frac{\partial^2E_z}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2E_z}{\partial\phi^2}+\frac{\partial^2E_z}{\partial z^2}+\frac{1}{r}\frac{\partial E_z}{\partial r}=-\omega^2\mu\epsilon E_z.$$

Let $$E_z=P(r)\,Q(\phi)\,e^{-\gamma z}=E_z^0\,e^{-\gamma z},$$

which when substituted in above wave equation gives

$$Q\frac{d^2P}{dr^2}+\frac{Q}{r}\frac{dP}{dr}+\frac{P}{r^2}\frac{d^2Q}{d\phi^2}+PQ\gamma^2+\omega^2\mu\epsilon PQ=0$$

or $$\frac{1}{P}\frac{d^2P}{dr^2}+\frac{1}{rP}\frac{dP}{dr}+\frac{1}{Qr^2}\frac{d^2Q}{d\phi^2}+h^2=0$$

or $$r^2\left[\frac{1}{P}\frac{d^2P}{dr^2}+\frac{1}{rP}\frac{dP}{dr}+h^2\right]=-\frac{1}{Q}\frac{d^2Q}{d\phi^2}=n^2,$$

where n is a constant. Thus two equations are formed, *viz.*

$$\frac{d^2Q}{d\phi^2}=-n^2Q,$$

the solution of which is

$$Q = A_n \cos n\phi + B_n \sin n\phi,$$

and equation

$$\frac{d^2P}{dr^2} + \frac{1}{r}\frac{dP}{dr} + \left(h^2 - \frac{n^2}{r^2}\right)P = 0.$$

Dividing this equation by h^2, we find

$$\frac{d^2P}{d(rh)^2} + \frac{1}{rh}\frac{dP}{d(rh)} + \left(1 - \frac{n^2}{r^2h^2}\right)P = 0,$$

which is Bessel equation in terms of (rh). Using only the solution that is finite at $(rh)=0$, we write the solution

$$P\,(rh) = J_n\,(rh),$$

where $J_n\,(rh)$ is Bessel function of first kind of order n.

Therefore, we find that

$$E_z^0 = P\,(r)\;Q\,(\phi)$$

$$= J_n\,(rh)\;(A_n \cos n\phi + B_n \sin n\phi). \qquad \ldots(3)$$

The solution of H_z will have exactly the same form as that of E_z and can, therefore, be written as

$$H_z^0 = J_n\,(rh)\;(C_n \cos n\phi + D_n \sin n\phi). \qquad \ldots(4)$$

The variations of E_z and H_z in the z-direction and with time are accounted if equations (3) and (4) are multiplied by a factor $e^{(j\omega t - \gamma z)}$. But as we have discussed previously in the case of rectangular guides, γ should be imaginary for transmission to occur, *i.e.*, $\gamma = j\beta$. Thus for waveguide case, we should multiply by $e^{j(\omega t - \beta z)}$. Therefore,

$$E_z = J_n\,(rh)\;(A_n \cos n\phi + B_n \sin n\phi)\; e^{j(\omega t - \beta z)}$$

and

$$H_z = J_n\,(rh)\;(C_n \cos n\phi + D_n \sin n\phi)\; e^{j(\omega t - \beta z)}.$$

Further, the relative amplitudes of A_n and B_n (also of C_n and D_n) determine the orientation of the field in the guide, and for a circular guide and for any particular value of n, the $\phi = 0$ axis can always be oriented to make either A_n or B_n equal to zero. Let us choose B_n (also D_n) equal to zero, so that

$$E_z = J_n\;rh)\;A_n \cos n\phi\; e^{j(\omega t - \beta z)} \qquad \ldots(5)$$

and

$$H_z = J_n\,(rh)\;C_n \cos n\phi . e^{j(\omega t - \beta z)}. \qquad \ldots(6)$$

27·5-1. TM WAVES IN CYLINDRICAL GUIDES :

For *TM* waves, H_z is identically zero and the wave equation for E_z is used. Thus from equations (2), and using equation (5), taking only the real part, we write expressions of fields for *TM* waves as follows :

$$E_z = A_n J_n(rh) \cos n\phi \cos(\omega t - \beta z)$$
$$H_z = \frac{A_n \omega \epsilon n}{h^2 r} J_n(rh) \sin n\phi \sin(\omega t - \beta z)$$
$$H_\phi = \frac{A_n \omega \epsilon}{h} J_n'(rh) \cos n\phi \sin(\omega t - \beta z) \quad \ldots(7)$$
$$E_r = \frac{\beta}{\omega \epsilon} H_\phi$$
$$E_\phi = -\frac{\beta}{\omega \epsilon} H_r$$
$$H_z = 0.$$

For *TM* wave it is necessary that E_z and E_ϕ be zero at the pipe wall where $r=a$, because of infinite conductivity of the wall. This requirement, from equation (5), gives

$$J_n(ha) = 0, \quad \ldots(8)$$

where a is the radius of the guide. Equation (8) possesses an infinite number of roots but since for the propagation to be possible γ is to be imaginary, *i.e.*, h should be small ($h^2 \ll \omega^2 \mu \epsilon$) or else extremely high frequencies will be required, only first few roots of eqn. (8) will be of practical interest. Few roots of J_n are given in the following table :

TABLE

Roots of $J_n = 0$

m	$n=0$	$n=1$	$n=2$	$n=3$
1	2·40	3·83	5·14	6·38
2	5·52	7 02	8·42	9 76
3	8·65	10 17	11·62	13·02
4	11·79	13·32	14·80	16·22

In the table m signifies the particular root meant. Thus

$$\left.\begin{aligned} (ha)_{01} &= 2{\cdot}40 \quad (ha)_{11} = 3{\cdot}83 \\ (ha)_{02} &= 5{\cdot}52 \quad (ha)_{12} = 7{\cdot}02 \end{aligned}\right\} \quad \ldots(9)$$

The first subscript refers to the value of n and the second refers to m *i.e.*, to the particular root under consideration. The various *TM* waves will be referred to as TM_{01}, TM_{12}, etc., *i.e.*, in general TM_{nm}. It is obvious that since there is no root $(ha)_{00}$, TM_{00} waves cannot exist.

Physically n is the number of cycles of variation of E_r found as ϕ varies around the complete cylinder or through 2π radians. The subscript m indicates the number of zeros of electric field along a radial path from the centre to the inside surface of the outer guide wall.

Let us represent relations (9) as follows in a general form

$$(ha)_{nm}=\tau_{nm},$$

where τ_{nm} is a root of $J_n=0$. Since $h=\sqrt{(\gamma^2+\omega^2\mu\epsilon)}$, we have

$$a\sqrt{(\gamma^2+\omega^2\mu\epsilon)}=\tau_{nm},$$

so that

$$\gamma_{nm}=\sqrt{\left\{\left(\frac{\tau_{nm}}{a}\right)^2-\omega^2\mu\epsilon\right\}}$$

For propagation to occur, γ_{nm} should be imaginary, *i.e.*,

$$\gamma_{nm}=j\beta_{nm}=j\sqrt{\left\{\omega^2\mu\epsilon-\left(\frac{\tau_{nm}}{a}\right)^2\right\}}.$$

The cut off frequency, above which transmission takes place, can be calculated by putting $\beta_{nm}=0$, *i.e.*

$$\therefore \quad \omega_c^2\mu\epsilon=\left(\frac{\tau_{nm}}{a}\right)^2$$

$$\therefore \quad f_c=\frac{1}{2\pi\sqrt{(\mu\epsilon)}}\cdot\frac{\tau_{nm}}{a}=\frac{\tau_{nm}\,v_0}{2\pi a}. \qquad ...(10)$$

The cutoff wavelength is

$$\lambda_c=\frac{v_0}{f_0}=\frac{2\pi a}{\tau_{nm}}.$$

Since $2\pi a$ is the inner circumference of the outer wall of the guide, the cut-off wavelength is that at which the circumference is equal to $\tau_{n,\,m}$ wavelengths.

The phase velocity is

$$v=\frac{\omega}{\beta_{nm}}=\frac{\omega}{\sqrt{(\omega^2\mu\epsilon-\omega_0^2\mu\epsilon)}}=\frac{v}{\sqrt{(1-\omega_0^2/\omega^2)}}$$

$$=\frac{v_0}{\sqrt{(1-f_c^2/f^2)}} \qquad ...(11)$$

and the group velocity is

$$v_g=\frac{1}{\dfrac{d\beta_{n,\,m}}{d\omega}}=v_0\sqrt{\left(1-\frac{f_0^2}{f^2}\right)} \qquad ...(12)$$

27·5-2. TE WAVES IN CYLINDRICAL GUIDES :

E_z is zero. Using equations (2) along with equation (6), taking

only the real part, we write the expression of fields for *TE* waves as follows :

$$\left.\begin{aligned} H_z &= C_n J_n (rh) \cos n\phi \cos (\omega t - \beta z) \\ H_r &= \frac{\beta C_n}{h} J_n' (rh) \cos n\phi \sin (\omega t - \beta z) \\ H_\phi &= -\frac{n\beta C_n}{h^2 r} J_n (rh) \sin n\phi \sin (\omega t - \beta z) \\ E_r &= \frac{\omega\mu}{\beta} H_\phi \\ E_\phi &= -\frac{\omega\mu}{\beta} H_r \\ E_z &= 0. \end{aligned}\right\} \quad \ldots(13)$$

It is known that at $r=a$, the H_r or normal field value must be zero which, as we can infer from equation (13), means

$$J'_n (ha) = 0.$$

The derivative J'_n with respect to r of the Bessel function may be obtained from

$$\frac{\partial J_n (hr)}{\partial r} = J'_n (hr) = \frac{n}{hr} J_n (hr) - J_{n+1} (hr)$$

and a few roots of $J_n' = 0$ are given in the following table.

TABLE

Roots $J_n' = 0$

m	$n=0$	$n=1$	$n=2$
1	3·83	1·84	3·05
2	7·02	5·33	6·71
5	10·17	8·54	9·97

The corresponding *TE* waves are referred to as $TE_{0,1}$; TE_{11} and so on.

The equations for β_{nm}, f_c, λ_c, v_p and v_g are the same as those obtained for *TM* mode except for the use of τ'_{nm} in place of τ_{nm}.

FIELD CONFIGURATIONS FOR TE_{nm} AND TM_{nm} MODES :

We shall show field configuration for $TE_{0,1}$ mode. Since $E_z = 0$. electric field will be completely normal to the axis of tube *i.e.*, along the cross section.

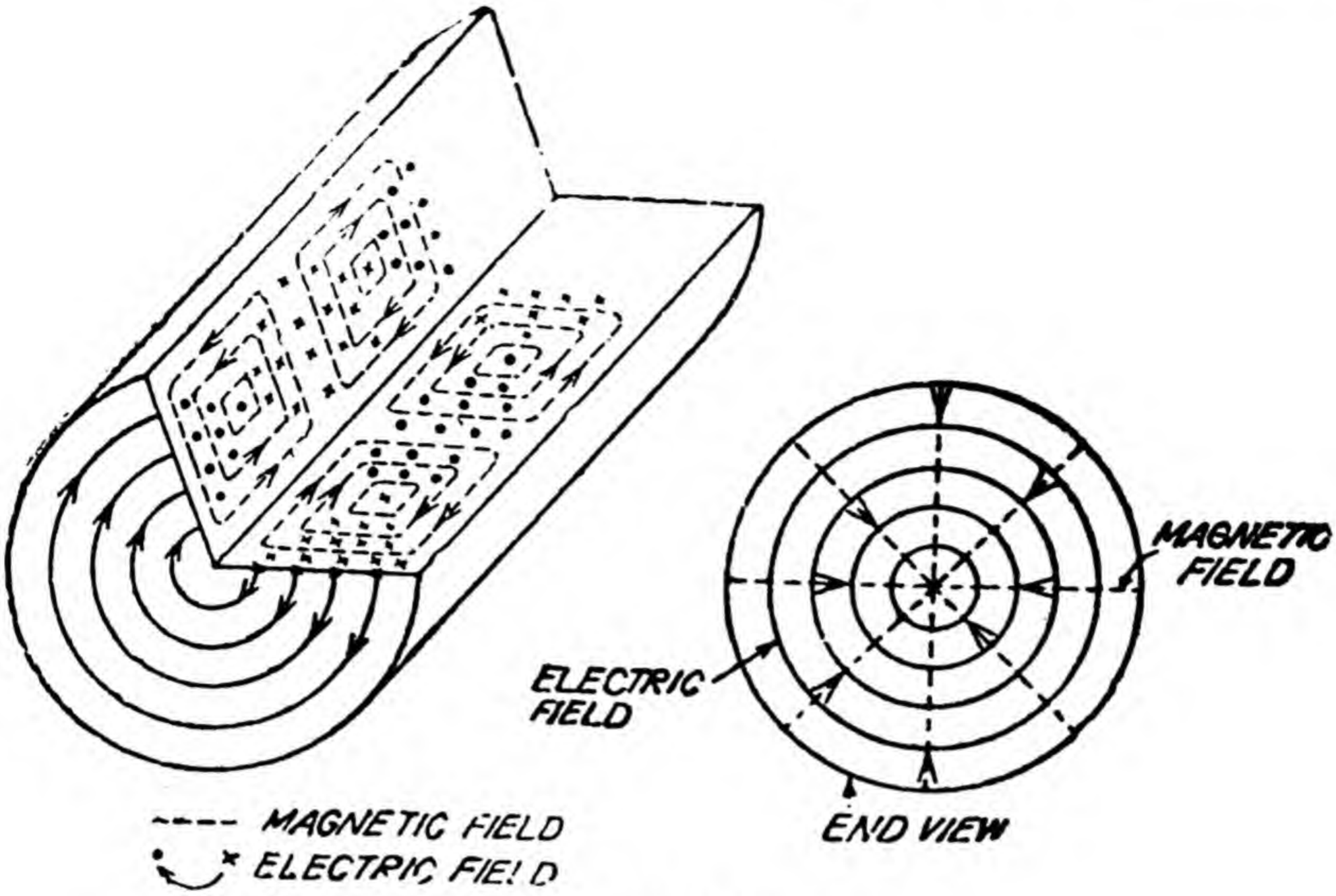


Fig. 10. TE_{10} wave field configuration.

27·6. IMPOSSIBILITY OF TEM WAVES IN WAVEGUIDES (Single Conductor) :

We know that *TE* and *TM* waves propagate within hollow metal pipes of rectangular or circular cross section. Each of these waves has an axial component of the field, *i.e.*, *TE* has H_z while *TM* has E_z component of fields along the axis—the direction of propagation. *TEM* wave, for which there is no axial component either *E* or *H*, can not possibly propagate within a single conductor waveguide for the reasons explained as follows :

Suppose a *TEM* wave exists within a waveguide. The lines of magnetic field *H* must lie in a plane perpendicular to the direction of propagation *i.e.*, in the transverse plane. Also in a non magnetic material,

$$\nabla.\mathbf{H}=0,$$

which requires that the lines of *H* be closed loops. It means, if a *TEM* wave exists inside the guide, the lines of *H* will be closed loops in a plane perpendicular to the axis. Further from Maxwell's m.m.f. relation

$$\oint \mathbf{H}.\mathbf{dl}=\int (j\omega\mathbf{D}+\mathbf{J})\,\mathbf{dS},$$

it is obvious that the magnetomotive force around each of these closed loops must equal the axial current (conduction or displacement) through the loop. In the case of a "guide" with an inner conductor, *e.g.*, a coaxial transmission line, the axial current through *H* loops is the conduction current in the inner conductor. However, for a hollow waveguide having no inner conductor, this axial current must be a displacement current (current arising due

to dielectric polarisation). But an axial displacement current requires an axial component of E, ($D_z=\epsilon E_z$) which is not possible in *TEM* waves and *hence TEM wave cannot exist in a single conductor*.

Characteristics of TEM waves :

In order to discuss *TEM* waves, it is thus obvious that one should consider a 'guide' with two conductors. For the present discussion, let us choose two parallel conducting planes. We can easily obtain by solving Maxwell's field equations for such a 'guide', the field components of E and H for *TE* as well as for *TM* waves as stated below* :

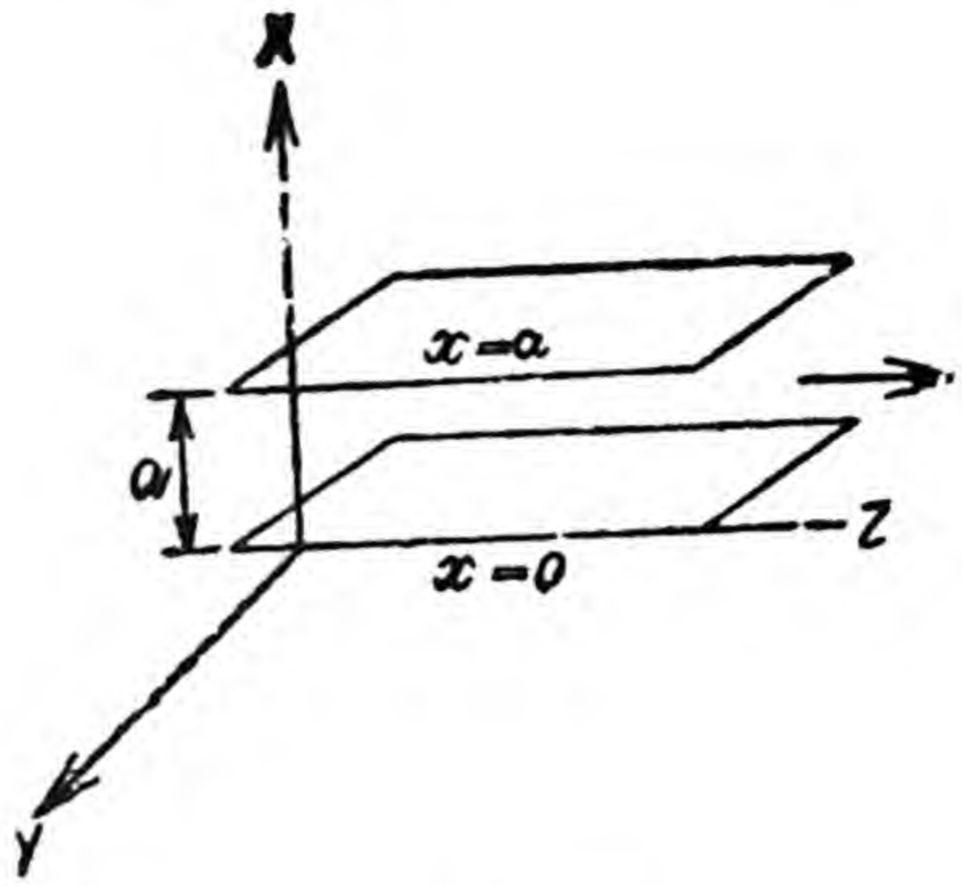


Fig. 11.

TE waves

$$E_y=C_1 \sin\left(\frac{m\pi x}{h}\right) e^{-\gamma z},$$

$$H_z=-\frac{m\pi}{j\omega\mu a} C_1 \cos\left(\frac{m\pi x}{a}\right) e^{-\gamma z},$$

$$H_x=-\frac{\gamma}{j\omega\mu} C_1 \sin\left(\frac{m\pi x}{a}\right) e^{-\gamma z},$$

TM waves

$$H_y=C_4 \cos\left(\frac{m\pi x}{a}\right) e^{-\gamma z},$$

$$E_x=\frac{\gamma C_4}{j\omega\epsilon} \cos\left(\frac{m\pi x}{a}\right) e^{-\gamma z}$$

$$E_z=\frac{jm\pi}{\omega\epsilon a} C_4 \sin\left(\frac{m\pi x}{a}\right) e^{-\gamma z}.$$

We should, in above equations, put $\gamma=j\beta$ for the range of frequencies in which wave propagation occurs. We note from above field equations that from $m=0$, all field components for *TE* wave become zero and therefore the lowest order *TE* wave is $TE_{1,0}$ wave. But we observe that for $m=0$, *TM* wave field components are

$$\left.\begin{aligned} H_y&=C_4 e^{-j\beta z},\\ E_x&=\frac{\beta}{\omega\epsilon} C_4 e^{-j\beta z},\\ E_z&=0 \end{aligned}\right]$$

and thus all the field components are not zero and TM_{00} wave is possible. We note that when $m=0$, $E_z=0$ and also for *TM* wave $H_z=0$ and hence with $m=0$, we obtain condition for such a wave

---

*For two parallel plane plates of infinite dimension in y direction, the direction of propagation being z, we put $\partial/\partial y=0$, so that the boundary condition $E_{tangential}=0$ at $x=0$. $x=a$ will be applied to

$$\frac{1}{X}\frac{d^2X}{dx^2}+h^2=0$$

In solution only one constant, h, ($=m\pi/a$) will occur.

which do not consist H_z, as well as E_z *i.e.*, electromagnetic field is entirely transverse. Such a wave is called transverse electromagnetic wave (*TEM* wave). It is a familiar type of wave propagated along all ordinary two conductor transmission lines when operating in their low frequency (customary) manner. It is usually called *a principal wave*.

We note in the case of *TM* or *TE* waves for propagation to be possible

$$\gamma=j\beta=j\sqrt{\left\{\omega^2\mu\epsilon-\left(\frac{m\pi}{a}\right)^2\right\}}$$

or

$$\beta=\sqrt{\left\{\omega^2\mu\epsilon-\left(\frac{m\pi}{a}\right)^2\right\}}.$$

For some value of ω, β may become zero. Let this be ω_c, then

$$\omega_c^2\mu\epsilon=\sqrt{\left(\frac{m\pi}{a}\right)^2}$$

or

$$f_o=\frac{m}{2a\sqrt{(\mu\epsilon)}}.$$

For frequencies greater than f_c, β will have some value making the propagation possible, f_o is called cutoff frequency. It depends upon a, the separation between the planes and thus propagation of *TE* and *TM* wave *requires a condition that is imposed on the separation between the planes.* For *TEM* waves, we put $m=0$ therefore

$$f_o=0,$$

which means that the cutoff frequencies for *TEM* wave is zero, *i.e.* for *TEM* waves, all frequencies down to zero can propagate along the guide (and that is why on a two conductor transmission line at which *TEM* wave exists, propagation at a low frequencies is possible). Further *TEM* wave propagation does not demand any condition to be imposed upon the separation between the plates, *i.e.*, we can reduce the distance between the conductors to a certain minimum.

In case of *TE* and *TM* waves, velocity of propagation along the 'guide' is

$$(v_g)_z=\frac{\omega}{\beta}=-\frac{\omega}{\sqrt{\left\{\omega^2\mu\epsilon-\left(\frac{m\pi}{a}\right)^2\right\}}}$$

$$=\frac{\omega}{\sqrt{(\omega^2\mu\epsilon-\omega_c^2\mu\epsilon)}}$$

$$=\frac{1}{\sqrt{(\mu\epsilon)}\sqrt{(1-f_o^2/f^2)}}$$

$$=\frac{v_0}{\sqrt{(1-f_c^2/f^2)}}$$

At frequencies $f > f_c$, at which propagation is possible, wave velocity within guide is greater than v_0, the free space velocity. For *TEM* waves $f_c=0$ and therefore wave velocity is equal to free space velocity. It means, unlike *TE* and *TM* waves, velocity of *TEM* wave is independent of frequency.

27·7. TRANSMISSION LINE ANALOGY FOR WAVEGUIDES :

It can be easily shown that electric and magnetic fields of *TE* and *TM* waves bear an analogy with those of voltage and current on a suitably loaded transmission line. The analogy helps in drawing 'equivalent circuits' for the waveguides. The concept of a waveguide as an equivalent transmission line with a certain characteristic impedance and propagation constant helps a lot in solving several waveguide problems because with such 'equivalent circuit plot', one can obtain the solution by means of well-known circuit transmission line theory.

EQUIVALENT CIRCUIT FOR TM WAVES :

For *TM* wave, we should substitute $H_z=0$ in the field equations :

$$\left.\begin{aligned} \nabla \times \mathbf{H} &= j\omega\epsilon\mathbf{E} \quad ...(a) \\ \text{and} \quad \nabla \times \mathbf{E} &= -j\omega\mu\mathbf{H} \quad ...(b) \end{aligned}\right\} \quad ...(1)$$

Since $H_z=0$,

$$(\nabla \times \mathbf{E})_z = 0$$

That is, curl of elcctric field in x, y plane vanishes and therefore E is conservative in this plane. We can then represent E_x and E_y as a gradient of some scalar potential V, *i.e*,

$$\left.\begin{aligned} E_x &= -\frac{\partial V}{\partial x} \\ E_y &= -\frac{\partial V}{\partial y} \end{aligned}\right\} \quad ...(2)$$

From equation (3) of art. 27·4 we know that, when H_z is zero,

$$H_y = -\frac{j\omega\epsilon}{h^2}\frac{\partial E_z}{\partial x} \quad .. (3)$$

(1) Current equation : Writing x components of equation 1 (*a*) with $H_z=0$, we get

$$\frac{\partial H_y}{\partial z} = -j\omega\epsilon E_x.$$

Putting the value of E_x and H_y from equations (2) and (3) in above equations, we get

$$\frac{\partial}{\partial z}\left(\frac{j\omega\epsilon}{h^2}\cdot\frac{\partial E_z}{\partial x}\right) = -j\omega\epsilon\frac{\partial V}{\partial x}$$

or

$$\frac{\partial}{\partial z}\left(\frac{j\omega\epsilon}{h^2}E_z\right) = -j\omega\epsilon V.$$

The quantity $j\omega\epsilon E_z$ (as $D_z=\epsilon E_z$) is the longitudinal displace-

ment current density, and $1/h^2$ has the dimensions of an area* so $\frac{j\omega\epsilon E_z}{h^2}$ represents a current in z direction and, therefore, above equation is

$$\frac{\partial I_z}{\partial z} = -j\omega\epsilon V. \qquad \text{.. (4)}$$

(ii) Voltage equation : Writing y component of equation 1 (b), we get

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -j\omega\mu H_y.$$

On putting the value of H_y from equation (3), we get

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -\frac{\omega^2\mu\epsilon}{h^2}\frac{\partial E_z}{\partial x}$$

or
$$\frac{\partial E_x}{\partial z} = -\left(\frac{\omega^2\mu\epsilon}{h^2} - 1\right)\frac{\partial E_z}{\partial x}.$$

Further $E_x = -\frac{\partial V}{\partial x}$ so that

$$\frac{\partial}{\partial z}\left(-\frac{\partial V}{\partial x}\right) = -\left(\frac{\omega^2\mu\epsilon}{h^2} - 1\right)\frac{\partial E_z}{\partial x}$$

or
$$\frac{\partial V}{\partial z} = -\left(\frac{\omega^2\mu\epsilon}{h^2} - 1\right)E_z$$

$$= -\left(j\omega\mu + \frac{h^2}{j\omega\mu}\right)\left(\frac{j\omega\epsilon}{h^2}\ E_z\right)$$

or
$$\frac{\partial V}{\partial z} = -\left(j\omega\mu + \frac{h^2}{i\omega\epsilon}\right)I_z. \qquad \text{..(5)}$$

by the arrangement of equation (4).

The transmission line equations taking the direction of propagation as z, are

$$\left.\begin{aligned}\frac{\partial V}{\partial z} &= -(R + j\omega L)\ I = -ZI \\ \text{and}\qquad \frac{\partial I}{\partial z} &= -(G + j\omega C)\ V = -YV,\end{aligned}\right\} \qquad \text{...(6)}$$

the negative sign signifies the loss of voltage and current. Comparing equations (4) and (5) with equation (6), we find that the series impedance per unit length of the 'equivalent line' is

$$Z = j\omega\mu + h^2/j\omega\epsilon, \qquad \text{...(7)}$$

and shunt admittance per unit length of 'equivalent line' is

$$Y = j\omega\epsilon. \qquad \text{...(8)}$$

---

*$h^2 = B^2 + A^2 = \left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2$, unit is thus cm^{-2} so that $1/h^2$ will have cm^2 as unit.

Therefore the equivalent transmission line circuit* can be shown as in the figure 12.

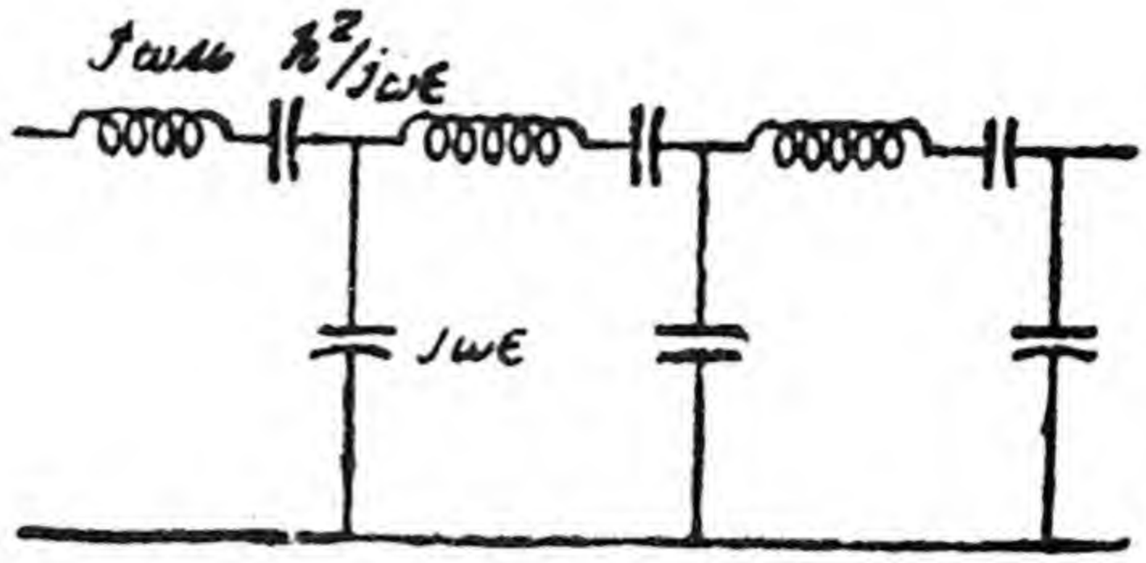


Fig. 12. Equivalent transmission line circuit representation for *TM* waves

The cutoff frequency of this 'equivalent circuit' occurs when the series reactance equals zero. *i.e.* $j\omega\mu + h^2/j\omega\epsilon = 0$, *i.e.*, when $h^2 = \omega^2\mu\epsilon$ which is same derived in the theory of *TM* waves.

The characteristic impedance of this 'equivalent transmission line' will be

$$Z_0\,(TM) = \sqrt{\left(\frac{Z}{Y}\right)} = \sqrt{\left(\frac{j\omega\mu + h^2/j\omega\epsilon}{i\omega\epsilon}\right)}$$

$$= \sqrt{\left(\frac{\mu}{\epsilon}\right)}\sqrt{\left(1 - \frac{\omega_0^2}{\omega^2}\right)}, \qquad \ldots (9)$$

which is equal to the wave impedance of waveguide in *TM* mode. This can be shown as follows :

The wave impedance which is seen when looking in the direction of propagation, that is, along z axis is denoted by Z_z. From the transverse field components of a *TM* waves in a rectangular waveguide, we find that

$$\frac{E_x}{H_y} = -\frac{E_y}{H_x} = \frac{\sqrt{(E_x^2 + E_y^2)}}{\sqrt{(H_x^2 + H_y^2)}} = \frac{|E_{trans.}|}{|H_{trans.}|} = \frac{\beta}{\omega\epsilon},$$

on using eq. (13) of art. 27·4.

Therefore,

$$Z_{xy} = Z_{yx} = \frac{\beta}{\omega\epsilon} = Z_z,$$

i e., the wave impedances looking in the z direction are equal and may be put equal to Z_z, where

$$Z_z = \frac{E_{trans.}}{H_{trans.}}.$$

Now we know that for *TM* case in rectangular waveguide,

$$\beta = (\omega^2\mu\epsilon - h^2)$$

$$= \sqrt{(\omega^2\mu\epsilon - \omega_0^2\mu\epsilon)}$$

$$\therefore \quad Z_z = \frac{\sqrt{(\omega^2\mu\epsilon - \omega_0^2\mu\epsilon)}}{\omega\epsilon} = \sqrt{(\mu/\epsilon)}\sqrt{(1 - \omega_c^2/\omega^2)}.$$

---

*ϵ bears a relation to capacitance and μ to inductance Refer to Art. 26 9 eq. (5) and (8) $C = \epsilon b/a$ and $L = \mu a/b$.

Comparing it with equation (9), we find that

$$Z_0\ (TM)=Z_z$$

as has been stated earlier.

EQUIVALENT CIRCUIT FOR TE WAVES :

For *TE* waves, we should substitute $E_z=0$ in the field equations :

$$\left.\begin{aligned}\nabla\times\mathbf{H}&=j\omega\epsilon\mathbf{E} \quad (a)\\ \nabla\times\mathbf{E}&=-j\omega\mu\mathbf{H} \quad (b)\end{aligned}\right\} \qquad ...(10)$$

and Since $E_z=0$

$$(\nabla\times\mathbf{H})_z=0.$$

That is, curl of magnetic field in xy plane vanishes and therefore **H** is conservative in this plane. We can then represent H_x and H_y as a gradient of scalar magnetic potential ϕ. *i.e.*,

$$\left.\begin{aligned}H_x&=-\frac{\partial\phi}{\partial x}\\ H_y&=-\frac{\partial\phi}{\partial y}\end{aligned}\right] \qquad ...(11)$$

Further, from eq. (3) of art. 27·4, with $E_z=0$, we write

$$E_x=-\frac{j\omega\mu}{h^2}\frac{\partial H_z}{\partial y} \qquad ...(12)$$

Writing x component of eq. 10 (*a*) and y component of eq. 10 (*b*), we get

$$\frac{\partial H_z}{\partial y}-\frac{\partial H_y}{\partial z}=j\omega\epsilon E_x \qquad ...13\ (a)$$

and

$$\frac{\partial E_x}{\partial z}-0=-j\omega\mu H_y. \qquad ...13\ (b)$$

(*i*) Current equation :

Let us put eqs. (11) and (12) in eq. 13 (*a*) to get

$$\frac{\partial H_z}{\partial y}-\frac{\partial}{\partial z}\left(-\frac{\partial\phi}{\partial y}\right)=j\omega\epsilon\times-\frac{j\omega\mu}{h^2}\frac{\partial H_z}{\partial y}$$

or

$$H_z+\frac{\partial\phi}{\partial z}=\frac{\omega^2\mu\epsilon}{h^2}H_z$$

or

$$\frac{\partial\phi}{\partial z}=\left(\frac{\omega^2\mu\epsilon}{h^2}-1\right)H_z$$

or

$$\frac{\partial\phi}{\partial z}=-\left(\frac{h^2}{j\omega\mu}+j\omega\epsilon\right)\left(\frac{j\omega\mu}{h^2}H_z\right) \qquad ...14\ (a)$$

This equation corresponds to current equation of transmission line.

(*ii*) Voltage equation :

Further, putting eqs. (11) and (12) in eq. 13 (*b*), we get

$$\frac{\partial}{\partial z}\left(\frac{-j\omega\mu}{h^2}\frac{\partial H_z}{\partial y}\right)=j\omega\mu\frac{\partial\phi}{\partial y}$$

or $$\frac{\partial}{\partial z}\left(\frac{j\omega\mu H_z}{h^2}\right)=-j\omega\mu\phi. \qquad \ldots 14\,(b)$$

This equation will correspond to the voltage equation of transmission line.

The quantity $\frac{j\omega\mu H_z}{h^2}$ has the dimensions of voltage† and ϕ has the dimensions of current††, so that the equations (14) take the form

$$\left.\begin{aligned}\frac{\partial I}{\partial z}&=-\left(\frac{h^2}{j\omega\mu}+j\omega\epsilon\right)V\\ \text{and}\quad \frac{\partial V}{\partial z}&=-j\omega\mu I\end{aligned}\right] \qquad \ldots(15)$$

which, on comparing with transmission line equation (6), gives that

$$\left.\begin{aligned}Z&=j\omega\mu\\ \text{and}\quad Y&=j\omega\epsilon+\frac{h^2}{j\omega\mu}.\end{aligned}\right] \qquad \ldots(16)$$

Basing on equation (16) 'equivalent transmission line' for *TE* waves can be plotted as shown in figure 13.

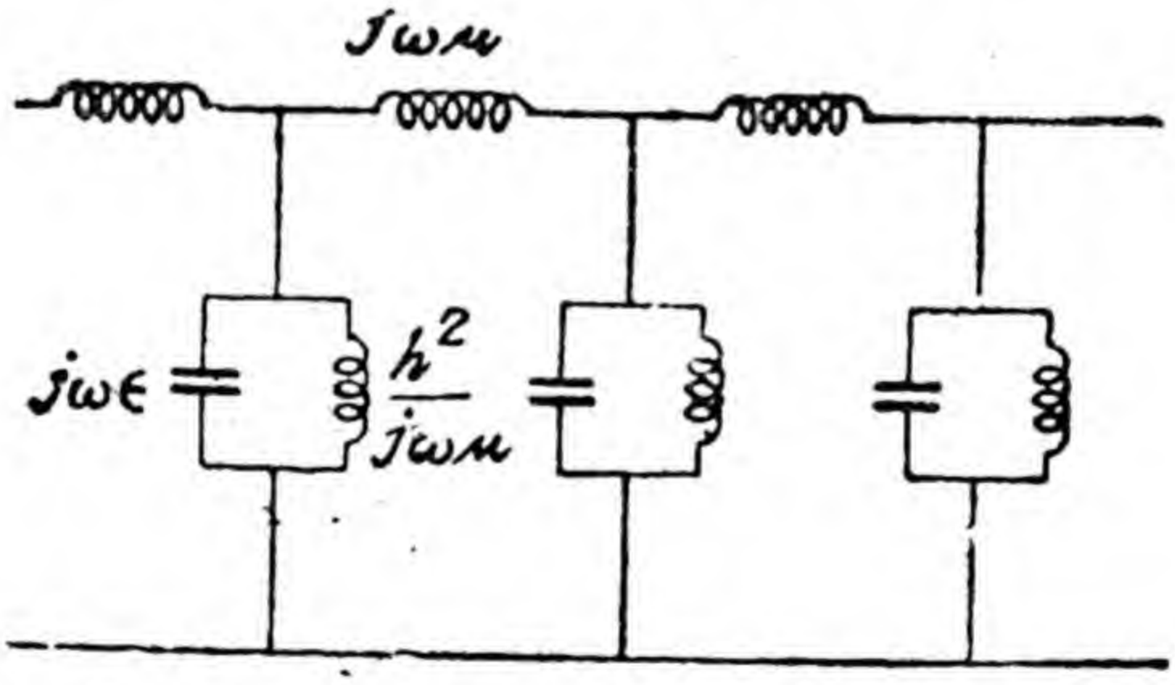


Fig. 13. Equivalent transmission line circuit representation for *TE* waves.

The cutoff frequency of this 'equivalent circuit' occurs when the shunt susceptance equals zero. But

$$\text{shunt susceptance}=\frac{1}{\text{shunt reactance}}$$

$$=Y \text{ (since no real part in admittance)}$$

and, therefore, for cutoff frequency, we have

$$Y=0,$$

---

†From equation (12)

$$E_x=-\frac{\partial}{\partial y}\left(\frac{j\omega\mu H_z}{h^2}\right)=-\frac{\partial V}{\partial y}$$

††From equation (11) $H_z=-\partial\phi/\partial x$, since variation of current should be associated with magnetic field, ϕ has the dimensions of current.

or $$j\omega\epsilon+\frac{h^2}{j\omega\mu}=0.$$

or $$h^2=\omega_c^2\mu\epsilon,$$

which is same as derived in the theory of *TE* waves.

The characteristic impedance of the 'equivalent transmission line' is

$$Z_0\,(TE)=\sqrt{\left(\frac{Z}{Y}\right)}$$

$$=\left(\frac{j\omega\mu}{j\omega\epsilon+h^2/j\omega\mu}\right)=\frac{\sqrt{(\mu/\epsilon)}}{\sqrt{(1-\omega_o^2/\omega^2)}}, \qquad ...(17)$$

which is equal to the wave impedance of the waveguide in *TE* mode as shown below :

The wave impedance which is seen when looking in the direction of propagation, that is, along z-direction is denoted by Z_z and is given as

$$Z_z=Z_{xy}=Z_{yx}=\left|\frac{E_{trans.}}{H_{trans.}}\right|$$

But for *TE* waves (putting $E_z=0$ in equation (3) of art. 27·4), we find that

$$Z_{xy}=\frac{E_x}{H_y}=\frac{\omega\mu}{\beta},$$

on using equations (27) and (30) of art 27·4.

Therefore, $$Z_z=\frac{\omega\mu}{\sqrt{(\omega^2\mu\epsilon-\omega_c^2\mu\epsilon)}}$$

$$=\frac{\sqrt{(\mu/\epsilon)}}{\sqrt{(1-\omega_o^2/\omega^2)}}.$$

Comparing with equation (17) we, thus, find that

$$Z_0\,(TE)=Z_z,$$

i.e. the characteristic impedance of 'equivalent transmission line' is equal to the wave impedance for *TE* waves.

In the above discussion of 'equivalent circuit' representation, we find that a waveguide operating in a particular mode *i.e.* *TE* or *TM* mode can be represented by an 'equivalent transmission line circuit' which gives

(a) a characteristic impedance equal to the wave impedance,
(b) same condition for cutoff frequency,
(c) identical voltage and current equations.

27·8. ATTENUATION FACTOR AND Q OF WAVEGUIDES :

First we shall find out attenuation factor due to imperfect conductors in the guide.

While dealing with the theory of waveguides, we have assumed that the walls of the guide are perfectly conducting so

that resistance R offered by them is zero and also that dielectric is lossless so that $G=0$. With these assumptions, a condition was obtained that when γ is imaginary, propagation is possible and $\alpha=0$. However, whereas the dielectric within the guide may be very nearly perfect (hence lossless), guide walls would certainly have some resistance. This means even above cutoff frequency, α, the attenuation constant, will have some value and is, therefore, not zero. The attenuation factor in the range of propagation is given by

$$\alpha=\frac{\text{Power lost per unit length of the wall}}{2\times\text{Power transmitted}}$$

$$=\frac{P_{lost}}{2\times P_{trans.}}.$$

In order to determine the power lost per unit length of the guide, we should first calculate the linear current density in the walls. If it is $|J_S|$,

$$|J_S|=|H_{trans}|$$

so that power lost per unit length of the guide is

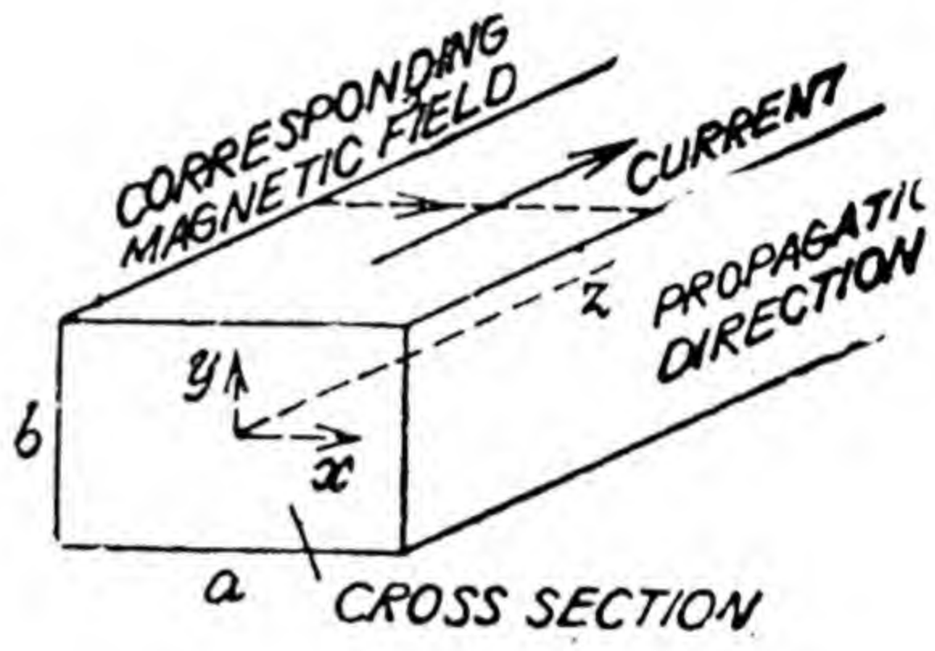


Fig. 14.

$$P_{lost}=\tfrac{1}{2}R_S\int_{surf.}|H_{tang.}|^2\,da,$$

where the intergration is taken over the wall surface of unit length of the guide and R_S is the actual surface resistance of the walls. Let us take a rectangular guide operation in TM mode. For this type, we have shown that

$$H_x^\circ=\frac{\omega\epsilon C}{h^2}\cdot A\sin Bx\cos Ay$$

$$H_y^\circ=\frac{-\omega\epsilon C}{h^2}B\cos Bx\sin Ay$$

$$E_x^\circ=\frac{-\beta C}{h^2}B\sin Bx\sin Ay$$

$$E_y^\circ=\frac{-\beta C}{h^2}A\sin Bx\cos Ay$$

In which $A=\frac{n\pi}{b}$ and $B=\frac{m\pi}{a}$. Also at cutoff, $h^2=\omega_c{}^2\mu\epsilon$.

Now H_x° is tangential to xz plane and its maximum value at $y=0$ or $y=b$ will be when $\cos Ay=1$. Therefore maximum value of H_x° is

$$H_x^\circ=\frac{\omega\epsilon C}{h^2}A\sin Bx$$

Similarly H_y° is tangential to yz plane and its maximum value at $x=0$ or $x=a$ is

$$H^\circ_y = \frac{-\omega\epsilon C}{h^2} B \sin Ay$$

With these values of tangential components of magnetic field we can write

$$P_{lost} = R_S \left(\int_0^a H_x^{\circ 2} dx + \int_0^b H_y^{\circ 2} dy \right) \text{ watts/meter of length.}$$

The factor $\frac{1}{2}$ being removed by the fact that there are two faces acted upon by each field. Thus

$$P_{lost} = R_S \left(\frac{\omega\epsilon C}{h^2}\right)^2 \left[A^2 \int_0^a \sin^2 \left(\frac{m\pi x}{a}\right) dx + B^2 \int_0^b \sin^2 \left(\frac{n\pi y}{b}\right) dy \right]$$

$$= R_S \left(\frac{\omega\epsilon C}{h^2}\right)^2 \left[A^2 \frac{a}{2} + B^2 \frac{b}{2} \right] \quad \text{.. (1)}$$

Now we proceed to calculate $P_{trans.}$ The power passing any point in the guide is

$$= \tfrac{1}{2} (\mathbf{E} \times \mathbf{H}^*)_z$$

$$= \tfrac{1}{2} (E_x H_y - E_y H_x)$$

$$= \tfrac{1}{2} \left[\frac{\beta\omega\epsilon C^2}{h^4} \cdot B^2 \cos^2 Bx \sin^2 Ay + \frac{\beta\omega\epsilon C^2}{h^4} A^2 \sin^2 Bx \cos^2 Ay \right]$$

$$= \frac{\beta\omega\epsilon C^2}{2h^4} [B^2 \cos^2 Bx \sin^2 Ay + A^2 \sin^2 Bx \cos^2 Ay]$$

The total power transmitted may be obtained by integration over the cross sectional area of the guide. Thus

$$P_{trans.} = \frac{\beta\omega\epsilon C^2}{2h^4} \left[B^2 \int_0^a \int_0^b \cos^2 \left(\frac{m\pi x}{a}\right) \sin^2 \left(\frac{n\pi y}{b}\right) dx\, dy + A^2 \int_0^a \int_0^b \sin^2 \left(\frac{m\pi x}{a}\right) \cos^2 \left(\frac{n\pi y}{b}\right) dx\, dy \right]$$

$$= \frac{\beta\omega\epsilon C^2}{h^4} \left[B^2 \frac{ab}{4} + A^2 \frac{ab}{4} \right] \quad \text{...(2)}$$

From equations (1) and (2), we write the attenuation factor as

$$\alpha = \frac{P_{lost}}{2 \times P_{trans.}} = \frac{R_S \left(\frac{\omega\epsilon C}{h^2}\right)^2 \left[A^2 \frac{a}{2} + B^2 \frac{b}{2} \right]}{2 \times \frac{\beta\omega\epsilon C^2}{2h^4} \left[B^2 \frac{ab}{4} + A^2 \frac{ab}{4} \right]}$$

$$= \frac{2R_S\omega\epsilon [A^2 a + B^2 b]}{\beta ab [B^2 + A^2]}$$

$$= \frac{2R_S\omega\epsilon}{\beta h^2} [A^2/b + B^2/a]$$

$$=\frac{2\ R_S\ \omega\epsilon}{\beta h^2}\left[\frac{n^2\pi^2}{b^3}+\frac{m^2\pi^2}{a^3}\right]$$

Putting $h^2=4\pi^2 f_c^2\mu\epsilon$

$$\beta=\omega\sqrt{(\mu\epsilon)}(1-f_c^2/f^2)^{1/2}$$

$$R_S=\sqrt{\left(\frac{\omega\mu_m}{2\sigma_m}\right)},$$

we get

$$\alpha_{TM}=\frac{2.\sqrt{\left(\frac{\omega\mu_m}{2\sigma_m}\right)}.\ \omega\epsilon}{\omega\sqrt{(\mu\epsilon)}(1-f_c^2/f^2)^{1/2}}\left[\frac{n^2\pi^2}{b^3}+\frac{m^2\pi^2}{a^3}\right].\frac{1}{4\pi^2 f_c^2\mu\epsilon}$$

$$=\frac{\left(\frac{2\pi f\mu_m}{2\sigma_m}\right)^{1/2}.v.}{(1-f_c^2/f^2)^{1/2}.2f_o^2\mu}\left[\frac{n^2}{b^3}+\frac{m^2}{a^3}\right]$$

$$=\frac{\left(\frac{\pi\mu_m}{\sigma_m}\right)^{1/2}.\sqrt{f}.v^2}{2\eta f_c^2\ (1-f_o^2/f^2)^{1/2}}\left[\frac{n^2}{b^3}+\frac{m^2}{a^3}\right]$$

where

$$v=\frac{1}{\sqrt{(\mu\epsilon)}}\text{ and }\eta=\sqrt{(\mu/\epsilon)}$$

Therefore we put attenuation factor in the form

$$\alpha_{TM}=\sqrt{\left(\frac{\pi\mu_m}{\sigma_m}\right)}\ \frac{v^2\sqrt{f}}{2\eta f_o^2}\ \frac{(n^2/b^3+m^2/a^3)}{\sqrt{(1-f_o^2/f^2)}}\text{ nepers/meter.}$$

By the same method attenuation factors for other forms of the guide are :

Rectangular guide $TE_{m,n}$ *mode* :

$$\alpha_{TE}=\sqrt{\left(\frac{\pi\mu_m}{\sigma_m}\right)}.\frac{2\sqrt{f}}{b\eta\sqrt{(1-f_o^2/f^2)}}\left[\left(1-\frac{f_o^2}{f^2}\right)\frac{bm^2/an^2+1}{bm^2/an^2+a/b}+\left(1+\frac{b}{a}\right)\left(\frac{f_c}{f}\right)^2\right]\text{ nepers/}m.$$

Cylindrical guide $TE_{n,m}$ *mode* :

$$\alpha_{TE}=\sqrt{\left(\frac{\pi\mu_m}{\sigma_m}\right)}\frac{\sqrt{f}}{\sqrt{(1-f_o^2/f^2)}}\left[\left(\frac{f_o}{f}\right)^2+\frac{n^2}{\tau_{n,\,m}{}^2-n^2}\right]\text{ nepers/}m.$$

Cylindrical guide $TM_{n,m}$ *mode* :

$$\alpha_{TM}=\sqrt{\left(\frac{\pi\mu_n}{\sigma_m}\right)}\frac{\sqrt{f}}{b\eta\sqrt{(1-f_o^2/f^2)}}\text{ nepers/}m.$$

These are the expressions from which the attenuation in the guide due to imperfect conductors can be found.

A plot of attenuation factors as a function of frequency for some modes for cylindrical copper guide of 12·7 cm. diameter is shown in the fig. 15.

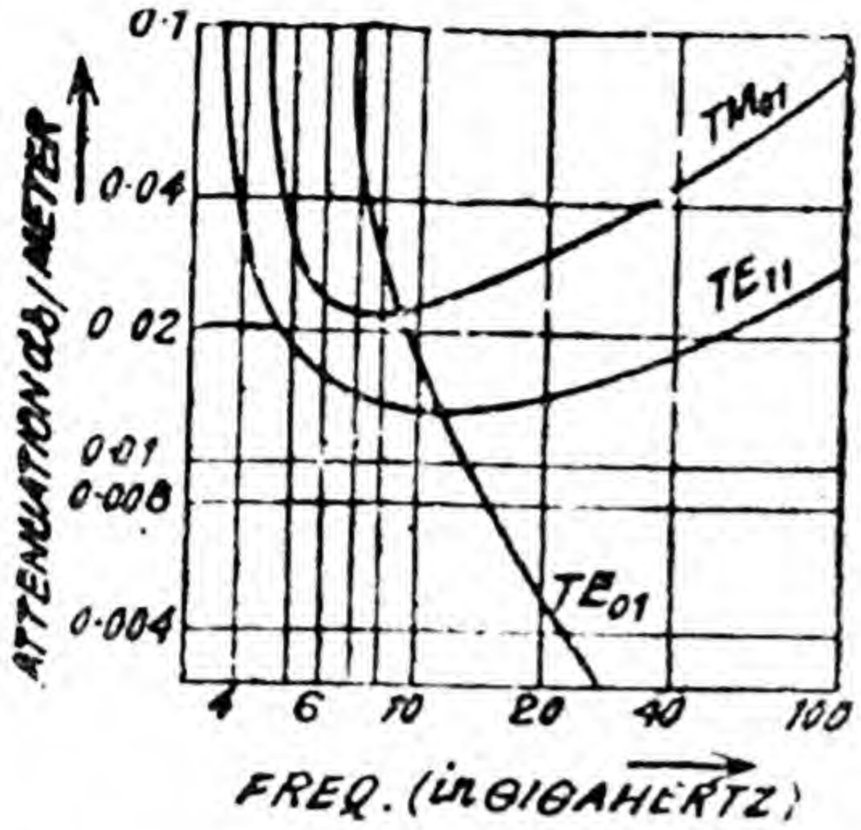


Fig. 15. α for various modes of a cylindrical guide of 12·7 cm. diameter.

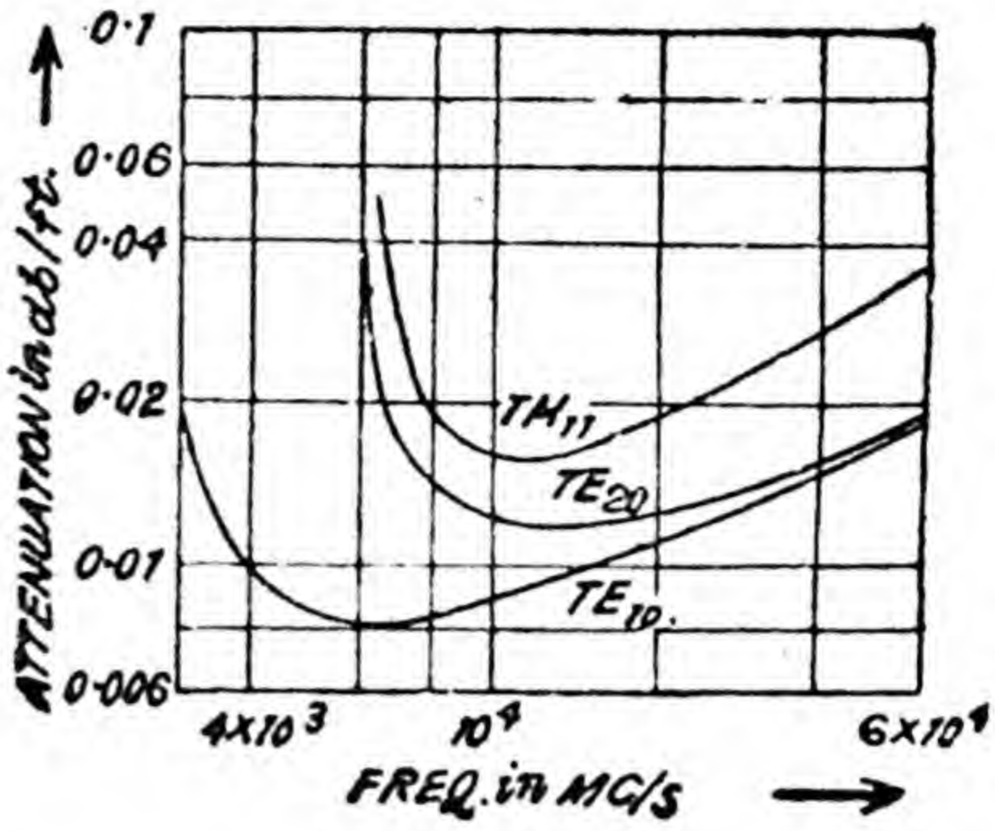


Fig. 16. Attenuation in rectangular copper waveguide for several modes ($a=2''$, $b=1''$).

Attenuation factor as function of frequency for a rectangular copper waveguide for several modes is plotted in fig. 16.

Q OF WAVEGUIDE :

To the attenuation factor, α, another factor Q called the 'quality factor' is closely related and is defined as

$$*Q=\omega\times\frac{\text{energy stored per unit length}}{\text{energy lost per unit length per second}}.$$

We know that transmission per second of 'energy stored per unit length' gives energy transmitted per second and hence energy transmitted per second is

$$=v\times(\text{energy stored per unit length}),$$

where v is the group velocity in the waveguide.

Or energy stored per unit length

$$=\frac{1}{v}\times\text{power transmitted}$$

$$\therefore\quad Q=\frac{\omega}{v}\left(\frac{\text{power transmitted}}{\text{power lost per unit length}}\right)$$

$$=\frac{\omega}{2\alpha v}.$$

For waveguides, group velocity v and phase velocity v_g in the guide are related as

$$v=\frac{v_0^2}{v_g}$$

where $v=\dfrac{1}{\sqrt{(\mu\epsilon)}}$, wave velocity in free space.

---

*As an analogy with circuit concept

$$Q=\frac{\omega L}{R}=\omega\ \frac{LI^2}{RI}=\omega.\ \frac{\text{energy stored}}{\text{energy lost}}.$$

Therefore,

$$Q=\frac{\omega v_g}{2\alpha {v_0}^2}.$$

We find

$$v_g=\frac{\omega}{\beta}=\frac{\omega}{\sqrt{(\omega^2\mu\epsilon-\omega^2{}_c\mu\epsilon)}}=\frac{v_0}{\sqrt{(1-\omega_c{}^2/\omega^2)}}.$$

so that

$$Q=\frac{\omega}{2\alpha v_c\sqrt{(1-\omega_c{}^2/\omega^2)}}.$$

27·9. SINGLE-CONDUCTING-WAVEGUIDE :

Electromagnetic waves can be guided along a bare cylindrical metal wire. A surface wave mode field distribution along such a wire is shown in fig. 17. According to Gobau, for a bare copper wire of 1 m.m. radius stretched in air and excited at 3000 mc/s, the attenuation is about 0·023 db per meter and the extent of field is such that 75 percent of transmitted energy is conveyed within a circle 18 cm. in radius. For a wire of 1 cm. radius the attenuation is 0 0026 db per meter and the radius of 75 percent power circle is 80 cm. The attenuation of 0·0026 *db* per meter compares favourably with that of rectangular waveguide but the extent of field is prohibitively large, which can be reduced by employing a thin layer coating of high-dielectric constant material over the wire.

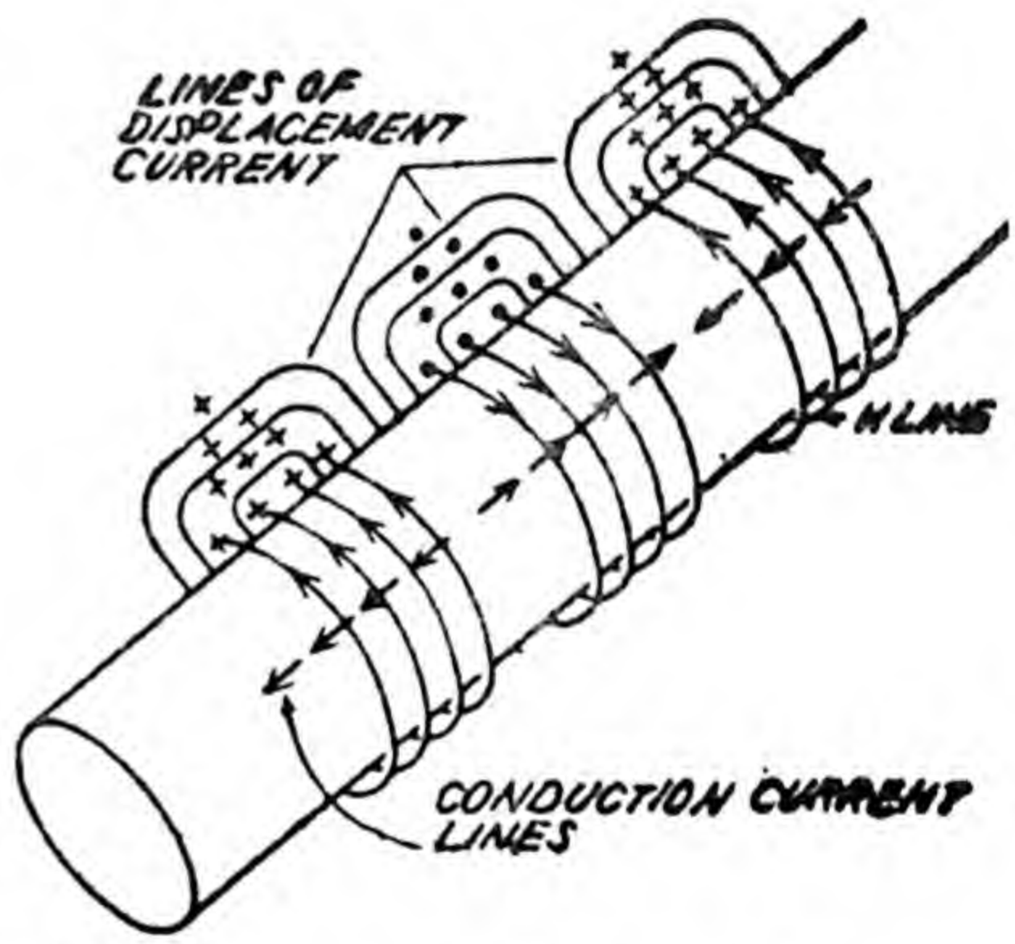


Fig. 17. Field distribution for surface wave mode.
+*H* lines going into the paper
°*H* lines coming out of paper.

27·10. CHOICE OF TYPF OF WAVEGUIDE AND OF WAVE-GUIDE DIMENSIONS :

It is always desired that propagation in a guide be affected by dominant mode without the presence of higher order modes because dominant mode propagation has the following advantages over the higher order mode propagation :

(*i*) It has lower power dissipation.

(*ii*) It requires smaller, lighter, and cheaper guiding structures at the same excitation frequency.

(*iii*) It requires simple components for coupling power into and out of the guide and for impedance matching.

The presence of higher modes will result in more power

dissipation, and since coupling device is designed for dominant mode, the power transported in higher order modes may not be coupled to load. Though the waveguide may be excited in a single mode but due to the presence of discontinuities in the conducting surfaces, higher order mode may be generated. Any modes for which the cut-off frequency is below the excitation frequency will be propagated in both directions from the discontinuity.

At microwaves, rectangular waveguide and coaxial lines are frequently employed. In circular wave guides the discontinuities change the angle of polarisation of the wave as the wave moves down the guide and since proper position of coupling probe depends upon the angle of polarisation, difficulty arises in coupling the energy to the receiver. Therefore circular waveguides are not so frequently used.

At frequencies below 3000 mc/s the dimensions of the guide for dominant mode propagation has to be quite large. This increases weight and cost of the guide. Above 3000 mc/s. guide becomes convenient and economical to use.

27·11. THE TEM WAVE IN THE COAXIAL LINE :

(a) Field Equations :

We have already discussed that it is a special case of *TM* mode in which $m=0$. It is shown that for *TEM* wave $E_z=H_z=0$. Accordingly for all fields to be varying in time and propagating in z direction, we can write

$$\mathbf{H}=\mathbf{H}^{\circ}\, e^{-\gamma z} \sin \omega t$$

$$\mathbf{E}=\mathbf{E}^{\bullet}\, e^{-\gamma z} \sin \omega t$$

Then in cylindrical co-ordinates in a non-conducting region, Maxwell's equations are

$$\left.\begin{aligned}
&\frac{1}{r}\cdot\frac{\partial H_z^{\circ}}{\partial \phi}+\gamma H_{\phi}^{\circ}=j\omega\epsilon E_r^{\circ}\\
&-\gamma H_r^{\circ}-\frac{\partial H_z^{\circ}}{\partial r}=j\omega\epsilon E_{\phi}^{\circ}\\
&\frac{1}{r}\cdot\frac{\partial}{\partial r}(rH_{\phi}^{\circ})-\frac{1}{r}\frac{\partial H_r^{\circ}}{\partial \phi}=j\omega\epsilon E_z^{\circ}
\end{aligned}\right\}$$

$$\left.\begin{aligned}
&\frac{1}{r}\frac{\partial E_z^{\circ}}{\partial \phi}+\gamma E_{\phi}^{\circ}=-j\omega\mu H_r^{\circ}\\
&-\gamma E_r^{\circ}-\frac{\partial E_z^{\circ}}{\partial r}=-j\omega\mu H_{\phi}^{\circ}\\
&\frac{1}{r}\frac{\partial}{\partial r}(rE^{\circ}{}_{\phi})-\frac{1}{r}\frac{\partial E_r^{\circ}}{\partial \phi}=-j\omega\mu H_z^{\circ}
\end{aligned}\right\}$$

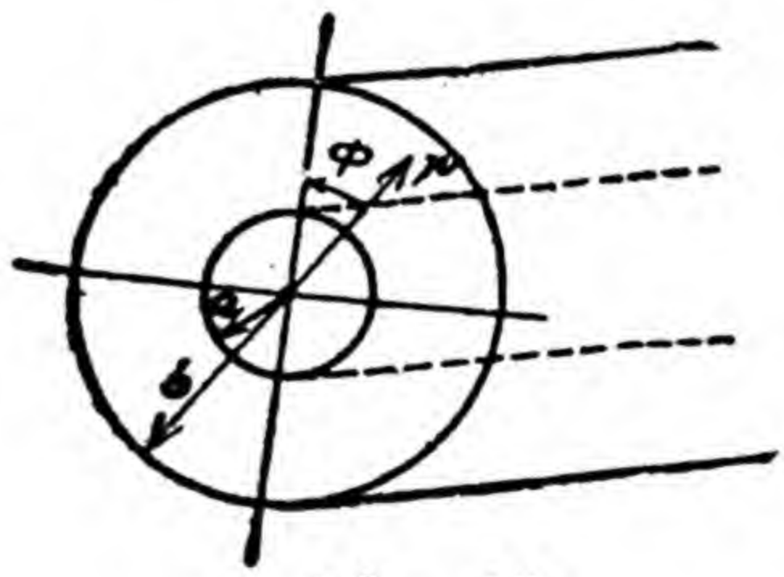


Fig. 18.

Since the electrostatic field will be totally radial, $E_{\phi}=0$. Similarly the magnetic field will be in the form of closed loop

around the inner conductor, $H_r=0$. With $E_z=E_\phi=0$, and $H_z=H_r=0$, the only components of the field left are E_r and H_ϕ. From above equations we, therefore, arrive at

$$\left.\begin{aligned} \gamma H_\phi^\circ &= j\omega\epsilon E_r^\circ \\ \frac{\partial}{\partial r}(rH_\phi^\circ) &= 0 \\ \gamma E_r^\circ &= j\omega\mu H_\phi^\circ \\ \frac{\partial E_r^\circ}{\partial \phi} &= 0 \end{aligned}\right]$$

Thus
$$H_\phi^\circ = \frac{j\omega\epsilon}{\gamma} E_r^\circ$$

also
$$H_\phi^\circ = \frac{\gamma}{j\omega\mu} E_r^\circ$$

From these two equations we can then put

$$\frac{j\omega\epsilon}{\gamma} = \frac{\gamma}{j\omega\mu}$$

or
$$\gamma^2 = -\omega^2\mu\epsilon$$

or
$$\gamma = j\beta = j\omega\sqrt{(\mu\epsilon)}$$

which is the same expression as we have obtained for phase constant β in theory of dissipationless ($\alpha=0$) transmission line. Consequently *TEM* wave propagation in the coaxial cable can be approximated with the propagation on dissipationless open wire line.

In parallel plane guide, we have obtained cutoff frequency by putting $\beta=0$ and got,

$$\beta = \sqrt{\left[\omega^2\mu\epsilon - \left(\frac{m\pi}{a}\right)^2\right]} = 0.$$

Simila ly here cutoff frequency is

$$\beta = j\omega_c\sqrt{(\mu\epsilon)} = 0$$

or
$$f_c = 0.$$

That means all the frequencies down to zero can be propagated along the cable by *TEM* mode.

(b) Voltage and Current Equations :

One of the field equations is

$$\frac{\partial}{\partial r}(rH_\phi^\circ) = 0$$

which gives

$$H_\phi^\circ = \frac{k}{r},$$

where k is not the function of r. Now by Ampere's law, if i is the instantaneous current enclosed by a path and H_ϕ is the corresponding component of instantaneous magnetic field then

$$\oint H\,dl = i$$

Suppose in the inner conductor $i = I$ and we apply the above law to the circumference of the inner conductor then

$$2\pi a H_\phi{}^\circ = I$$

or

$$H_\phi{}^\circ = \frac{I}{2\pi a},$$

so that comparison of it with

$$H_\phi{}^\circ = \frac{k}{r}$$

at $r = a$ will give

$$k = \frac{I}{2\pi}.$$

Then the magnetic field intensity interms of inner conductor current will be

$$A_\phi{}^\circ = \frac{I}{2\pi r}$$

Further from field equation

$$\gamma E_r{}^\circ = j\omega\mu H_\phi{}^\circ,$$

On putting the value of γ, we get

$$E_r{}^\circ = \frac{j\omega\mu}{j\omega\sqrt{(\mu\epsilon)}}\, H^\circ = \sqrt{\left(\frac{\mu}{\epsilon}\right)}\, H_\phi{}^\circ$$
$$= \eta H_\phi{}^\circ.$$

The voltage drop from the inner conductor to the outer conductor of the coaxial line may be written interms of the maximum value $|V|$ as

$$|V| = \int_a^b E_r{}^\circ\, dr = \int_a^b \eta \frac{|I|}{2\pi r}\, dr = \eta \frac{|I|}{2\pi} \log_e \frac{b}{a}.$$

Inserting propagation factor and time function, this value is

$$V = \eta . \frac{|I|}{2\pi} . \log_e \frac{b}{a}\, e^{-j\beta z} \sin \omega t$$

so that

$$\frac{\partial V}{\partial z} = -j\beta\eta \frac{|I|}{2\pi} \log_e \frac{b}{a}\, e^{-j\beta z} \sin \omega t$$
$$= \frac{-j.\omega\sqrt{(\mu\epsilon)}}{2\pi} \cdot \sqrt{\left(\frac{\mu}{\epsilon}\right)} . \log_e \frac{b}{a} . |I|\, e^{-j\beta z} \sin \omega t$$
$$= -\frac{j\omega\mu I}{2\pi} \log_e \frac{b}{a},$$

where

$$I = |I|\, e^{-j\beta z} \sin \omega t.$$

With this expression for current, we can find

$$\frac{\partial I}{\partial z} = -j\beta |I|\, e^{-j\beta z} \sin \omega t$$

$$= \frac{-j\beta \mid I \mid . V}{\eta \frac{\mid I \mid}{2\pi} \log_e \frac{b}{a}}$$

$$= \frac{-j2\pi\beta V}{\eta \log_e \frac{b}{a}}$$

$$= \frac{-j2\pi\omega\sqrt{(\mu\epsilon)}\, V}{\sqrt{\left(\frac{\mu}{\epsilon}\right)} \log_e \frac{b}{a}} = -\frac{j2\pi\omega\epsilon V}{\log_e \frac{b}{a}}.$$

In the article 'Line Parameters' of transmission line, it has already been shown that

$$L = \frac{\mu}{2\pi} \log_e \frac{b}{a} \text{ and } C = \frac{2\pi\epsilon}{\log_e \frac{b}{a}}.$$

Inserting these values in the expressions for voltage and current derivations, we get

$$\left.\begin{aligned} \frac{\partial V}{\partial z} &= -j\omega LI \\ \text{and} \quad \frac{\partial I}{\partial z} &= -j\omega CV \end{aligned}\right\}$$

which are same as we have already derived for a transmission line with $R=G=0$. Thus we get identical voltage and current equations for the two cases one for *TEM* mode on coaxial cable and the other for the wave on the dissipationless open wire line. The characteristic impedance can be calculated as

$$Z_0 = \frac{V}{I} = \frac{\eta \frac{\mid I \mid}{2\pi} \log_e \frac{b}{a} e^{-j\beta z} \sin \omega t}{\mid I \mid e^{-j\beta z} \sin \omega t}$$

$$= \frac{\eta}{2\pi} \log_e \frac{b}{a} \text{ ohms.}$$

If we put $\eta = 120\pi$ for free space, then

$$Z_0 = \frac{120\pi}{2\pi} \log_e \frac{b}{a} \text{ ohms.}$$

$$= 60 \log_e \frac{b}{a} \text{ ohms.}$$

(c) ATTENUATION IN THE COAXIAL LINE :

We have already mentioned that at $r=a$, the tangential magnetic field is

$$H_\phi^\circ = \frac{\mid I \mid}{2\pi a}$$

and at $r=b$, (for outer conductor) this value will be changed to

$$H_\phi^\circ = \frac{|I|}{2\pi b}.$$

Therefore current density for inner conductor

$$J_s = H_\phi^\circ = \frac{|I|}{2\pi a}$$

and for outer conductor

$$J_s = H_\phi^\circ = \frac{I}{2\pi b}.$$

Then at some particular point along z-axis, the average power entering through the surface of inner conductor will be

$$= \tfrac{1}{2} (J_s)^2 R_s$$

$$= \tfrac{1}{2} \frac{|I|^2}{4\pi^2 a^2} \sqrt{\left(\frac{\omega \mu_m}{2\sigma_m}\right)} \text{ watts/meter}^2,$$

where μ_m and σ_m correspond to the metal of the conductor. The area of surface of inner conductor is $2\pi a$ square meters per meter of length so that the power supplied to the inner conductor per meter of length is

$$P'_{lost} = 2\pi a \times \tfrac{1}{2} \frac{|I|^2}{4\pi a^2} \sqrt{\left(\frac{\omega \mu_m}{2\sigma_m}\right)}$$

$$= \tfrac{1}{2} \frac{|I|^2}{2\pi a} \sqrt{\left(\frac{\omega \mu_m}{2\sigma_m}\right)}.$$

Similarly power supplied to the outer conductor of radius b per meter of length is

$$P''_{lost} = \tfrac{1}{2} \frac{|I|^2}{2\pi b} \sqrt{\left(\frac{\omega \mu_m}{2\sigma_m}\right)}.$$

So that the total power lost into the conductors per meter of length is

$$P_{lost} = P'_{lost} + P''_{lost}$$

$$= \frac{1}{4\pi} |I|^2 \sqrt{\left(\frac{\omega \mu_m}{2\sigma_m}\right)} \cdot \left\{\frac{1}{a} + \frac{1}{b}\right\} \text{ watts.}$$

The power transmitted past a given point in the guide is

$$= \tfrac{1}{2} (\mathbf{E}^\circ \times \mathbf{H}^{\circ *})_z$$

$$= \tfrac{1}{2} E_r^\circ H_\phi^\circ.$$

Then the power transmitted over the space between the conductors per unit length is

$$P_{trans.} = \int_a^b \tfrac{1}{2} E_r^\circ H_\phi^\circ . 2\pi r \, dr$$

$$= \int_a^b \tfrac{1}{2} \eta H_\phi^{\circ 2} \, 2\pi r \, dr$$

$$= \int_a^b \tfrac{1}{2} \eta \frac{|I|^2}{4\pi^2 r^2} \, 2\pi r \, dr$$

$$= \frac{|I|^2}{4\pi} \eta \log_e \frac{b}{a} \text{ watts.}$$

Therefore attenuation factor is

$$\alpha = \frac{P_{lost}}{2 \times P_{trans.}} = \frac{\sqrt{\left[\pi \mu_m / \sigma_m\right]} \left[\frac{1}{a} + \frac{1}{b}\right] \cdot \sqrt{f}}{2\eta \log_e \frac{b}{a}} \text{ nepers/meter,}$$

which shows that attenuation is proportional to the square root of frequency.

COMPARISON OF WAVEGUIDES AND COAXIAL LINES :

Waveguides are used for the transmission of radio frequency power from one point to another. In radar, they deliver r.f. power from transmitter to antenna during the transmission of pulses, and from antenna to the receiver during the reception of echoes. These are most suited at frequencies in the *microwave region* (ultra high frequencies). For the transmission of radio frequency power at longer wavelengths, large guides would be required which is quite inconvenient.

Advantages of Waveguides :

1. They are quite simple in construction and no inner conductor, as in coaxial lines, is needed.

2. Since the electric and magnetic fields are confined to the space within the guides, there is a little loss of power due to radiation.

3. Guides are normally filled with air due to which dielectric loss is not of any practical importance.

4. From (2) and (3) points we conclude that overall attenuation of a waveguide is lower. Moreover, the power lost as heat in the walls of a standard size waveguide is quite small.

5. The peak power capacity is greater than that of coaxial lines. The maximum peak power than can be transmitted by a 1 by $\frac{1}{2}$ in, waveguide is 1 *MW*.

Disadvantages of waveguides :

1. Its use is restricted to higher frequencies In practice, they are seldom employed below 3000 mc/s. The physical size of the guide to use below this frequency becomes quite bulky.

2. Installation of waveguides is much more difficult than for other types of transmission lines. Care should be taken in soldering sections of the guide.

27·12. RESONANT CAVITIES :

We have already discussed the use of open and short circuited lines as resonant circuits at high frequencies Waveguides can also be used as fixed or tunable resonators. Waveguide resonators are

ordinarily short-circuited at both ends by fixed walls or movable plungers. They differ from open wire or coaxial cable resonators in the following aspects :

(*i*) they have a cut off frequency below which propagation does not take place, and

(*ii*) the phase velocity of propagation is a function of the ratio of the frequency to the cut off frequency.

As a result of these characteristics, the resonance wavelengths of waveguide resonators are not directly proportional to the resonator length, and the resonance frequencies do not approach zero as the resonator length is increased, but approach a limiting value which is equal to the cut off frequency in all but the lowest-order mode of resonance.

When an excited waveguide is closed by a perfectly conducting sheet at some point, a standing wave pattern will set up along the axis of propagation. The tangential electric intensity will be zero at the end wall closure, and also at points half wavelength back along the guide (wavelength being measured along the guide wall and hence interms of phase velocity). If a perfectly conducting plane is inserted at the points $\lambda/2$ distant from the closure wall then, electric intensity being zero there, field distribution present between the plane and the end wall will not be disturbed, the exciting antenna being considered as present between the shorting planes. A volume or cavity is then formed which has resonant properties similar to those of the parallel resonant circuit since only definite frequencies can produce the necessary standing wave field conditions. The lowest resonant frequency field configuration corresponds to dominant mode.

Since such a cavity has other dimensions also terminated in conducting planes, there is a possibility of resonance at other frequencies, depending on excitation and orientation of the cavity.

The shape of the cavity resonator is governed by such considerations as the desired values of resonance frequency, shunt conductance, and Q ; by the purpose for which they are to be used ; and by the manner in which they are to be tuned.

RESONANT FREQUENCY OF THE CAVITY :

We shall now proceed to find the resonant frequency of the cavity. We assume that there is a rectangular cavity with perfectly conducting walls and good dielectric, and excited in TE_{mn} mode. The initial electric fields would be (refer to rectangular wave guide TE_{mn} mode field equations) :

$$E_x = \frac{\omega\mu}{h^2} CA \cos Bx \sin Ay \, e^{-j\beta z}$$

$$E_y = -\frac{\omega\mu}{h^2} CB \cos Ay \sin Bx \, e^{-j\beta z}$$

$$E_z = 0$$

Given

$$\frac{v_{gr.}}{v_0}=0{\cdot}9$$

and free space wavelength

$$\lambda_0=\frac{3\times10^{10}}{6\times10^9}=5 \text{ cm.}$$

We know that (See ex. 4)

$$\frac{v_{gr.}}{v_0}=\frac{\lambda_0}{\lambda_g}$$

so that wavelength within the guide is

$$\lambda_g=\lambda_0\left(\frac{v_0}{v_{gr}}\right)=\frac{5\times10}{9}=\frac{50}{9}=5{\cdot}55 \text{ cm.}$$

The cut off wavelength is given by

$$\frac{1}{\lambda_c^2}=\frac{1}{\lambda_0^2}-\frac{1}{\lambda_g^2}$$

$$=\frac{1}{(5)^2}-\frac{1}{(5{\cdot}55)^2}$$

$$=\cdot04-\cdot0324$$

$$=\cdot0075.$$

or $$\lambda_c=11{\cdot}52 \text{ cm.}$$

Therefore the breadth of the guide is given by

$$a=\frac{\lambda_c}{2}=\frac{11{\cdot}52}{2}=5{\cdot}76 \text{ cm.}$$

Ex. 6. *A 6 Gc/s signal is being propagated in a waveguide. The breadth of the guide is 6 cm. Calculate the characteristic wave impedance of this rectangular waveguide for the first two* $TE_{m,0}$ *modes.*

Refer to art. 27·7, we have obtained an expression of wave impedance of waveguide operating in *TE* mode as given by

$$Z_z=\frac{\sqrt{(\mu/\epsilon)}}{\sqrt{(1-\omega_c^2/\omega^2)}}=\frac{120\pi}{\sqrt{(1-\lambda_0^2/\lambda_c^2)}}$$

Free space wavelength

$$\lambda_0=\frac{3\times10^{10}}{6\times10^9}=5 \text{ cm.}$$

Cut off wavelength

$$\lambda_c=\frac{2a}{m}=\frac{2\times6}{m}$$

(*i*) Z_z for $TE_{1,0}$

$$\lambda_0=5 \text{ cm.}$$

$$\lambda_o=\frac{2\times 6}{1}=12 \text{ cm.}$$

$$Z_z=\frac{120\pi}{\sqrt{\{1-(5/12)^2\}}}=\frac{120\times 3{\cdot}14}{\sqrt{\left(1-\frac{25}{144}\right)}}$$

$$=414{\cdot}49 \text{ ohm.}$$

(*ii*) Z_z for $TE_{2,0}$

$$\lambda_o=\frac{2\times 6}{2}=6 \text{ cm.}$$

$$Z_z=\frac{120\times 3{\cdot}14}{\sqrt{\{1-(5/6)^2\}}}=681{\cdot}65 \text{ ohm.}$$

Ex 7. *A waveguide has an internal breadth equal to 4 cm. and carries the dominant mode. If the characteristic wave impedance is 500 ohm. Calculate the frequency.*

Cut off wavelength is

$$\lambda_c=\frac{2a}{m}=\frac{2\times 4}{1}=8 \text{ cm.}$$

Further wave impedance is given by

$$Z_z=\frac{120\pi}{\sqrt{(1-\lambda_0^2/\lambda_o^2)}}$$

or
$$500=\frac{120\times 3{\cdot}14}{\sqrt{(1-\lambda_0^2/64)}}$$

or
$$1-\frac{\lambda_0^2}{64}=\left(\frac{120\times 3{\cdot}14}{500}\right)^2=0{\cdot}56791$$

$$\lambda_0=8\times\sqrt{(1-0{\cdot}56791)}$$

$$=5{\cdot}258 \text{ cm.}$$

Therefore corresponding frequency of the wave will be

$$f=\frac{v_0}{\lambda_0}=\frac{3\times 10^{10}}{5{\cdot}258}=5{\cdot}70\times 10^9 \text{ c/s}$$

$$=5{\cdot}7 \text{ Gc/s.}$$

Ex. 8. *Calculate the cut off wavelength, and the guide wavelength of a circular waveguide whose internal radius is 2·5 cm. and operates at 9 Gc/s signal propagating in TM_{11} mode.*

Refer to art. 27·5–1. It is given there in the table that

$$\tau_{nm}=(ha)_{11}=3{\cdot}83$$

and cut off wavelength is

$$\lambda_o=\frac{2\pi a}{\tau_{nm}},$$

Where a is the internal radius of the guide. Therefore for TM_{11} mode,

$$\lambda_c = \frac{2 \times 3{\cdot}14 \times 2{\cdot}5}{3{\cdot}83}$$

$$= \mathbf{4{\cdot}1\ cm.}$$

Guide wavelength is given by

$$\lambda_g = \frac{\lambda_0}{\sqrt{(1 - \lambda_0^2/\lambda_c^2)}}$$

$$= \frac{3{\cdot}33}{\sqrt{\left\{1 - \left(\frac{3{\cdot}33}{4{\cdot}1}\right)^2\right\}}} \qquad \left(\lambda_0 = \frac{3 \times 10^{10}}{9 \times 10^9}\right)$$

$$= \mathbf{5{\cdot}70\ cm.}$$

EXERCISES AND PROBLEMS

1. Discuss *TE* and *TM* wave propagation between two parallel conducting planes. Obtain expressions for (*a*) group velocity, (*b*) cut off frequency.
2. Obtain expression for the attenuation factor in case of *TE* and *TM* mode propagation between conducting planes and show their variation with frequency.
3. Discuss $TE_{1,0}$ mode propagation of the wave in a rectangular guide. Obtain field component expressions and sketch the field configuration.
4. Obtain expressions for *TE* or *TM* mode propagation of the wave in cylindrical waveguides.
5. Deduce expression for the attenuation factor in the case of wave propagation within a rectangular guide operating in *TM* mode. Show that Q of the waveguide is

$$Q = \frac{\omega}{2\alpha v_0 \sqrt{(1 - \omega_c^2/\omega^2)}},$$

where the symbols have their usual meaning.

6. Let f_c be the cut off frequency of the propagating mode inside rectangular waveguide and let $f\ (> f_c)$ be the frequency of operation. The wavelengths corresponding to f and f_c in an unbounded medium having the same constituent parameters as that enclosed by the walls of the waveguide are denoted by λ and λ_c respectively. Let λ_g be the wavelength inside the guide. Then shows that

$$\frac{1}{\lambda_g^2} + \frac{1}{\lambda_c^2} = \frac{1}{\lambda^2}.$$

7. Discuss how the waveguide section can be used as resonant cavity. Obtain the expressions for (*a*) resonant frequency and (*b*) Q of the resonant cavity.

8. Discuss the wave propagation along a coaxial line and discuss how the attenuation varies with frequency.
9. Obtain the expressions for the field vectors, cut off wavelength and the characteristic wave impedance in the case of TE_{10} mode. Explain how these waveguides are employed in microwave work.
10. Show that the propagation of *TEM* wave is not possible in a single conductor wave guide. Discuss the characteristic of *TEM* wave propagating between two parallel conducting planes.
11. Give the theory of propagation of microwaves between two parallel conducting planes. Considering the TE_{mo} mode, show that cut off wavelength $\lambda_o = 2a/m$ where a is the separation between the planes.

28 ANTENNA

Transmission line, as discussed previously, is a device to transmit radio frequency energy from the transmitter to a radiator usually known as 'antenna'. The transference of energy for radio communication is affected by the propagation of elec romagnetic waves. These waves represent the electrical energy that has escaped into the free space from a conductor carrying high frequency currents.

It may be mentioned here that for effective radiation from an antenna system frequency used should be high enough because power propagated into space by an electromagnetic wave is given by the Poynting vector

$$\mathbf{P}=\mathrm{E}\times\mathbf{H}.$$

As we shall snow that magnitud of **E** and **H** are directly proportional to frequency f (or inversely to wavelength λ), magnitude of energy vector **P** will be proportional to f^2. Hence high frequency waves will have high energy content. That is why they are useful for long distance communication. In free space though all electromagnetic waves, whether of low or high frequency, will travel with the velocity of light but they will differ in energy content.

The radio frequency, energy corresponding to the intelligence to be transmitted, is fed to a transmitting antenna. The r.f. currents, as a result of applied r.f. electrical energy, flow along the length of the antenna. A portion of the energy, fed to it, is released or radiated into space as electromagnetic waves which then travel with the velocity of light. At the receiver, energy is abstracted from the wave by the receiving antenna. The transmission and reception of electromagnetic waves used for radio communication are, therefore, accomplished by radiators and collectors exposed in space. These radiators and collectors are named as transmitting and receiving antennas respectively. An antenna is a device composed of a system of one or more linear conductors, the electrical dimensions of which vary from a fraction to several wavelength, used to couple a high frequency a.c. generator or receiver to space. In this section of the chapter, we shall discuss transmitting antennas in detail.

The study of antenna should be divided according to following steps :

(*i*) How the energy is released from an antenna ?

(ii) What are the factors that concern with the radiation efficiency of an antenna (radiation and loss resistances, radiated power, mutual coupling, and ground effect etc.)

(iii) Design of an antenna system or of an antenna for the use at different frequencies and for the purpose of greater directivity.

28·1. ANTENNA ACTION :

We have written above that an antenna is a current carrying conductor. If we take a length, *much smaller than the wavleength of the radio frequency wave in operation*, of this conductor then it is called a current element. It can be seen that the electric field of the alternating current element contains the terms that correspond to the field of an oscillating electric dipole. It is not merely by coincidence but we shall show that a current element can, infact, be approximated with an oscillating electric dipole.

An electric dipole is formed of a pair of charges of equal magnitude and of opposite signs as shown in fig 1 (*b*). Its dipole moment is $\mathbf{p}=qd\mathbf{l}$. For oscillating electric dipole, we should have $q=q_0\ e^{j\omega t}$. Then moment of such a dipole would be

$$\mathbf{p}=qd\mathbf{l}=q_0\ e^{j\omega t}\ d\mathbf{l}.$$

Such an electric dipole whose moment is a sinusoidal function of time, can be made up of a pair of conductors of finite capacitance, such as spheres or plates, joined by a thin wire of negligible capacitance. The charges will then be situated mostly on the electrodes, but can still flow from one to the other.

Suppose $I\ (=I_0\ \cos\ \omega t)$ is the current across the element. Equation of continuity (or conservation of charge) requires that there be an accumulation of charge at the ends of the element, and is given by

$$\frac{dq}{dt}=I\cos\omega t.$$

That is, charge at one end of the element is increasing and at the other end decreasing, by the amount of the current flow (coulombs per second). This requires as if there be charge accumulating devices at the two ends of the element, one transferring charge to the other. Therefore, in order to obtain a physical approximation of an isolated current element, one can terminate the current element in two small spheres or discs on which the charge can accumulate. If the wire is very thin compared with the radius of the spheres, then its distributed capacitance would be negligible as compared with the capacitance between the spheres, and then the current in the wire will be uniform. Also, the radii of the spheres should be small compared with $d\mathbf{l}$, and their distance apart. We find that the picture corresponds to that of an oscillating electric dipole as discussed above. The current element will look like as shown in fig. 1 (*a*).

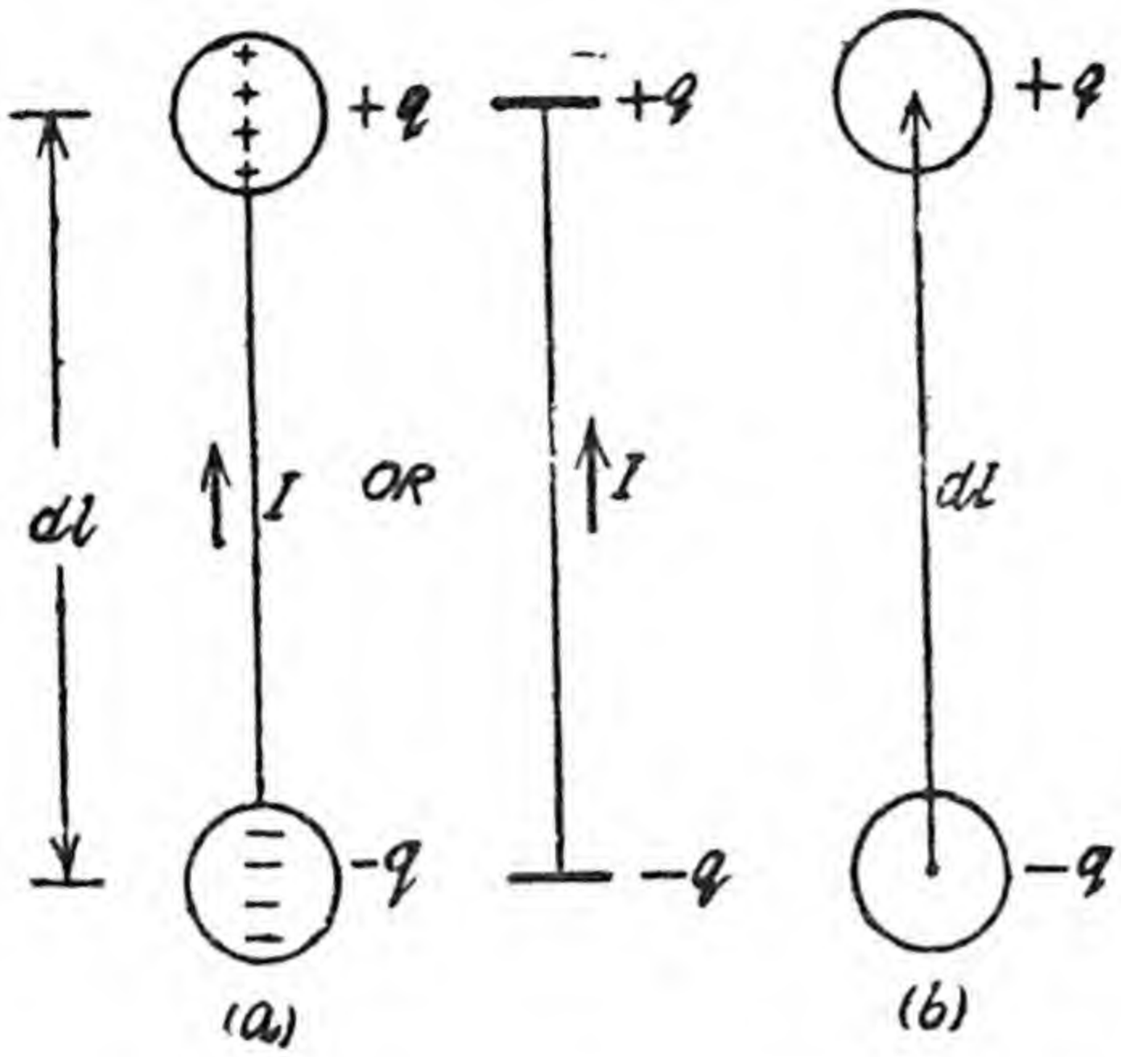


Fig. 1 (a) Current element as an electric dipole (Hertzian dipole).

Fig. 1 (b) Electric dipole. Two opposite and equal charges separated by a small distance.

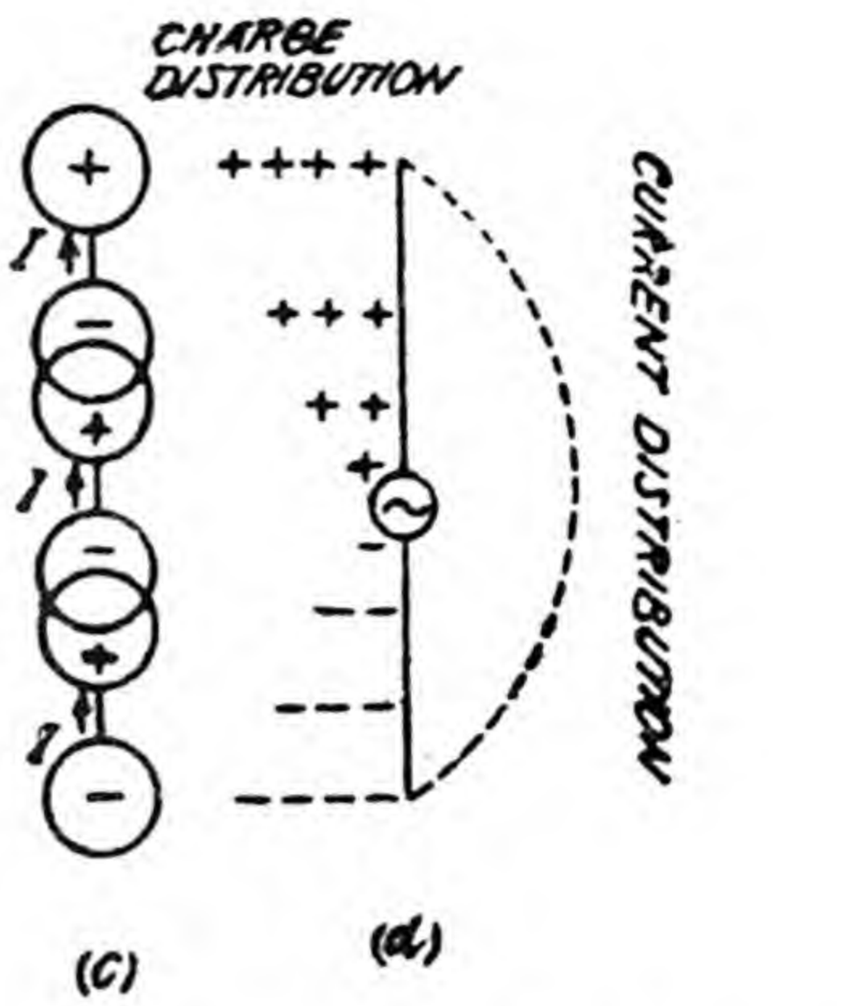


Fig 1 (c) chain of Hertzian dipoles.

Fig. 1 (d) Current and charge distribution on a linear antenna.

A current element which is not so short, say $\lambda/4$ long, can be thought as made up of a chain of Hertzian dipoles. We note that positive charge at one end of one dipole is just cancelled by an equal amount of negative charge at the opposite end of the adjacent dipole (fig. 1*c*). However, if the current along the antenna is not uniform along its length but is distributed as shown in fig. 1 (*d*), then antenna can be represented as a chain of current elements or Hertzian dipoles having slightly different amplitudes. In this case the adjacent charges do not completely cancel, and there is an accumulation of charge on the surface of the wire as shown in fig. 1 (*d*).

In fig. 1 (*a*), we represented current element as an electic dipole. At high frequencies to which the antenna is put, the wire connecting the plates will have a certain amount of self inductance. Therefore, we correctly represent a current element as shown in fig. 1 (*e*). It is fed by a balanced transmission line, as shown.

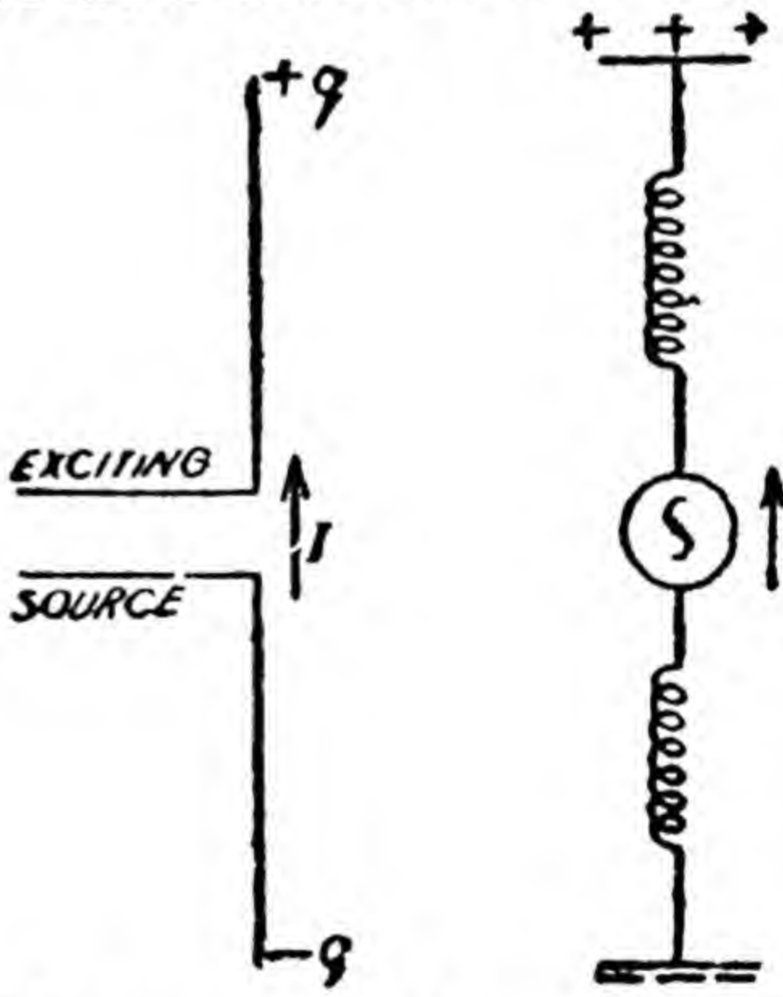


Fig 1 (e) Representation of antenna system.

It is assumed that the transmission line does not radiate and therefore, its presence will be disregarded. Radiation from the end plates is also considered to be negligible. Taking this representation, we shall now proceed to explain that how the energy is released from a current element—the antenna action.

Now we shall consider what changes in the field occur during

one cyele of applied e.m.f. and how the electric field is distributed around such an aerial.*

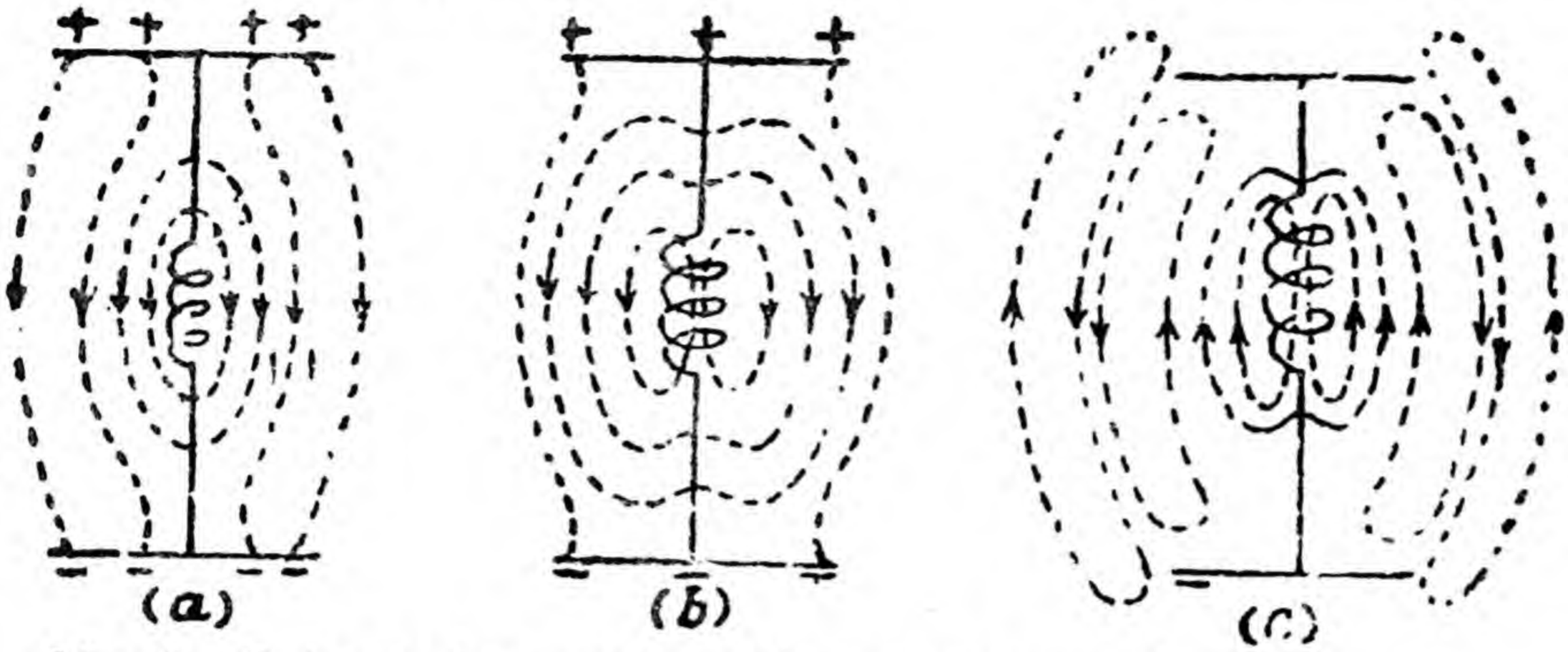


Fig. 2. Release of electromagnetic energy from an antenna system.

Consider that the cycle starts when the condenser is charged to maximum potential difference, so that the upper plate is positive with respect to the lower plate. The current in the wire is zero. At this instant, the field in the vicinity of antenna may be regarded as purely electric. The lines of this electric stress, connecting the positive charge on the upper plate to its opposite number negative charge on the lower plate, can be shown as in fig. 2 (*a*).

After the passage of the moment of maximum potential difference, the electrons start to flow upwards, constituting a current. The electric field begins to collapse which is obvious from the ends of the lines of force coming together along the wire in fig. 2 (*b*). Due to electric inertia, the *current continues to flow even after the potential difference across the condenser is reduced to zero**. As a result the condenser starts to charge in opposite direction. This will set a new electric field between the plates. *If the initial electric field lags on the changes in the potential which causes it, then this new electric field will start to build up before the first one has disappeared.* As a result, the first disturbance is repulsed outward in the form of closed loops by the new electric field as shown in fig. 2 (*c*) because the lines of force have the same direction in the inner surface of initial field and the outer surface of the new electric field.

Ia addition to the electric field, the changes which are shown in fig. 2, we must regard the circuit as being surrounded by rings of changing magnetic stress in a direction normal to the direction of electric field (space quadrature). The intensity of this magnetic stress at any point will vary with the current strength and its direction alternates.

---

*The analogy can be had easily by the mechanism of generation of oscillation in parallel *LC* circuit.

† $V = V_0 \sin \omega t$

$$I = \frac{dq}{dt} = \frac{d}{dt}(CV) = C\omega V_0 \cos \omega t$$

So when V is zero I is not zero.

The study of the process explains that magnetic field is produced only when the electric field is changing, or in other words any changing or moving electric field must have a magnetic field associated with it. Therefore, the movement of the detached closed loops becomes contingent on the association of electric and magnetic fields. These electric and magnetic fields can not exist separately, but bear the same conditions ; if one is zero, then so must the other. Thus *there is no time quadrature between the radiative fields, but essential space quadrature is always maintained.* This mean that radiative fields are in the time phase with each other although at right angles in space as shown in figure 3. It may be pointed out here that inductive magnetic field and electrostatic field in the case of parallel *LC* circuit are in space quadrature as well as in time quadrature with respect to each other.

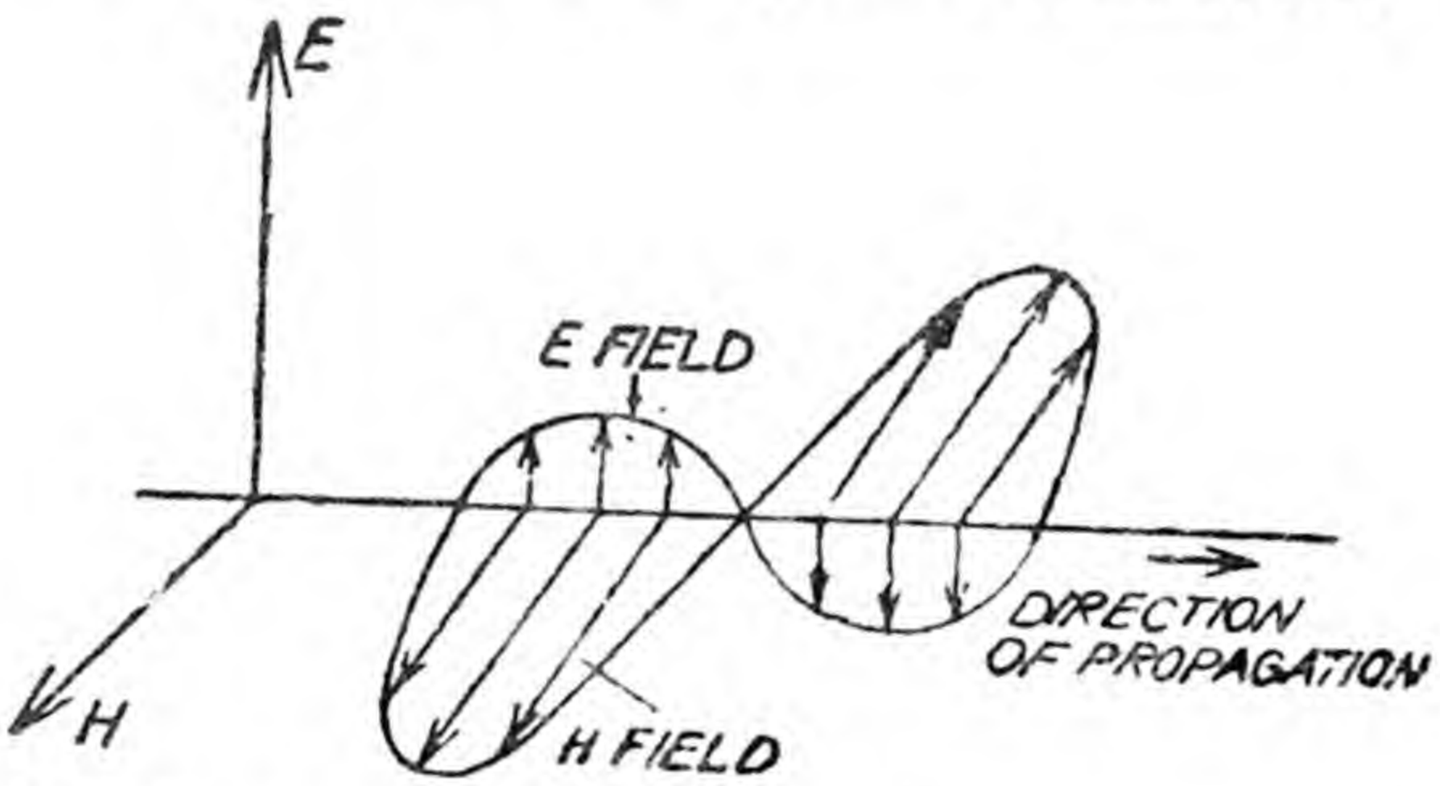


Fig. 3. Distribution of field densities with distance along the direction of propagation.

Fig. 3 also explains the distribution of field intensities with distance along the direction of propagation.

As the direction of current reverses periodically, under the influence of alternating e.m.f., the fields expand and contract, and a substantial amount of energy is released as electromagnetic waves.

It the antenna is fed with radio frequency energy (carrier wave), then the current will change the periodicity several thousand times in a second. This results in effective radiation from antenna The waves are polarised vertically if the antenna is vertical with respect to earth, and horizontally if the antenna is horizontal.

The above explanation thus gives an idea of the way in which a portion of the energy is detached from an antenna and radiated into space. It is obvious that *the detachment of the energy occurs whenever the current in the circuit changes.*

The release of energy from an antenna can also be understood on quantum lines When a radio frequency alternating e.m f. of high magnitude is applied to the antenna circuit, then the orbits of the planetary electrons of the atom of the dielectric around the aerial wire are subjected to strong straining action—first in one direction and then in opposite direction. The result is that many of them are knocked from their normal orbits to the orbits nearer

the central nucleus of the atom. Corresponding to each electron knocked into a new orbit, a quantum of energy is radiated, which then propagates with the velocity of 1,86,000 miles/sec.

28·2. SHORT ELECTRIC DOUBLET :

A short linear conductor is often called a short doublet*. Since any linear antenna can be considered as consisting of a large number of very short conductors in series, it will be therefore useful to study first the radiation properties of short conductors.

(a) Radiative field strengths due to short doublet : A complete discussion will be given in art. 28·3. Here we shall calculate the radiative magnetic field strength at a point *P* due to short doublet. Let the length of the doublet be *dl*, that is short as compared with wavelength λ. We assume that the capacitance associated with each end of the doublet, is such that the current through its length everywhere has the same value *I*.

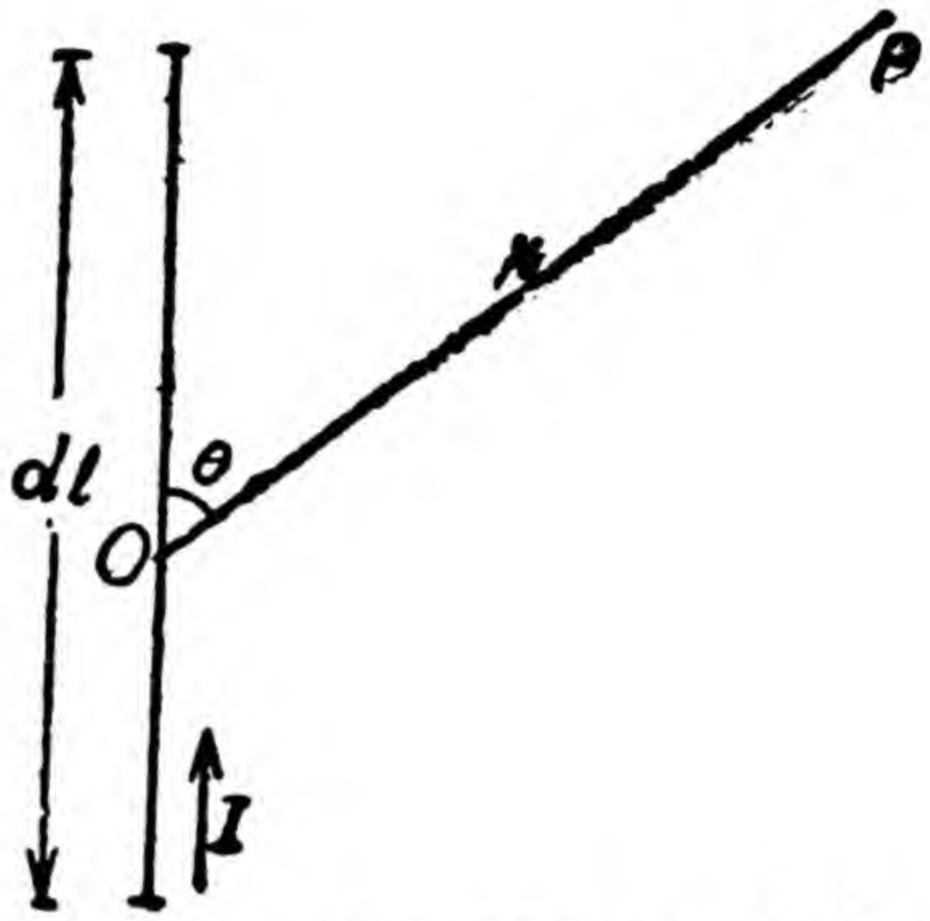


Fig. 4. Short doublet length *dl* carrying a current *I*.

The potential at a point distant *r* from the centre of the doublet can be represented in the form

$$dA=\frac{I\,dl\,\sin\theta}{r} \qquad \ldots(1)$$

The magnetic field strength due to doublet can be found by differentiating (1) with respect to *r*, *i.e.*,

$$dH=-\frac{d}{dr}\,(dA). \qquad \ldots(2)$$

The current at any instant in the doublet is given by

$$I=I_0\sin\omega t.$$

Since a disturbance originated at *O* on the doublet at the time *t* will arrive at point *P* at a time *r*/*c* seconds later where *c* is the velocity of light, the potential at *P* can be expressed as

$$dA=\frac{I_0\sin\omega\,(t-r/c)\,dl\,\sin\theta}{r}.$$

Substituting this value of potentiai in equation (2), we should have

$$dH=-\frac{d}{dr}\left[\frac{I_0\sin\omega\,(t-r/c)\,dl\,\sin\theta}{r}\right]$$

$$=I_0\,dl\,\sin\theta\left[\frac{\sin\omega\,(t-r/c)}{r^2}+\frac{\omega\cos\omega\,(t-r/c)}{rc}\right]. \qquad \ldots(3)$$

In equation (3) the first term gives the induction field. This

*A doublet consists of a length of conductor (antenna) short compared with a wavelength, and capacity areas associated with its ends are so large as to keep the current same at every point of the conductor.

term represents the energy stored in the magnetic field surrounding the conductor. This energy is alternately stored in the field and returned to the source during each half cycle. The induction field diminishes as the square of the distance and is only appreciable in the vicinity of the conductor. Therefore for transmission purposes we do not consider this field.

The second term, called the radiation field, can be represented by the expression

$$dH=\frac{I_0\, dl \sin\theta\, \omega \cos\omega\,(t-r/c)}{rc}$$

$$=\frac{2\pi\, I_0\, dl \sin\theta}{\lambda r}\cos\omega\left(t-\frac{r}{c}\right) \text{ in c.g.s. system of units.}$$

In the above expression current is in *ab* amperes or in e.m.u. while dl, λ and r are in centimeters. Expressing current in amperes, the radiative magnetic field in e.m.u. is given by the expression,

$$dH=\frac{0{\cdot}2\pi I_0\, dl \sin b}{\lambda r}\cos\omega\left(t-\frac{r}{c}\right) \text{ oersted} \qquad \ldots(4)$$

or in M.K.S. system of units

$$dH=\frac{I_0\, dl \sin\theta}{2\lambda r}\cos\omega(t-r/c) \text{ amp. turn/meter.} \qquad \ldots(4a)$$

Note that this magnetic field is maximum when current is minimum. We have also interpreted in previous article that when current is minimum, electrostatic field is maximum. This implies that radiative magnetic field, H and electrostatic field, E, both are maximum at the same instant *i.e.* no time quadrature.

The unit may be amp. turn/meter or amp. turn/cm. depending upon whether dl, λ and r are measured in meters or in centimeters.

Associated with this radiative magnetic field, there is an equivalent electrostatic field dE in time phase with it. In m.k.s. system of units

$$dH=\frac{I_0\, dl \sin\theta}{2\lambda r}\cos\omega(t-r/c) \text{ amper turn/meter.}$$

The corresponding electric field will be

$$dE=\sqrt{(\mu_0/\epsilon_0)}\, dH=377\, dH=120\pi\, dH,$$

where μ_0 and ϵ_0 are the free space permeability and permittivity. Then

$$dE=120\pi.\frac{I_0\, dl \sin\theta}{\lambda r}\cos\omega(t-r/c)$$

$$=\frac{60\pi I_0\, dl \sin\theta}{\lambda r}\times\cos\omega(t-r/c) \text{ volt./meter} \qquad \ldots(5)$$

(b) Power Pattern : The plot of power density dP of the field as a function of dependent co-ordinates in a particular co-ordinate system is termed as radiation pattern of the radiating element.

The power density $dP = dE\, dH = (dE)^2/120\pi$ [since in m.k.s. system $dE = 120\pi\, dH$] and therefore, a plot of antenna pattern in terms of dP will convey the same information as a plot of the magnitude of dE. When the radiation is expressed as field strength, E volt per meter, the radiation pattern is a field strength pattern. If the radiation in a given direction is expressed interms of power per unit solid angle, the resulting pattern is a power pattern. A power pattern is proportional to the square of the field strength pattern.

Consider that an elemental dipole or doublet is located at the centre of a spherical co-ordinate system. Then dP will be the function of the co-ordinates r, θ and ϕ. If we want to determine power through the surface of a sphere at so ne fixed distance r, then dP will be a function of θ and ϕ. Since $dP = dE^2/377$, the measurement of electric field intensity over the surface of a sph :re at some fixed distance r will also determine the radiation p .ttern.

From equations (4) and (5) it is obvious that $dP = dE\, dH$ is a function of r and θ and hence independent of third co-ordinate ϕ. The plot will be as shown in fig. 5. In XZ plane, since θ varies, pattern varies with different values of r and θ but in YZ plane, pattern is symmetrical since dP is independent of ϕ.

Figure 5 represents the radiation pattern of an elementary doublet.

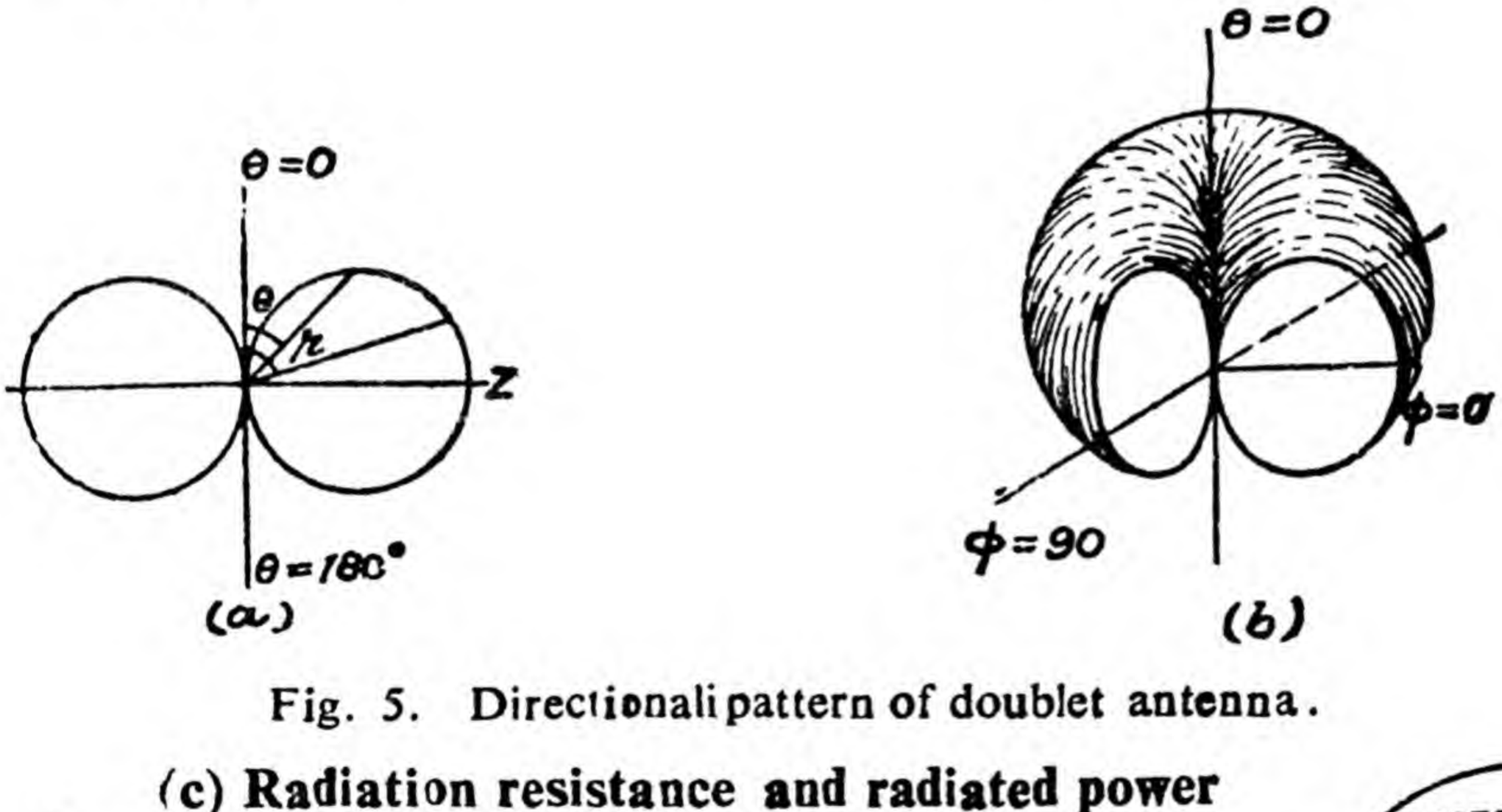


Fig. 5. Directionali pattern of doublet antenna.

(c) Radiation resistance and radiated power for short electric doublet : The power radiated by a doublet can be calculated by assuming it to be placed at the centre of a large sphere, the surface of which will be penetrated continuously by the outgoing moving fields. It has been assumed that the doublet is radiating as a point source located at the centre O so that it radiates uniformly in all directions and there is complete absence of any directional field pattern.

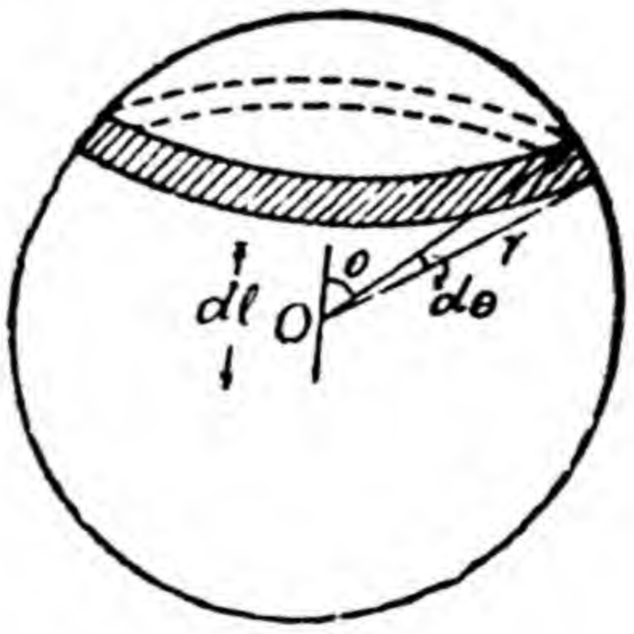


Fig 6. Power calculation radiated by doublet of length dl.

From Poynting's theorem, the mean energy crossing unit area at a point A, in m.k.s. system of units. is

$= EH$ watts/sq. metre.

$= 120\pi H^2$ watts/sq. metre.

$= 120\pi H^2$ watts/sq. cm.

(provided dl, λ and r are in cms.)

∴ Mean rate of energy flow through unit area at A

$$= 120\pi (\overline{H^2})$$

$$= 120\pi \times \left[\frac{I_0\, dl \sin\theta \cos\omega(t - r/c)}{2\lambda r} \right]^2$$

watts/sq. cm.

$$= 30\pi \frac{I_0^2\, dl^2 \sin^2\theta}{\lambda^2 r^2} \times \overline{\cos^2 \omega(t - r/c)}$$

$$= 30\pi \frac{I_0^2\, dl^2 \sin^2\theta}{\lambda^2 r^2} \times \tfrac{1}{2}.$$

Some mean rate or flow of total energy or power through sphere will be

$$= 30\pi \times \frac{I_0^2\, dl^2}{\lambda^2 r^2} \times \tfrac{1}{2} \times 2 \int_0^{\pi/2} 2\pi r^2 \sin^3\theta\, d\theta$$

$$= \frac{30.\pi I_0^2.dl^2}{\lambda^2 r^2} \times 2\pi r^2 \times \tfrac{2}{3}$$

$$= \frac{40\pi^2.I_0^2\, dl^2}{\lambda^2} \text{ watts.}$$

$$\therefore \quad dP = \frac{40\pi^2.I_0^2\, dl^2}{\lambda^2} \text{ watts.}$$

If we substitute the r.m.s. value of current $I_{r.m.s.}$ in place of I_0 which are related by

$$\frac{I_0}{\sqrt{2}} = I_{r.m.s.}$$

then

$$dP = \frac{80\pi^2 I^2_{r.m.s.}\, dl^2}{\lambda^2} \text{ watts.} \qquad \ldots(6)$$

(d) Radiation Resistance : If the radiation resistance of the short doublet antenna is denoted by dR, then power radiated can be expressed as

$$dP = I^2_{r.m.s.}\, dR \qquad \ldots(7)$$

Equating eqs. (7) and (8), we find that

$$dR = 80\pi^2 \left(\frac{dl}{\lambda}\right)^2 \text{ ohms} \qquad \ldots(8)$$

The above expression gives the radiation resistance of a short doublet antenna.

If $\dfrac{dl}{\lambda} = \dfrac{1}{10}$, then $dR = 7{\cdot}9$ ohms :

If $\dfrac{dl}{\lambda} = \dfrac{1}{100}$, then $dR = 0{\cdot}08$ ohm.

Thus the radiation resistance of a short doublet is small. In developing all the above relations for short doublet it has been assumed that $\lambda \gg dl$. This made it possible to neglect the phase difference of field contribution from different parts of the doublet.

28·3. RADIATION FROM A CURRENT ELEMENT:

In art. 28·2 we have discussed it which we are again taking up here.

Let us consider a thin element of length l, area of cross-section da and carrying the current i. Then the magnetic vector potential developed at a point distant r from the element will be

$$\mathbf{A}=\mu \int \frac{J\ dv}{4\pi r}, \qquad \ldots(1)$$

where J is the current density throughout the volume dv, μ is free space permeability. The total current, i, in the thin element of ler gth dl will be

$$i=J\ da,$$

But $$dv=da\ dl,$$

so that $$i=J\frac{dv}{dl}$$

or $$i\ dl=J\ dv.$$

Therefore from equation (1), we get

$$\mathbf{A}=\mu \int \frac{i\ dl}{4\pi r}. \qquad \ldots(2)$$

Integration of above equation implies that the potentials due to various elements of current $i\ d\mathbf{l}$ must be summed up. Since the velocity of propagation is finite, any disturbance originated at current element will reach a point distant r meters from the element after r/c seconds, where c is the velocity of light with which the electromagnetic radiation propagates. Therefore, if the current in element is sinusoidal

$$i=I_0 \sin \omega t,$$

it should be written as

$$i=I_0 \sin \omega(t-r/c),$$

for the reasons stated above. The expression (2) can then be expressed as

$$A=\mu \int \frac{I_0 \sin\ \omega\ (t-r/c)\ d\mathbf{l}}{4\pi r}. \qquad \ldots(3)$$

Fig. 7 (*a*) illustrates the element of current of length dl at the origin with spherical co-ordinates r, θ and ϕ. Though if we follow the convention used in art. 28·2, we must write quantities for small element of length dl by $d\mathbf{A}$ for magnetic vector potential. $d\mathbf{H}$ for magnetic field and $d\mathbf{E}$ for electric field but here it will complicate the analysis and therefore we shall denote them by **A**, **H** and **E** even

for short dipole. For such an element. equation (3) in y-direction can be written as

$$A_y=\frac{I_0 \sin \omega\,(t-r/c)\,dl}{4\pi r}. \qquad ...(4)$$

because it is the direction of current producing vector potential at point P distant r from the origin O.

Calculation of Radiative Magnetic field : Since A is completely along y-axis,

$$A_x=A_z=0.$$

so that from

$$\mathbf{B}=\mu\mathbf{H}=\nabla\times\mathbf{A}$$

we get

$$\left.\begin{aligned}\mu H_x&=-\frac{\partial A_y}{\partial z}\\ \mu H_y&=0\\ \mu H_z&=\frac{\partial A_y}{\partial x}\end{aligned}\right\} \qquad ...(5)$$

They give the components of magnetic field.
From fig. 7 (a) it can be seen that

$$\left.\begin{aligned}r&=(x^2+y^2+z^2)^{1/2}\\ x&=r\sin\theta\sin\phi\\ y&=r\cos\theta\\ z&=r\sin\theta\cos\phi\end{aligned}\right] \qquad ...(6)$$

From equations (4), (5) and (6), we have

$$\mu H_x=-\frac{\partial}{\partial z}\left[\frac{\mu\, I_0\, \sin\omega\left(t-\frac{\sqrt{(x^2+y^2+z^2)}}{c}\right)dl}{4\pi\sqrt{(x^2+y^2+z^2)}}\right]$$

$$H_x=-dl\left[\frac{-\frac{1}{2}(x^2+y^2+z^2)^{-3/2}\times 2z}{4\pi}\times I_0\sin\omega\left\{t-\frac{\sqrt{(x^2+y^2+z^2)}}{c}\right\}\right.$$

$$\left.+\frac{I_0}{4\pi\sqrt{(x^2+y^2+z^2)}}\times\cos\omega\left\{t-\frac{\sqrt{(x^2+y^2+z^2)}}{c}\right\}\times\frac{-\frac{1}{2}(\omega/c)\,2z}{\sqrt{(x^2+y^2+z^2)}}\right]$$

or $$H_x=-\frac{dl\, I_0}{4\pi}\left[\frac{-z}{r^3}\sin\omega\,(t-r/c)-\frac{\omega z}{cr^2}\cos\omega\,(t-r/c)\right] \qquad ...(7)$$

Similarly, from equations (4), (5) and (6) we can write

$$H_z=\frac{I_0\, dl}{4\pi}\left[\frac{-x}{r^3}\sin\omega\,(t-r/c)-\frac{\omega x}{cr^2}\cos\omega\,(t-r/c)\right] \qquad ...(8)$$

H_x and H_z lie in the same plane, horizontal here, and are perpendicular to each other. Angle in this plane is denoted by ϕ. The only spherical coordinate field component will be H_ϕ, given by

$$H_\phi=H_x\cos\phi-H_z\sin\phi \qquad \text{(see fig. 7-b)} \qquad ...(9)$$

Putting in equations (7) and (8), the value

$$\frac{z}{r}=\sin\theta\cos\phi \text{ and } \frac{x}{r}=\sin\theta\sin\phi$$

and then substituting the values of H_x and H_z in equation (9), we get

$$H_\phi=\frac{I_0\,dl}{4\pi}\left[\frac{\sin\theta\cos^2\phi}{r^2}\sin\omega\,(t-r/c)+\frac{\omega\sin\theta\cos^2\phi}{cr}\cos\omega(t-r/c)\right.$$
$$\left.+\frac{\sin\theta\sin^2\phi}{r^2}\sin\omega\,(t-r/c)+\frac{\omega\sin\theta\sin^2\phi}{cr}\cos\omega\,(t-r/c)\right]$$

or $$H_\phi=\frac{I_0\sin\theta\,dl}{4\pi}\left[\frac{\sin\omega\,(t-r/c)}{r^2}+\frac{\omega}{cr}\cos\omega\,(t-r/c)\right]$$

The first term represents induction field and is not appreciable at large distances from the element for it varies as $1/r^2$. For communication over long distance only the second term is important and is called *radiative magnetic field*. Therefore

$$\boxed{H_\phi=\frac{I_0\sin\theta\,dl}{4\pi}\cdot\frac{\omega}{cr}\cdot\cos\omega(t-r/c).} \qquad ...(10)$$

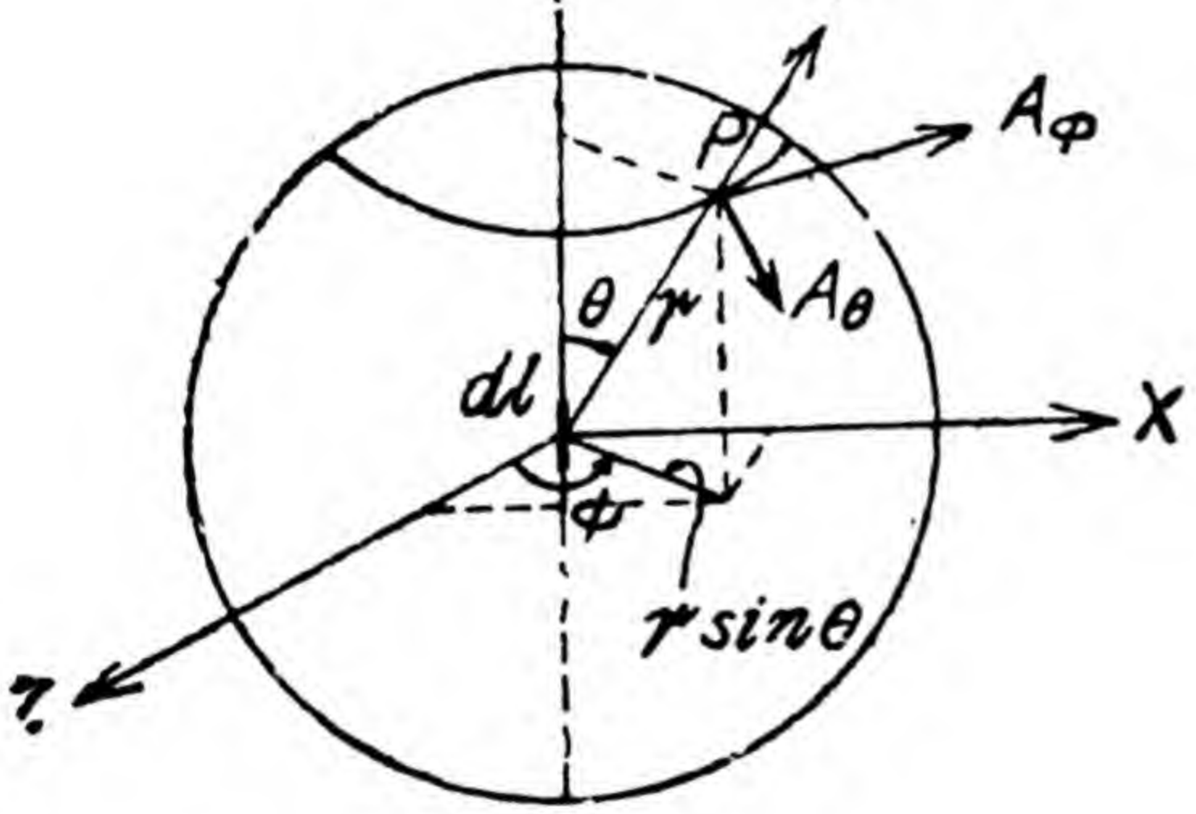


Fig. 7 (*a*)

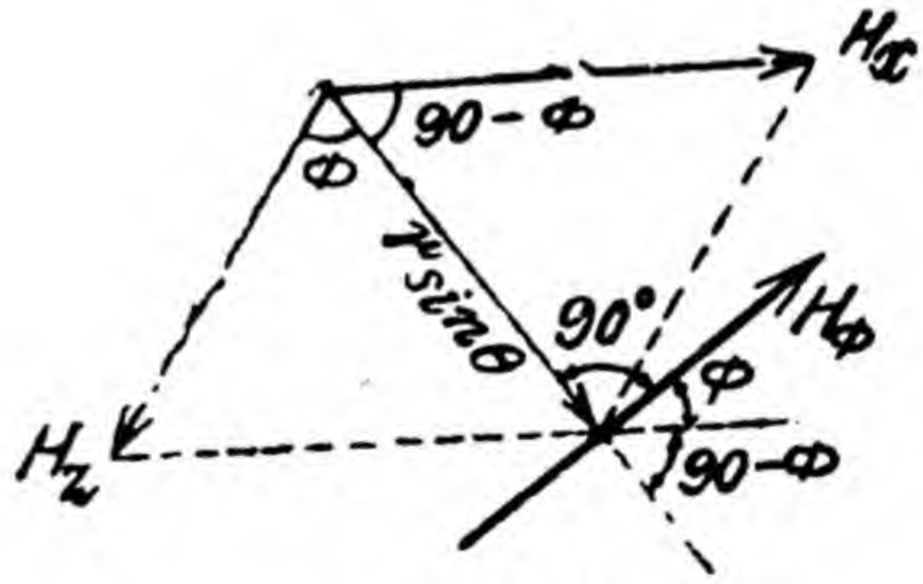


Fig. 7 (*b*)

Calculation of Radiative Electric field : The expression for corresponding electric field can be derived on using Maxwell's field equations with the following conditions :

(*i*) Since $H_y=0$, magnetic field lies purley in the horizontal plane. So that

$$H_\theta=H_r=0$$

and only H_ϕ exists.

(*i*) Field is symmetrical around current element in XZ plane (horizontal plane), so that

$$\frac{\partial}{\partial\phi}=0.$$

Then from Maxwell's field equations, we can arrive at

$$\epsilon \frac{\partial E_r}{\partial t} = \frac{1}{r \sin \theta} \left[\frac{\partial}{\partial \theta} (H_\phi \sin \theta) \right] \quad \ldots(11)$$

and $$\epsilon \frac{\partial E_\theta}{\partial t} = -\frac{1}{r} \left[\frac{\partial}{\partial r} (rH_\phi) \right], \quad \ldots(12)$$

because electric field will be purely in vertical plane by virtue of H_θ and H_r being zero, ϵ is the free space permittivity.

From equations (10) and (11), we get

$$\epsilon \frac{\partial E_r}{\partial t} = \frac{1}{r \sin \theta} \left[\frac{\partial}{\partial \theta} \left\{ \frac{I_0 \, dl \sin^2 \theta}{4\pi} \frac{\omega}{cr} \cos \omega \left(t - \frac{r}{c} \right) \right\} \right]$$

or $$\frac{\partial E_r}{\partial t} = \frac{2 I_0 \, dl \cos \theta}{4\pi\epsilon} \cdot \frac{\omega}{cr^3} \cos \omega \left(t - \frac{r}{c} \right).$$

Integrating it with respect to time,

$$E_r = \frac{I_0 \, dl \cos \theta}{2\pi\epsilon} \cdot \frac{\sin \omega (t - r/c)}{cr^2}.$$

Since it varies as $1/r^2$, this corresponds to induction magnetic field and is appreciable only in the vicinity of the element. Therefore, for radiation to be effective at far distances, this field is of least importance. Consequently, we neglect E_r.

Further from equation (12), we have

$$\epsilon \frac{\partial E_r}{\partial t} = -\frac{1}{r} \frac{\partial}{\partial r} \left[r \cdot \frac{I_0 \, dl \sin\theta}{4\pi} \cdot \frac{\omega}{cr} \cdot \cos \omega \left(t - \frac{r}{c} \right) \right]$$

or $$\frac{\partial E_\theta}{\partial t} = -\frac{I_0 \, dl \sin\theta}{4\pi\epsilon} \cdot \frac{\omega^2}{c^2 r} \sin \omega \left(t - \frac{r}{c} \right).$$

Integrating above expression with respect to t, we get

$$\boxed{E_\theta = \frac{I_0 \omega \, dl \sin \theta}{4\pi\epsilon} \cdot \frac{\cos \omega (t - r/c)}{c^2 r}.} \quad \ldots(13)$$

As it varies as $1/r$, E_θ corresponds to radiative magnetic field and is appreciable over large distances. This can be termed as *radiative electric field.*

Intrinsic impedance of space : If we compare equation (10) and (13), we can arrive at the conclusion that there is no time quadrature between H_ϕ and E_θ. Of course, they maintain the space quadrature. Also

$$\frac{E_\theta}{H_\phi} = \frac{1}{c\epsilon} = \frac{1}{\epsilon . 1/\sqrt{(\mu\epsilon)}} = \sqrt{\left(\frac{\mu}{\epsilon}\right)}$$

$$= 120\pi \text{ ohms}$$

and is known as the intrinsic impedance of the free space.

$$\boxed{E_\theta = 120\pi H_\phi.}$$

28·4. THIN LINEAR ANTENNA (Approximate analysis) :

(a) Current and Voltage Distribution :

be made of a large number of very short conductors connected in series.

In fig. (8-1) an aerial having length equal to $\lambda/2$ is shown.

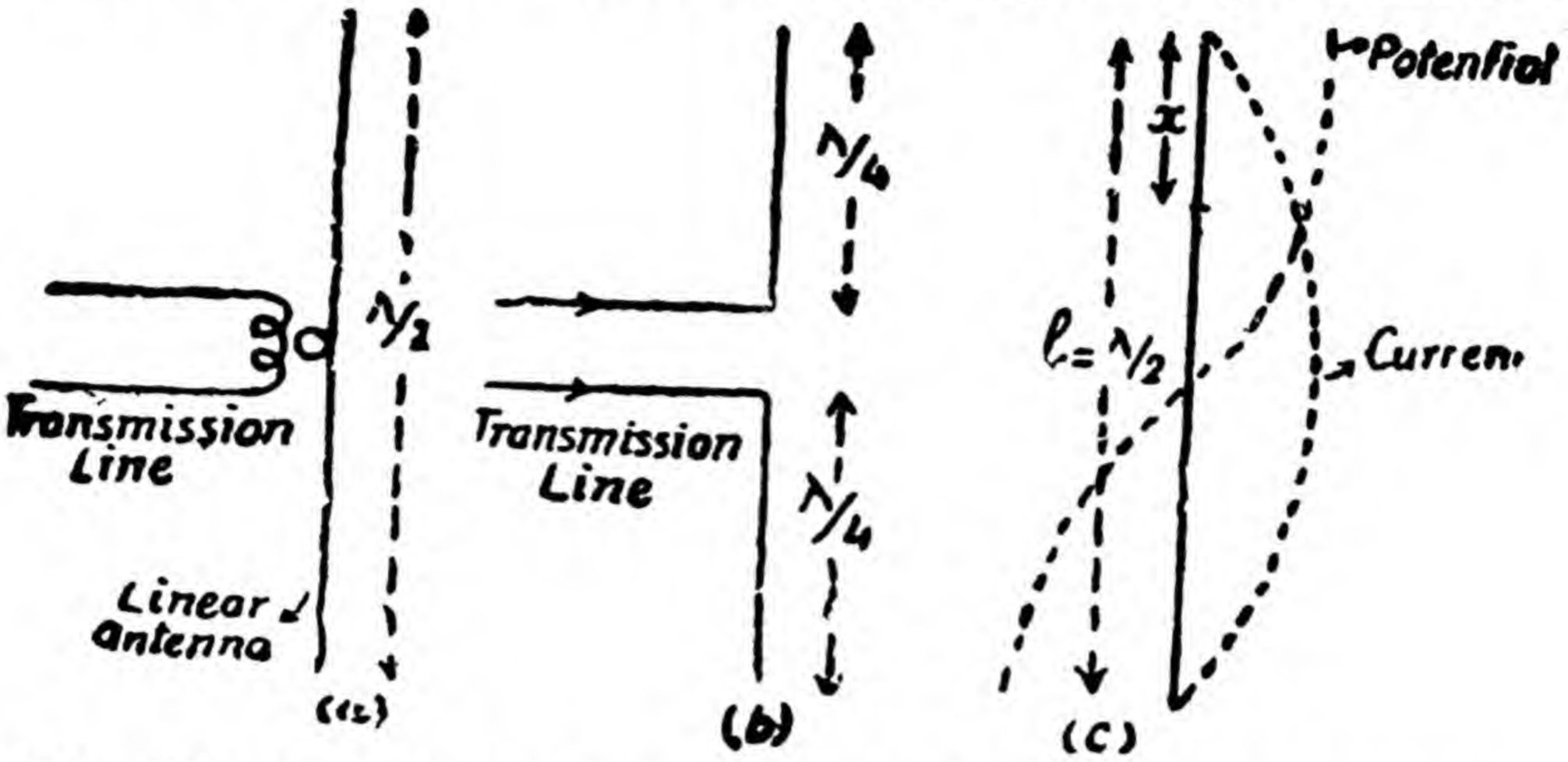


Fig. 8-1. Hertz aerial (half wave dipole) and distribution of potential and current in it.

The two ends of the aerial are at opposite potentials. The r.f. energy from the transmitter through transmission line is fed at its centre. Such a conductor, supported on insulators so as to be away from the ground, will have self inductance and self capacitance depending upon its length. At the open ends the current will be zero, *i.e.* current nodes exist at these ends For antenna to be symmetrical and excited in basic mode, the current maximum (*i.e.* antinode) should exist at its centre. The potential and current distribution in half wave dipole antenna is shown in fig. (8-1).

The current and voltage distributions on open ended long antenna wires are basically similar to the standing waves of current and voltage on an open ended two wire transmission

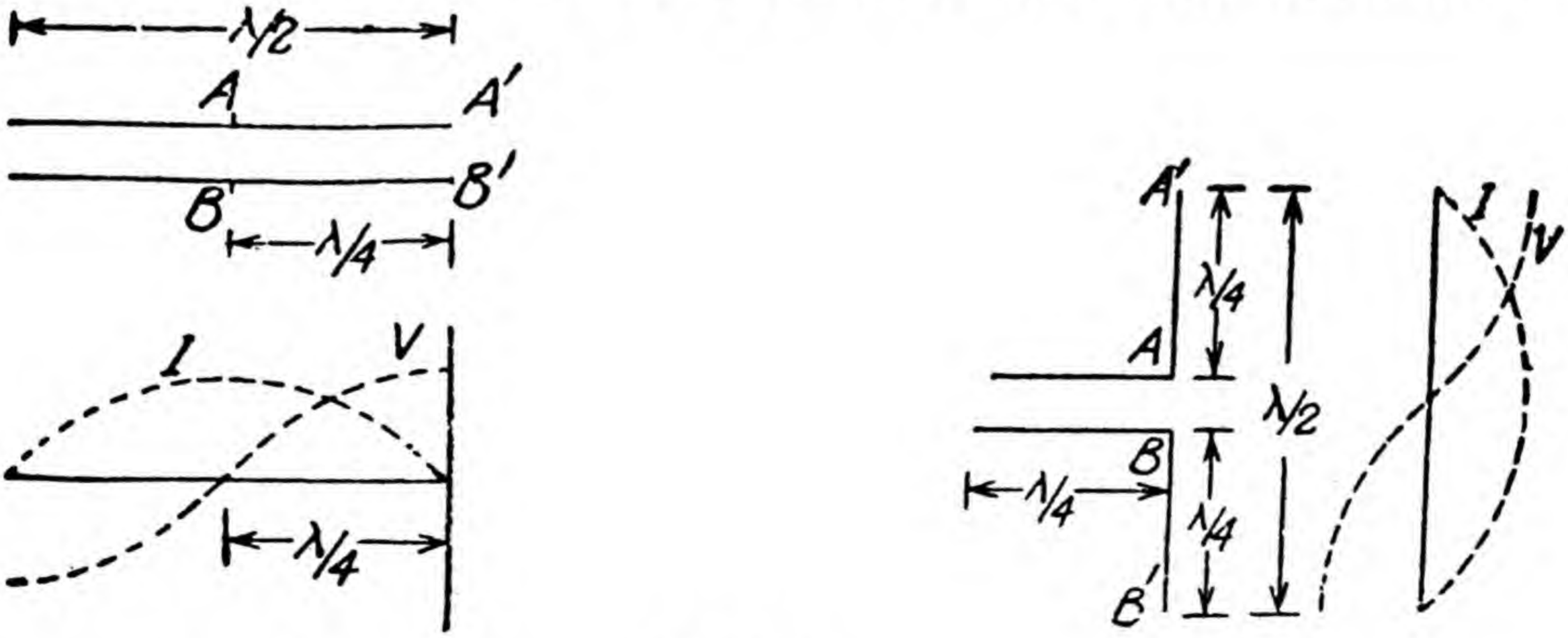


Fig 8-2.

line. Refer to art. 28·3 case (ii). For example, let us consider a half wave long open circuited transmission line [fig. 8·2 (*a*)]. We give a turn to AA' and BB' portion of the line so as to assume the shape shown in fig. 8-2 (*b*). Obviously, it constitutes

a half wave antenna fed with quarter wave line. Therefore, voltage and current distribution for half wave open circuited transmission line will be applicable to a half wave antenna.

Similarly, we can predict the voltage and current distributions on antenna of any length.

By this analogy, we get only approximate description of sinusoidal distributions of voltage and current on the antenna. In fact, antenna wires are not equivalent to a uniform transmission line because the radiation resistance as well as any actual resistance in the antenna wire results in a small component of current that is in phase with the voltage, rather than 90° out of phase which happens in case of uniform transmission line. But it has been found experimentally that the current distribution is almost sinusoidal for linear antennas that are quite thin compared to their length.

(b) Radiation Length (effective length) of Linear Antenna : In the previous deductions for short electric doublet, it was assumed that current through the doublet is uniform and given by $I = I_0 \sin \omega t$ at time t over the whole length of the doubet. This assumption, however cannot be made for the dipole aerial having length l, close to wavelength order, *e.g.*, quarter wave or half wave dipole.

If for a moment we assume that the amplitude of oscillatory current is same at all point in the radiating wire (antenna), then the expression for the radiative magnetic field H can be written as

$$H = \frac{I_0 l \sin \theta}{2\lambda r} \cos \omega \left(t - \frac{r}{c} \right).$$

$$| H | = \frac{I_0 \sin \theta \, l}{2\lambda r}. \qquad \ldots(1)$$

The above relation is not true because it is based on the assumption that amplitude and phase of oscillatory current is same at all points in the antenna which is not the case in practice as can be seen from current distribution given in figure 8-1 (*c*). The current distribution from the centre to the upper end may be either approximately (*a*) linear. or (*b*) sinusoidal. *The effect reduces the amount of power radiated and makes the aerial equivalent to a shorter one with the same current.* 'The actual length of equivalent shorter aerial' may easily be assessed by finding the 'mean value' of current over the length concerned and then expression (1) should be written as

$$| H | = \frac{I_{av} \, l \sin \theta}{2\lambda r}. \qquad \ldots(2)$$

Case (a) Linear distribution of current along the length of aerial : It is obvious that the 'equivalent length' or radiation length, will be $\frac{l}{2}$ because the mean value of current is $\frac{I_0}{2}$, *i.e.* putting $I_{av} = \frac{I_0}{2}$ in equation (2)

$$|H| = \frac{l \sin\theta}{2\lambda r} \cdot \frac{I_0}{2}$$
$$= \frac{I_0 \sin\theta}{2\lambda r} \cdot \frac{l}{2}$$
$$= \frac{I_0 \sin\theta}{2\lambda r} l_e, \qquad \ldots(3)$$

where l_e is the *effective length of the antenna.*

Case (b) When current distribution is sinusoidal :

Referring to figure (8-1), the current is zero at each end and has an instantaneous maximum value at the centre, I_0, at a time t; while the distribution over either half of the antenna length is represented by $I_0 \sin\left(\frac{x}{l/2}\right)\frac{\pi}{2}$, giving instantaneous current at a point distant x from the end ; x lies between 0 and $\frac{l}{2}$. The mean value of this current over half the length of aerial is

$$I_{av} = \frac{I_0 \int_0^{\pi/2} \sin\phi \, d\phi}{\int_0^{\pi/2} d\phi} = \frac{2I_0}{\pi}\Big(-\cos\phi \Big)_0^{\pi/2}$$
$$= \frac{2I_0}{\pi},$$

where $\phi = \left(\frac{x}{l/2}\right)\frac{\pi}{2}$.

Thus, for this case, expression for $|H|$ can be written as

$$|H| = \frac{l \sin\theta}{2\lambda r} \cdot \frac{2I_0}{\pi}$$
$$= \frac{I_0 \sin\theta}{2\lambda r} \cdot I_e, \qquad \ldots(4)$$

where $l_e = \frac{2l}{\pi}$ is the *effective length of the aerial.*

Now, on applying the average value of current, we shall show that how the expressions of power radiated and radiation resistance of doublet antenna can be made to suit the case of an antenna with sinusoidal current distribution.

(c) Total power radiated (by the linear antenna) : For short doublet it was shown that total radiated power is

$$dP = \frac{40\pi^2 I_0^2}{\lambda^2} dl^2 \text{ watts.}$$

On the basis of above expression, the total power P radiated by doublet of length l, when the current distribution is sinusoidal can be obtained by substituting $dl = \frac{2l}{\pi}$, the effective length.

Therefore, $P = \frac{40\pi^2 I_0^2}{\lambda^2}\left(\frac{2l}{\pi}\right)^2$

$$= \frac{160 l^2 I_0^2}{\lambda^2} \text{ watts.}$$

Putting I_0 in terms of $I_{r.m.s.}$, the r.m.s. value of current, we should have $P = \frac{60 l^2 . (\sqrt{2} I_{r.m.s.})^2}{\lambda^2}$

$$= \frac{320 . l^2 I^2_{r.m.s.}}{\lambda^2} \text{ watts.} \quad ...(5)$$

(d) Radiation Resistance (of thin linear antenna) : *It is the resistance which when inserted in series with antenna will consume the same amount of power as is actually radiated.* Its magnitude depends upon the configuration of antenna and upon the point at which it is considered as being inserted in the antenna system. We always *calculate the radiation resistance at a point of current maximum* (at current antinode) *at which the antenna may be fed.* Radiation resistance R can be expressed as the ratio of the power P, radiated by antenna to the current flowing, *i.e.*,

$$R = \frac{P}{I^2_{r.m.s.}}. \quad ...(6)$$

From relation (4) it is found to be equal to

$$R = \frac{320 l^2}{\lambda^2} \text{ ohms.}$$

Special Cases :

(*i*) *For half wave antenna* (*or half wave dipole*) $l = \lambda/2$.

$$R_{\lambda/2} \doteq \frac{320}{\lambda^2} . \left(\frac{\lambda}{2}\right)^2 = 80 \text{ ohms.}^*$$

Therefore the transmission line required to feed this aerial must have characteristic impedance equal to 80 ohms for good matching purposes. This value is roughly constant for $\lambda/2$ aerial remote from the ground.

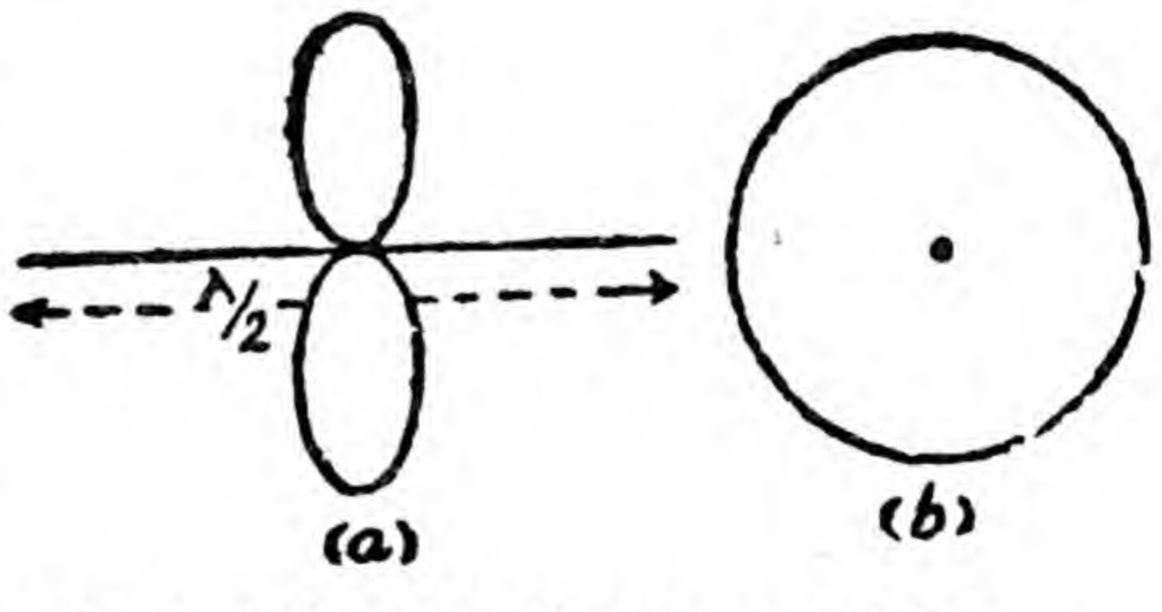


Fig. 9. Directional patterns of the radiation from a half wave antenna

Fig. 9 (*a*) shows radiation pattern representing the field strength in various directions of an open ended half wave antenna in free space. It is obvious that maximum radiation occurrs from the centre of the antenna in directions perpendicular to it. The complete pattern is of the shape which would be formed if fig. 9(*a*) is revolved about the antenna as axis. If we consider radiation pattern in a plane perpendicular to the antenna and passing through its centre, it would be circular as shown in figure 9 (*b*)

*For precise calculation follow art 28·5

(*ii*) *For quarter wave antenna* (*or quarter wave monopole*) : For monopole, we shall have to consider only the upper or lower half of the $\lambda/2$ aerial. The whole aerial is then quarter wavelength long and in this case, thus, the radiated power will be half the previous value, *i.e.*,

$$\frac{160I^2_{r.m.s.}}{\lambda^2}\, l^2 \text{ watts.}$$

Since only upper half of dipole is effective,

$$\frac{l}{2}=\frac{\lambda}{4},$$

$$l=\lambda/2.$$

Therefore, $P_{\lambda/4}=40I^2_{r.m.s.}$ watts.
and radiation resistance $R_{\lambda/4}=40$ ohms† .

28·5. THIN LINEAR ANTENNA IN SPACE :

We shall develop expressions for field patterns of thin linear antenna. The antenna, under consideration, may be of any length, but it is assumed that the current distribution is sinusoidal and it is fed at the centre by a balanced two wire transmission line (shown in fig. (10·1).

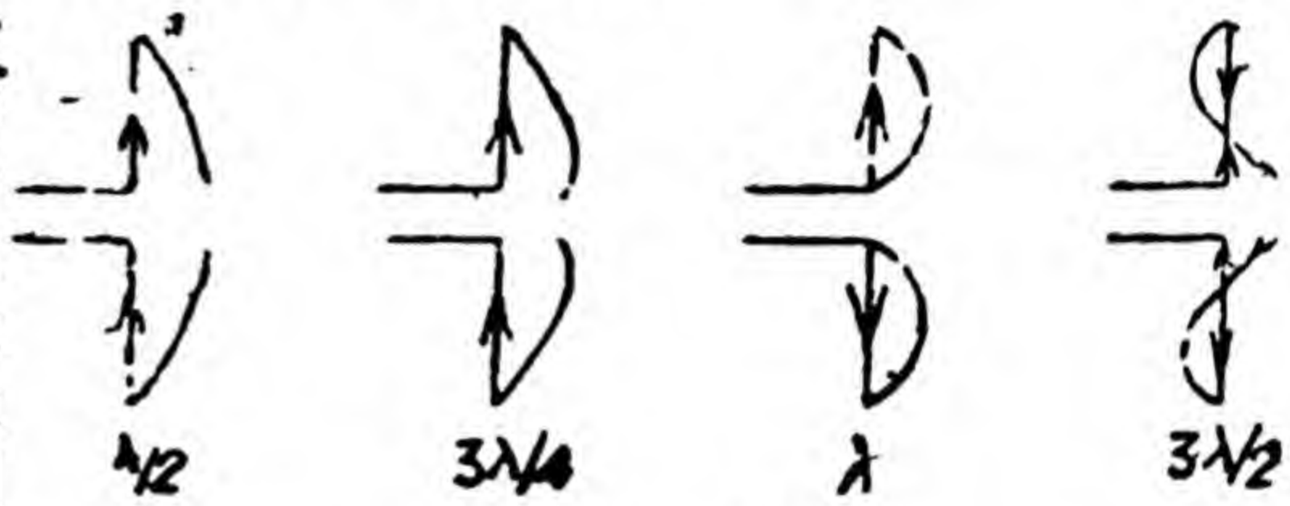


Fig. 10·1

Current and voltage distribution along any of the above atennas (fig. 10–1) can be found by approximating the antenna in question to equivalent length of open circuited transmission line (art 26·3-ii) For example $3\lambda/2$ antenna can be constructed by a line of length $3\lambda/4$.

We find from fig. 10 art. 26·3 that at $270°\left(\beta s=\frac{2\pi}{\lambda}.\frac{3\lambda}{4}=\frac{6\pi}{4}=270°\right)$ voltage is zero whereas current is maximum. Therefore at feed point of $3\lambda/2$ antenna voltage will be zero whereas current will be maximum; and at end points current will be zero where as voltage will be maximum (fig. 10-1).

(a) Calculation of Radiation Fields : The primary aim in discussing antenna system is to calculate the fields radiated by it. Once fields are known, it becomes quite easy to calculate power and radiation resistance

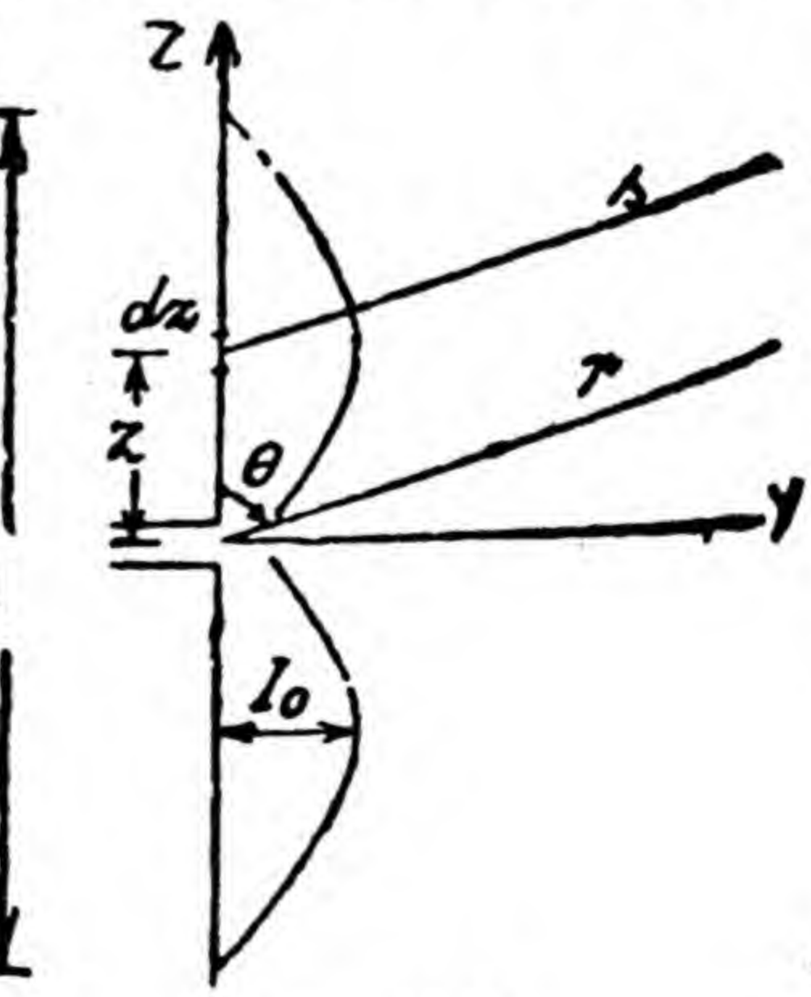


Fig 10·2.

Let us now consider a symmetrical, thin, linear, centre fed antenna of length *l*. The retarded value of the current at any point *z* on the antenna referred to a

†For precise calculation see art. 28·5.

point at a distance s is (fig. 10·2) given by

$$[I]=I_0 \sin \frac{2\pi}{\lambda}\left(\frac{l}{2}\pm z\right).e^{j\omega\,(t-s/c)} \qquad \ldots(1)$$

In above equation, function $\sin \frac{2\pi}{\lambda}\left(\frac{l}{2}\pm z\right)$ is the form factor for the current on the antenna. The expression $\left(\frac{l}{2}+z\right)$ is used when $z<0$ and $\left(\frac{l}{2}-z\right)$ is used when $z>0$. By regarding the antenna as made up of a series of infinitesimal dipoles of length dz, the field of the entire antenna may then be obtained by integrating the fields from all the dipoles making up the antenna. †The far fields dE_θ and dH_ϕ in *mks* units, at a distance s from the infinitesimal dipole dz are

$$dE_\theta=\frac{60\pi\,[I]\sin\theta\,dz}{s\lambda}\ \text{volt/meter.} \qquad \ldots(2)$$

$$dH_\phi=\frac{[I]\sin\theta\,dz}{2s\lambda}\ \text{amp. turn /meter.} \qquad \ldots(3)$$

Since $E_\theta=120\pi\,H_\phi$, it will suffice to calculate H_ϕ. The value of magnetic field H_ϕ for the entire antenna is the integral of equation (3) over the length of the antenna. Thus

$$H_\phi=\int_{-l/2}^{+l/2} dH_\phi. \qquad \ldots(4)$$

Putting the value of $[I]$ from equation (1) into equation (3) and then from equation (4),

$$H_\phi=\frac{I_0\sin\theta\,e^{j\omega t}}{2\lambda}\left\{\int_{-l/2}^{0}\frac{1}{s}\sin\frac{2\pi}{\lambda}\left(\frac{l}{2}+z\right).e^{-j(\omega s/c)}.dz+\int_{0}^{l/2}\frac{1}{s}\sin\frac{2\pi}{\lambda}\left(\frac{l}{2}-z\right).e^{j(\omega s/c)}.dz\right\}. \qquad (5)$$

In equation (5), $1/s$ affects only the amplitude and hence at a large distance it may be regarded as constant. Also at large distance the difference between s and r can be neglected in its effect on the amplitude although its effect on the phase must be considered. Further from fig. 10·2,

$$s=r-z\cos\theta. \qquad \ldots(6)$$

Putting equation (6) into equation (5) and also putting $s=r$ in the amplitude factor, equation (5) becomes

$$H_\phi=\frac{I_0\sin\theta\,e^{j\omega\,(t-r/c)}}{2\lambda r}\left\{\int_{-l/2}^{0}\sin\frac{2\pi}{\lambda}(l/2+z)\,e^{(j\omega z\cos\theta)/c}\,dz+\int_{0}^{l/2}\sin\frac{2\pi}{\lambda}(l/2-z)\,e^{(j\omega z\cos\theta)/c}.dz\right\} \qquad \ldots(7)$$

†Since it is vertical antenna E will be vertical and H horizontal. Therefore, they are written as E_θ and H_ϕ

Since $\beta=\frac{\omega}{c}=\frac{2\pi}{\lambda}$ equation (7) may be written as

$$H_\phi=\frac{I_0 \sin\theta . e^{j\omega(t-r/c)}}{2\lambda r}\left\{\int_{-l/2}^{0} e^{j\beta z\cos\theta}.\sin\beta\left(\frac{l}{2}+z\right).dz + \int_{0}^{l/2} e^{j\beta z\cos\theta}.\sin\beta\left(\frac{l}{2}-z\right).dz\right\} \quad ...(8)$$

$$=\frac{I_0 \sin\theta . e^{j\omega(t-r/c)}}{2\lambda r}\left[\frac{\cos\left(\frac{\beta l\cos\theta}{2}\right)-\cos\frac{\beta l}{2}}{\frac{\pi}{\lambda}\sin^2\theta}\right] \quad ..(9)$$

$$=\frac{[I_0]}{2\pi r}\left[\frac{\cos\left(\frac{\beta l\cos\theta}{2}\right)-\cos\frac{\beta l}{2}}{\sin\theta}\right], \text{ amp. turn/meter} \quad ..(10)$$

where $[I_0]=I_0\, e^{j\omega(t-r/c)}$.

Hence $E_\theta=120\pi\, H_\phi$

$$E_\theta=\frac{60\,[I_0]}{r}\left[\frac{\cos\left(\frac{\beta l\cos\theta}{2}\right)-\cos\frac{\beta l}{2}}{\sin\theta}\right], \text{ volt/meter} \quad ...(11)$$

Equations (10) and (11) are the expressions for the far fields, H_ϕ and E_θ of a symmetrical, centre fed, thin linear antenna of length l. From equations (10) and (11), we can draw the field patterns for the antenna concerned.

(b) Field Patterns for particular cases:

(*i*) *½ wavelength antenna*: When $l=\lambda/2$, the pattern factor becomes

$$E_\theta=\frac{\cos(\pi/2\cos\theta)}{\sin\theta}. \quad ...(12)$$

This pattern in vertical plane is shown in fig. 10·3. It is only slightly more directional than the pattern of an infinitesimal or short

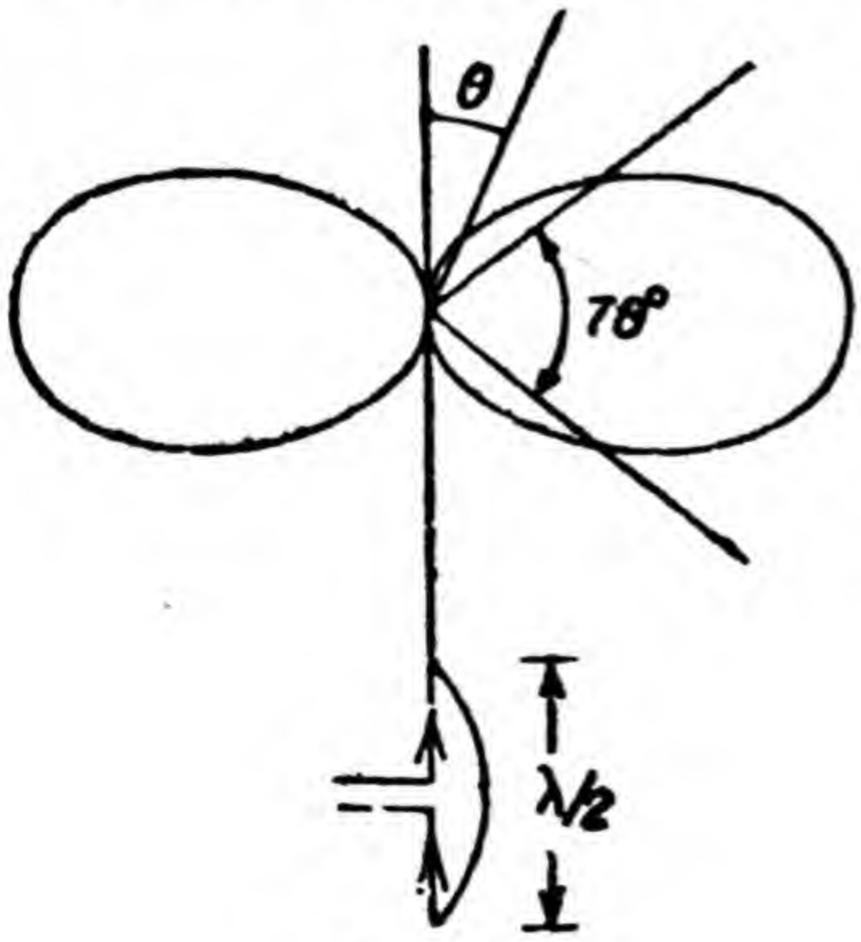


Fig. 10·3.

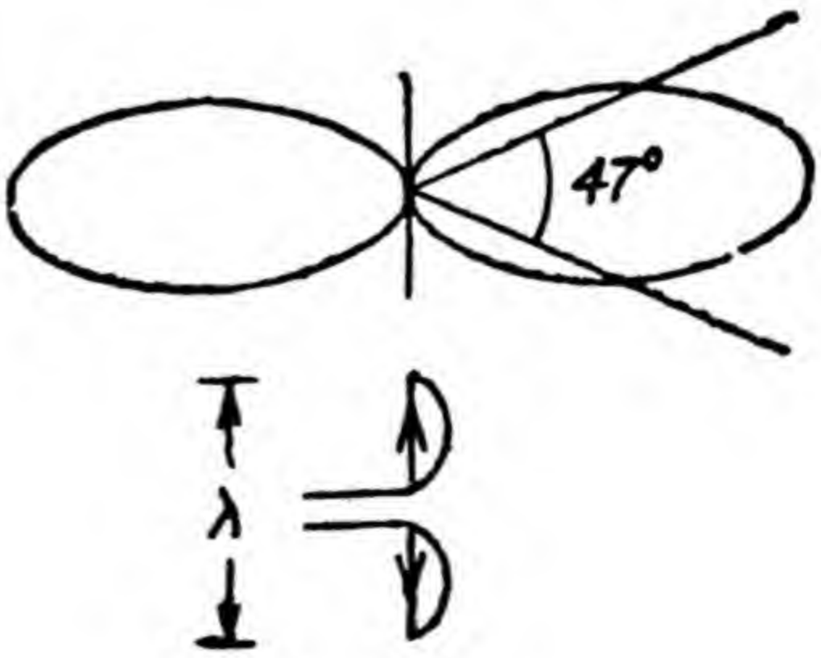


Fig. 10·4.

dipole which is given by sin θ. The beamwidth between half power points of $\lambda/2$ antenna is 78° as compared to 90° for the short dipole.

In horizontal plane $\theta=\pi/2$, form factor is unity, so that pattern is circular.

(*ii*) *Full wave antenna* : When $l=\lambda$, the pattern factor becomes

$$E_\theta=\frac{\cos(\pi\cos\theta)+1}{\sin\theta}. \qquad ...(13)$$

This pattern is shown in fig. 10·4. This is in vertical plane. For horizontal plane $\theta=90°$, we find that form factor is unity so that

$$E_{90}=\frac{60I_0}{r} \quad \text{or} \quad E(\phi)=\frac{60I_0}{r}.$$

This pattern is circular in horizontal plane.

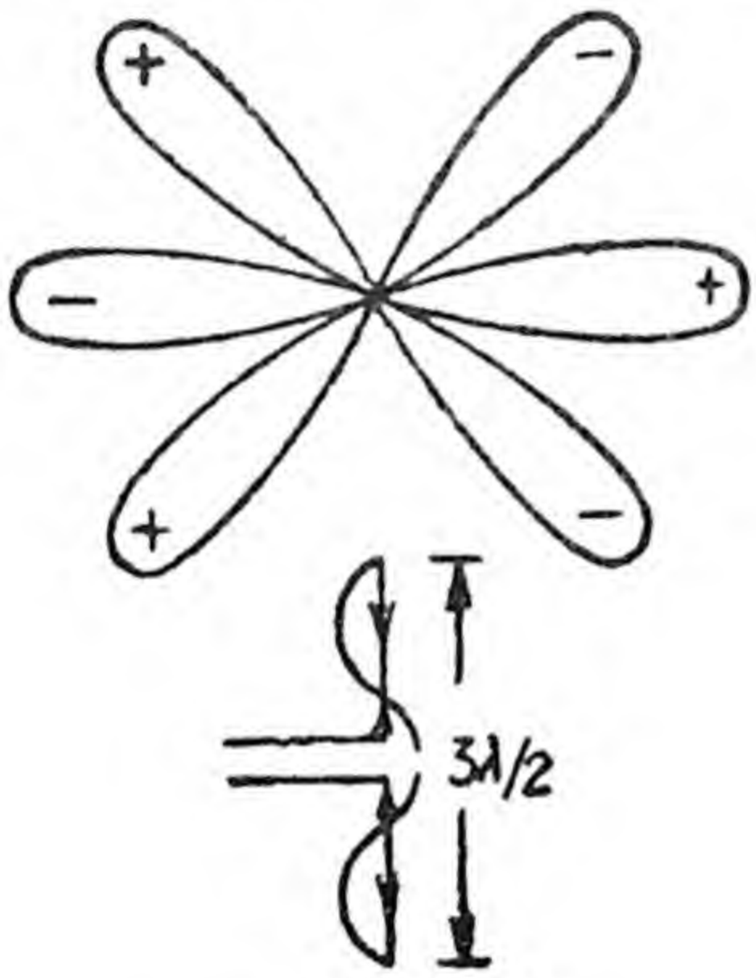


Fig. 10·5.

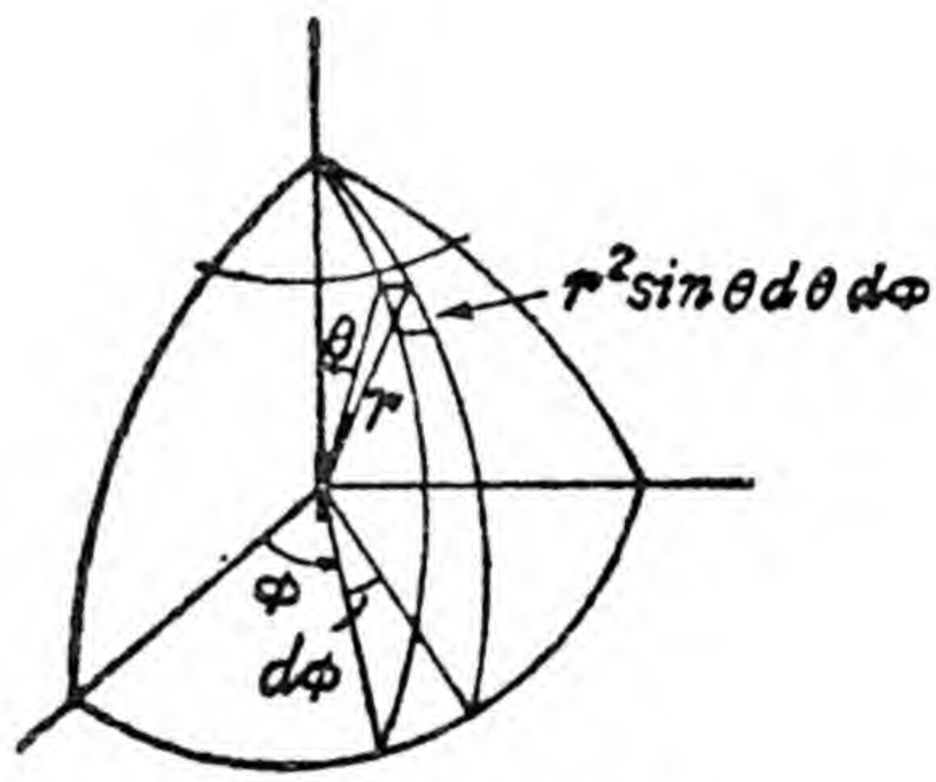


Fig. 10·6.

(*iii*) 3/2 *wavelength antenna* : When $l=3\lambda/2$, the pattern factor is

$$E_\theta=\frac{\cos(3\pi/2.\cos\theta)}{\sin\theta}. \qquad ...(14)$$

The pattern in vertical plane is represented in fig. (10 5). For horizontal plane $\theta=\pi/2$ so that form factor is unity and pattern is circular.

(c) Calculation of power radiated and radiation resistance :

Suppose the antenna is located at the centre of spherical co-ordinate system. Then the average power radiated through a small surface element of area† $r^2 \sin d\theta\, d\phi$ will be

$$dP=\tfrac{1}{2}E_\theta H_\phi r^2 \sin\theta\, d\theta\, d\phi \text{ watts/sec.}^*$$
$$=\tfrac{1}{2}\times 120\pi\, H_\phi H_\phi r^2 \sin\theta\, d\theta\, d\phi$$
$$=60\pi\, H_\phi^2 r^2 \sin\theta\, d\theta\, d\phi.$$

---

†Since $\int_0^{2\pi}\int_0^{\pi} r^2 \sin\theta\, d\theta\, d\phi = 2\int_0^{\pi/2} 2\pi r^2 \sin\theta\, d\theta$,

we can also take a ring area over a spherical surface as used in the case of doublet antenna art. 28·1(*c*).

*Factor $\frac{1}{2}$ occurs due to the average over one cycle *i.e.*, for $\cos^2\omega(t-r/c)$ factor, $\frac{1}{2}$ has been put.

Therefore, total power radiated into space will be

$$P=60\pi \int_0^{2\pi}\int_0^{\pi} H_\phi^2 r^2 \sin\theta \, d\theta \, d\phi,$$

Putting value of H from equation (10), we find

$$P=\frac{15I_0^2}{\pi}\int_0^{2\pi}\int_0^{\pi}\frac{\left[\cos\left(\frac{\beta l\cos\theta}{2}\right)-\cos\frac{\beta l}{2}\right]^2}{\sin\theta}d\theta\,d\phi$$

$$=30I_0^2\int_0^{\pi}\frac{\left[\cos\left(\frac{\beta l\cos\theta}{2}\right)-\cos\frac{\beta l}{2}\right]^2}{\sin\theta}d\theta$$

$$=60I^2{}_{r\ m\cdot s.}\int^{\pi}\frac{\left[\cos\left(\frac{\beta l\cos\theta}{2}\right)-\cos\frac{\beta l}{2}\right]^2}{\sin\theta}d\theta.$$

This is the expression for the power radiated by an antenna of any length l. We can make it particular by specifying the value of l.

Therefore, the corresponding radiation resistance will be

$$R_r=60\int_0^{\pi}\frac{\left[\cos\left(\frac{\beta l\cos\theta}{2}\right)-\cos\frac{\beta l}{2}\right]^2}{\sin\theta}d\theta. \qquad ...(15)$$

This is the general expression for the radiation resistance of any length l of antenna. For a particular case let us take first :

(*a*) *Half wave antenna* :

For half wave antenna $l=\lambda/2$,

$$R_{\lambda/2}=60\int_0^{\pi}\frac{\left[\cos\left(\frac{2\pi.\lambda/2}{\lambda.2}\cos\theta\right)-\cos\frac{2\pi}{2\lambda}\lambda/2\right]^2}{\sin\theta}d\theta$$

$$=60\int_0^{\pi}\frac{\cos^2\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta}d\theta$$

$$=60\times 1{\cdot}22\approx 73 \text{ ohms.}$$

Therefore power radiated is

$$P=RI^2{}_{r\cdot m\ s}=73I^2{}_{r\ m\cdot s}$$

$$=36{\cdot}5I_0^2.$$

(*b*) *Quarter wave antenna* : We know that equation (15) is fit for the calculation of R_r at feed point which is also point of current maximum.

But if we put $l=\lambda/4$ in equation (15), then, the case corresponds to fig. 10·7 (*b*) in which at feed point current is not maximum Therefore we cannot obtain the value of R_r from equation (15) simply by putting $l=\lambda/4$.

However, to apply it to the case of quarter wave antenna, we can think a quarter wave antenna to be equivalent to upper or lower half of a half wave antenna [fig. 10·7(*a*)] so that $l/2=\lambda/4$ and

current is maximum at the feed point. The power radiated will be half of the power radiated by half wave antenna. That is,

$$R_r = 30 \int_0^{\pi} \frac{\left[\cos\left(\frac{\beta l \cos\theta}{2}\right) - \cos\frac{\beta l}{2}\right]}{\sin\theta} d\beta.$$

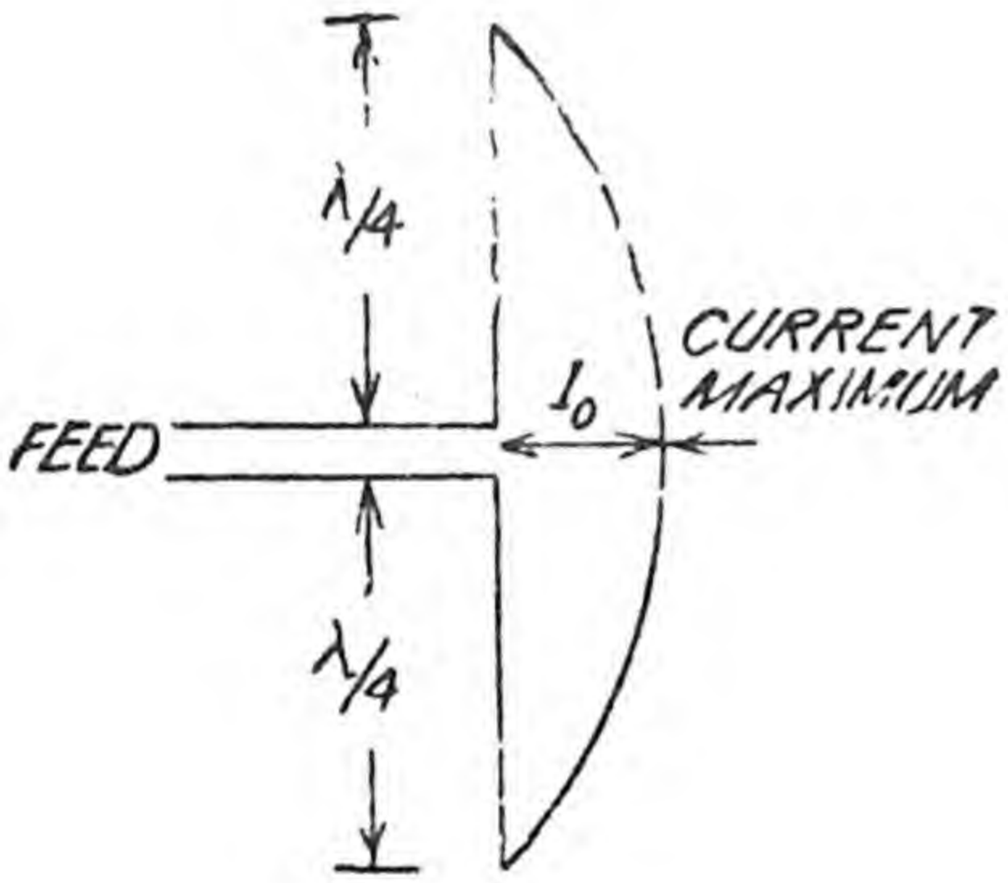


Fig. 10·7.

Let us put $l/2=\lambda/4$, so that

$$R_r = 30 \int_0^{\pi} \frac{\cos^2\left(\frac{1}{2}\pi\cos\theta\right)}{\sin\theta} d\theta,$$

$$R_r = 30 \times 1{\cdot}22$$

$$= 36{\cdot}6 \text{ ohms.}$$

28·6. CURRENT DISTRIBUTION ON RESONANT AND NON-RESONANT ANTENNA :

When both incident and reflected waves exist on an antenna, it is said to be resonant one Therefore, the current on the antenna will be expressed as a sum of incident and reflected components.

For example, $I = I_0 e^{+j\beta z} + I_0 e^{-j\beta z}$

$$= 2I_0 \cos\beta z$$

or $I = I_0 e^{+j\beta z} - I_0 e^{-j\beta z}$

$$= jI_0 \sin\beta z.$$

where $\beta = 2\pi/\lambda = \omega/c.$

This means current distribution for resonant type of antenna can be expressed by sine or cosine functions. In art. 28·5, we have discussed a resonant antenna.

If the antenna is top loaded such that incident wave is completely absorbed no reflected wave exists on the antenna then it is said to be nonresonant. Obviously, current on such an antenna can be expressed by

$$I = I_0 e^{-j\beta z}$$

only. This is also called antenna with progressive wave.

NON-RESONANT ANTENNA :

We shall calculate the power r adiated and radiation resistance of non-resonant antenna of any length. Let us consider a non-resonant antenna of length $2l$ shown in fig. 11. The current distribution may be represented by

$$|I| = I_0 e^{-j\beta z},$$

so that $$I = |I| e^{j\omega(t-r/c)}$$
$$= I_0 e^{-j\beta z} e^{-j\omega(t-r/c)}.$$

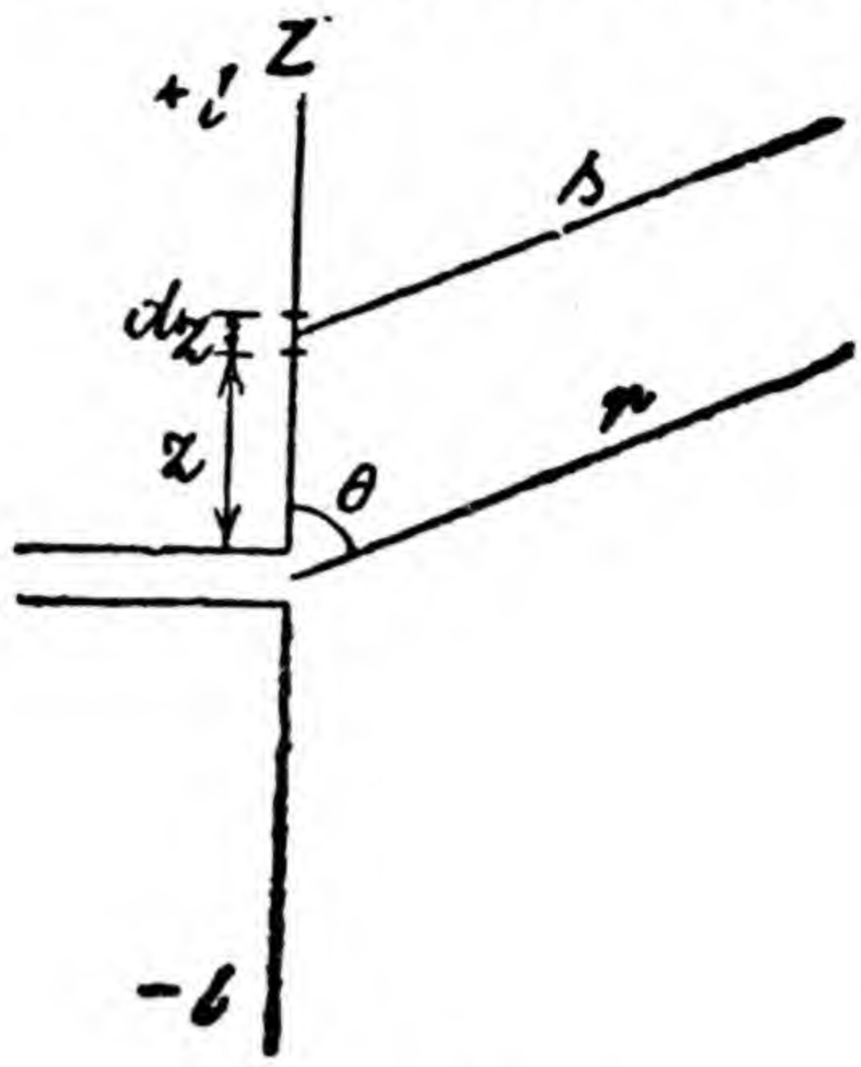


Fig. 11.

The radiative magnetic field at any point P distant s meters from an element of length dz is given by

$$dH_\phi = \frac{I\,dz \sin\theta}{2\lambda s}$$
$$= \frac{I_0 e^{-j\beta z}\,dz \sin\theta}{2\lambda s} e^{j\omega(t-s/c)}$$

amp. turn/meter.

In amplitude term, we can replace s by r, but in phase term, we shall have to put $s = r - z\cos\theta$,

$$dH_\phi = \frac{I_0 e^{-j\beta z}\,dz \sin\theta}{2\lambda r} e^{j\omega\{t-(r-z\cos\theta)/c\}}.$$

Therefore the magnetic field at point P due to the whole antenna will be

$$H_\phi = \int_{-l}^{+l} \frac{I_0 e^{-j\beta z}\,dz \sin\theta}{2\lambda r} e^{j\omega\{t-(r-z\cos\theta)/c\}}$$
$$= \frac{I_0 e^{j\omega(t-r/c)} \sin\theta}{2\lambda r} \int_{-l}^{+l} e^{-j\beta z} \cdot e^{j(\omega/c)\, z\cos\theta}.\, dz$$
$$= \frac{I_0 e^{j\omega(t-r/c)} \sin\theta}{2\lambda r} \int_{-l}^{+l} e^{-j\beta z(1-\cos\theta)}\, dz$$
$$= \frac{I_0 e^{j\omega(t-r/c)} \sin\theta}{2\lambda r} \cdot \frac{2}{\beta} \cdot \frac{\sin\{\beta l(1-\cos\theta)\}}{(1-\cos\theta)}$$
$$= \frac{I_0 e^{j\omega(t-r/c)} \sin\theta}{2\pi r} \cdot \frac{\sin\{\beta l(1-\cos\theta)\}}{(1-\cos\theta)} \quad \ldots(1)$$

Therefore, corresponding electrostatic field is

$$E_\theta = 120\pi . H_\theta = \eta H_\phi$$
$$= \eta . \frac{I_0 e^{j\omega(t-r/c)}}{2\pi r} \sin\theta \frac{\sin\{\beta l(1-\cos\theta)\}}{(1-\cos\theta)} \quad \ldots(2)$$

Power radiated in watts/unit area/sec. averaged over one cycle will be $P_{av} = \frac{1}{2} E_\theta H_\phi$

$$= \frac{1}{2} \frac{\eta I_0^2}{4\pi^2 r^2} \left[\frac{\sin\theta \sin\{\beta l(1-\cos\theta)\}}{(1-\cos\theta)}\right]^2,$$

in which factor $\frac{1}{2}$ appears because of the average of $e^{2j\omega(t-r/c)}$ taken over one cycle.

Therefore, power through an element of area $r^2 \sin\theta\, d\theta\, d\phi$ will be

$$P_{av}' = \frac{\eta I_0^2}{\varkappa \pi^2 r^2}\left[\frac{\sin\theta \sin\{\beta l\,(1-\cos\theta)\}}{(1-\cos\theta)}\right]^2 r^2 \sin\theta\, d\theta\, d\phi.$$

Then the total power radiated into space by the resonant antenna will be

$$P_{av} = \int_0^{2\pi}\int_0^{\pi} \frac{\eta I_0^2}{8\pi^2 r^2}\left[\frac{\sin\theta \sin\{\beta l\,(1-\cos\beta)\}}{(1-\cos\theta)}\right]^2 r^2 \sin\theta\, d\theta\, d\phi$$

$$= \frac{\eta I_0^2}{4\pi}\int_0^{\pi} \frac{\sin^3\theta \sin^2\{\beta l\,(1-\cos\theta)\}}{(1-\cos\theta)^2} d\theta.$$

Let us put

$$1-\cos\theta = u,$$

so that $\sin^2\theta = u(2-u)$ and $\sin\theta\, d\theta = du$.

Then

$$P_{av} = \frac{\eta I_0^2}{4\pi}\int_0^2 \frac{(2-u)}{u}\sin^2\beta l u\, du$$

$$= \frac{\eta I_0^2}{4\pi}\int_0^2 \left(\frac{2}{u}-1\right)\frac{(1-\cos 2\beta l u)}{2}\, du$$

$$= \frac{\eta I_0^2}{4\pi}\left[-\frac{1}{2}\int_0^2 (1-\cos 2\beta l u)\, du + \int_0^2 \frac{(1-\cos 2\beta l u)}{u}\, du\right]$$

$$= \frac{\eta I_0^2}{4\pi}\left[-\tfrac{1}{2}\left\{u - \frac{\sin 2\beta l u}{2\beta l}\right\}_0^2 + \int_0^{4\beta l}\frac{1-\cos v}{v}\, dv\right]$$

$$= \frac{\eta I_0^2}{4\pi}\left[-1 + \frac{\sin 4\beta l}{4\beta l} + \int_0^{4\beta l}\frac{1-\cos v}{v}\, dv\right],$$

in which $v = 2\beta l u$.

Special Case : For half wave antenna $2l = \lambda/2$. Power radiated is given by

$$P_{\lambda/2} = \frac{\eta I_0^2}{4\pi}\left[-1 + \frac{\sin 2\beta.\lambda/2}{2\beta\lambda/2} + \int_0^{4\beta\,\lambda/4}\left(\frac{1-\cos v}{v}\right)dv\right]^*$$

$$= \frac{\eta I_0^2}{4\pi}[-1+0+2{\cdot}4377]$$

$$= \frac{\eta I_0^2}{4\pi}\times 1{\cdot}4377$$

$$= \frac{\eta I_0^2}{2}\cdot\frac{0{\cdot}719}{\pi} = \frac{\eta}{\pi}.I^2_{r\cdot m\cdot s}.\ 0{\cdot}719$$

$$^*\int_0^{2\pi}\left(\frac{1-\cos v}{v}\right)dv = \left[\frac{v^2}{2.2!} - \frac{v^4}{4.4!} + \frac{v^6}{6.6!} - \frac{v^8}{8.8!} + \cdots\right]_0^{2\pi}$$

$$= 2{\cdot}4377$$

Therefore, radiation resistance of half wave antenna is

$$R_{\lambda/2} = \frac{\eta}{\pi} .0{\cdot}^{7}19$$

$$= \frac{120\pi \times 0{\cdot}^{7}19}{\pi}$$

$$= 86{\cdot}3 \text{ ohms.}$$

and is greater than that for half wave resonant antenna (73 ohms).

The ratio of the power radiated by the actual antenna, under consideration, to the power radiated by an isotropic radiator, called directive gain, G_p, is shown in fig. 12. The plot shows that G_p is

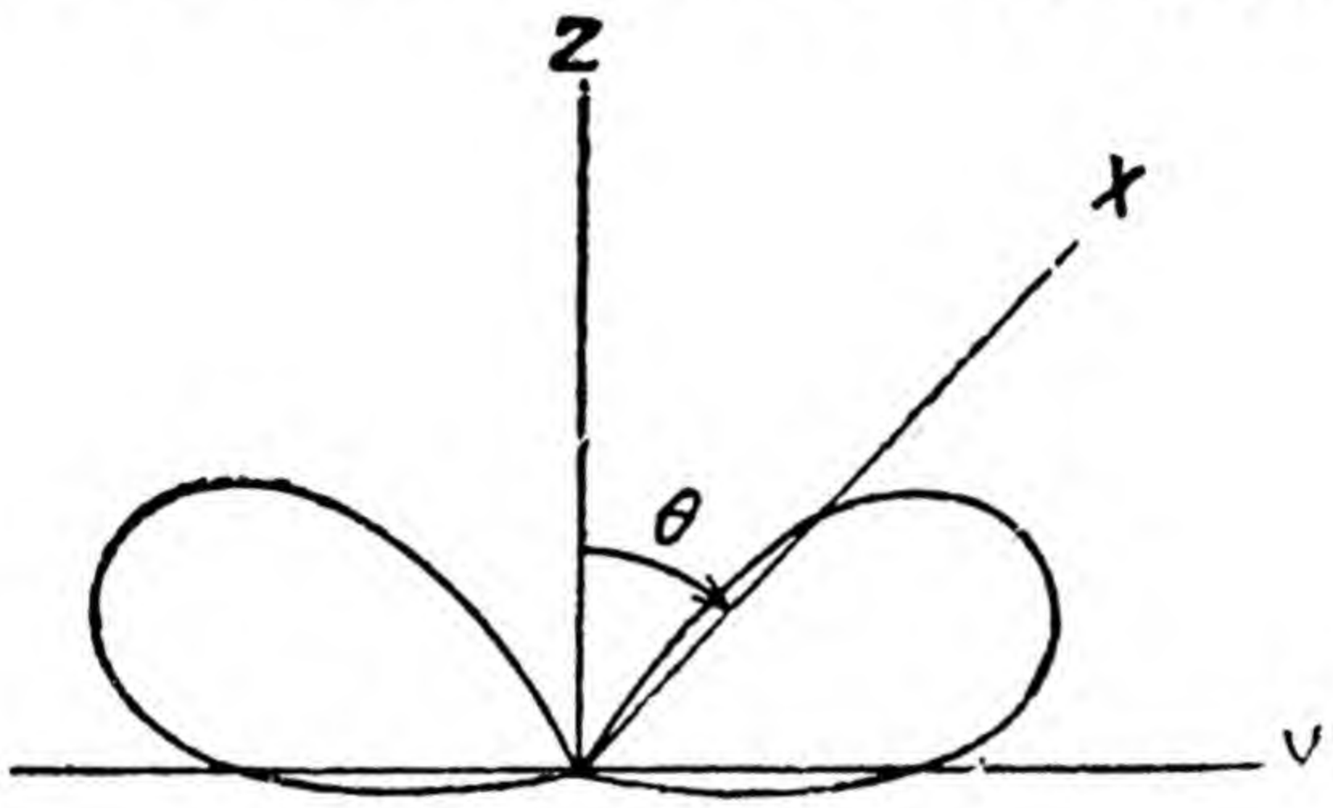


Fig. 12.

not symm trical about $\theta = \pi/2$ (note that it is symmetrical in case of resonant $\lambda/2$ antenna). We can, therefore, conclude that a $\lambda/2$ antenna with progressive current distribution tends to concentrate the power in the directions close to the direction of progression of current along the antenna.

28·7-1. EFFECT OF GROUND : IMAGE ANTENNAS :

In all previous discussions it has been assumed that antenna was remote from ground. When the antenna is grounded or ungrounded but close to earth, its directional characteristics are altered, since energy radiated towards the earth is reflected and thus the total field in any direction will be vector sum of direct and reflected waves. For the purpose of analysis. it is convenient to regard the reflected radiation as coming from an 'image' antenna located below the surface of earth as shown in fig. 13.

If we give a physical configuration to this fictitious antenna length equivalent to the mirror image of actual antenna. then the radiation characteristics of the equivalent system resemble with those of actual system The approximation involved in such a process is the assumption that the earth is a perfect reflector and absorbs no energy. The direction of current flow is such that the vertical component of the current in the image is in the same direction as the vertical component of the current in the corresponding parts of the actual antenna. while the horizontal component of

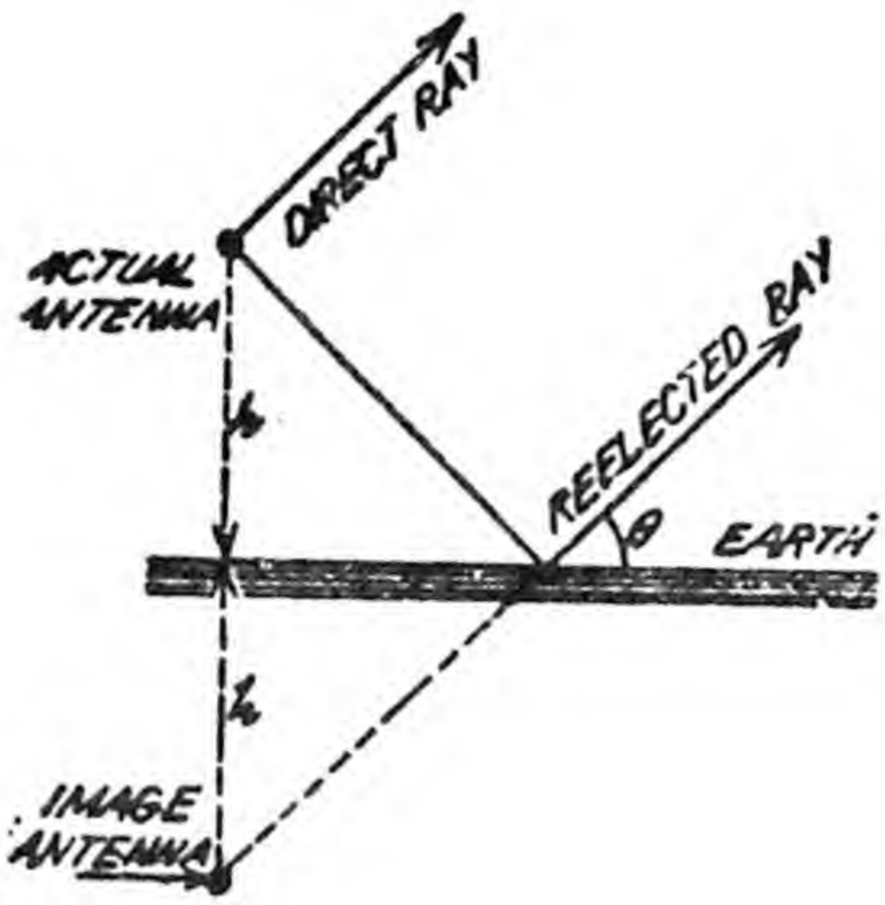


Fig. 13. Figure illustrating concept of image antenna.



Fig. 14 Relation of actual antenna and image antenna currents.

28·7-2. FIELD AT A POINT ABOVE THE SURFACE OF EARTH DUE TO HORIZONTAL AND VERTICAL ANTENNAS CLOSE TO EARTH :

(i) Ground affected horizontal antenna : Let AA' represent a horizontal antenna of length l and located at a height h above the ground. An image antenna BB' (fig. 15) should definitely be accounted at a depth equal to height h and having the same length as that of actual antenna. It should be pointed out here that in this case, the current in actual and in image antennas will be 180° out of phase (shown by waveform representation in the figure).

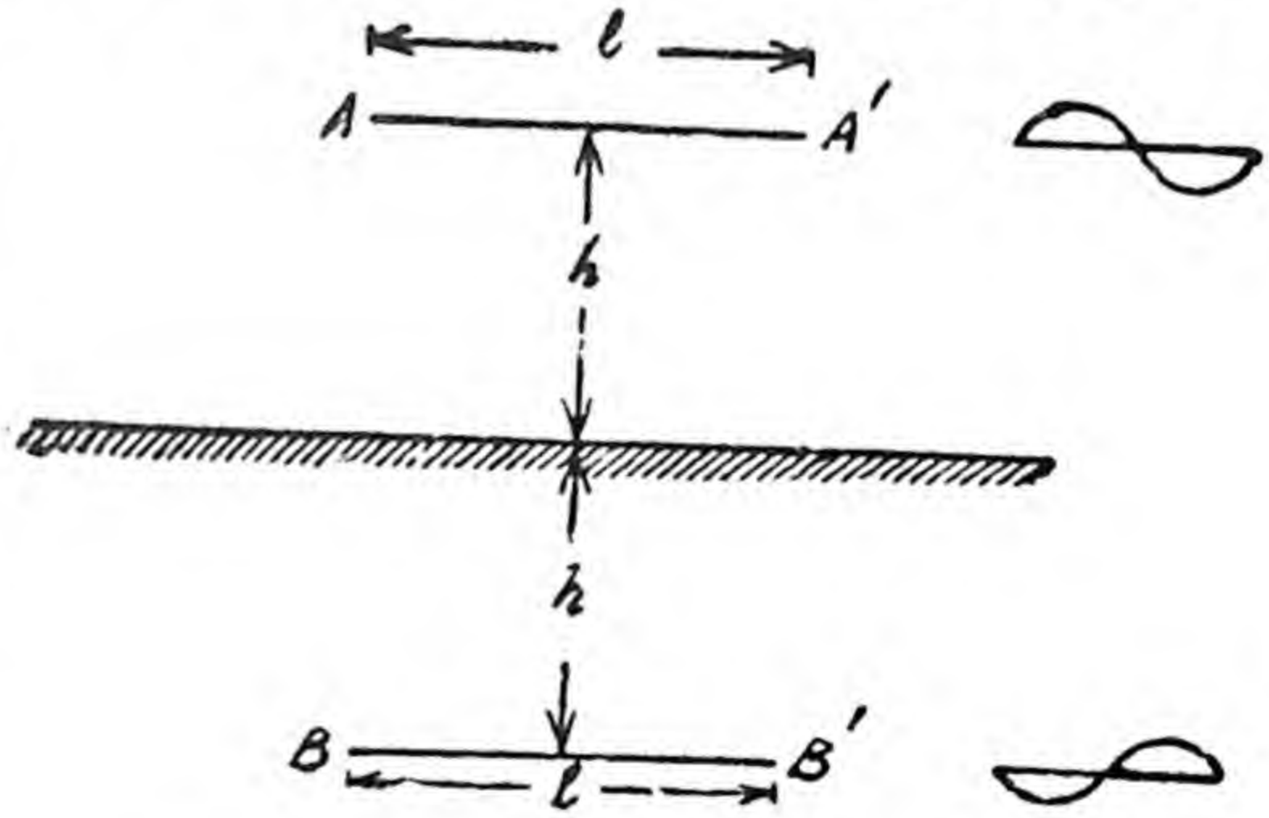


Fig. 15. Horizontal antenna close to ground.

For a point above the earth, radiations from actual antenna will reach sooner as compared to image antenna or simply radiations from BB' lag to those from AA'.

Let us calculate the field at a point P due to antenna AA' and its image antenna BB'. Let ϕ be the angle which the point P makes with a line perpendicular to antennas and θ that with horizontal. Obviously $\phi = 90° - \theta$ (fig. 16).

The path difference between the radiations from AA' and BB' will be $2h \cos \phi$ or phase difference equal to $\frac{2\pi}{\lambda} . 2h \cos \phi$. Furthermore, let the phase difference between radiations from two antennas due to the currents flowing in them be δ (the

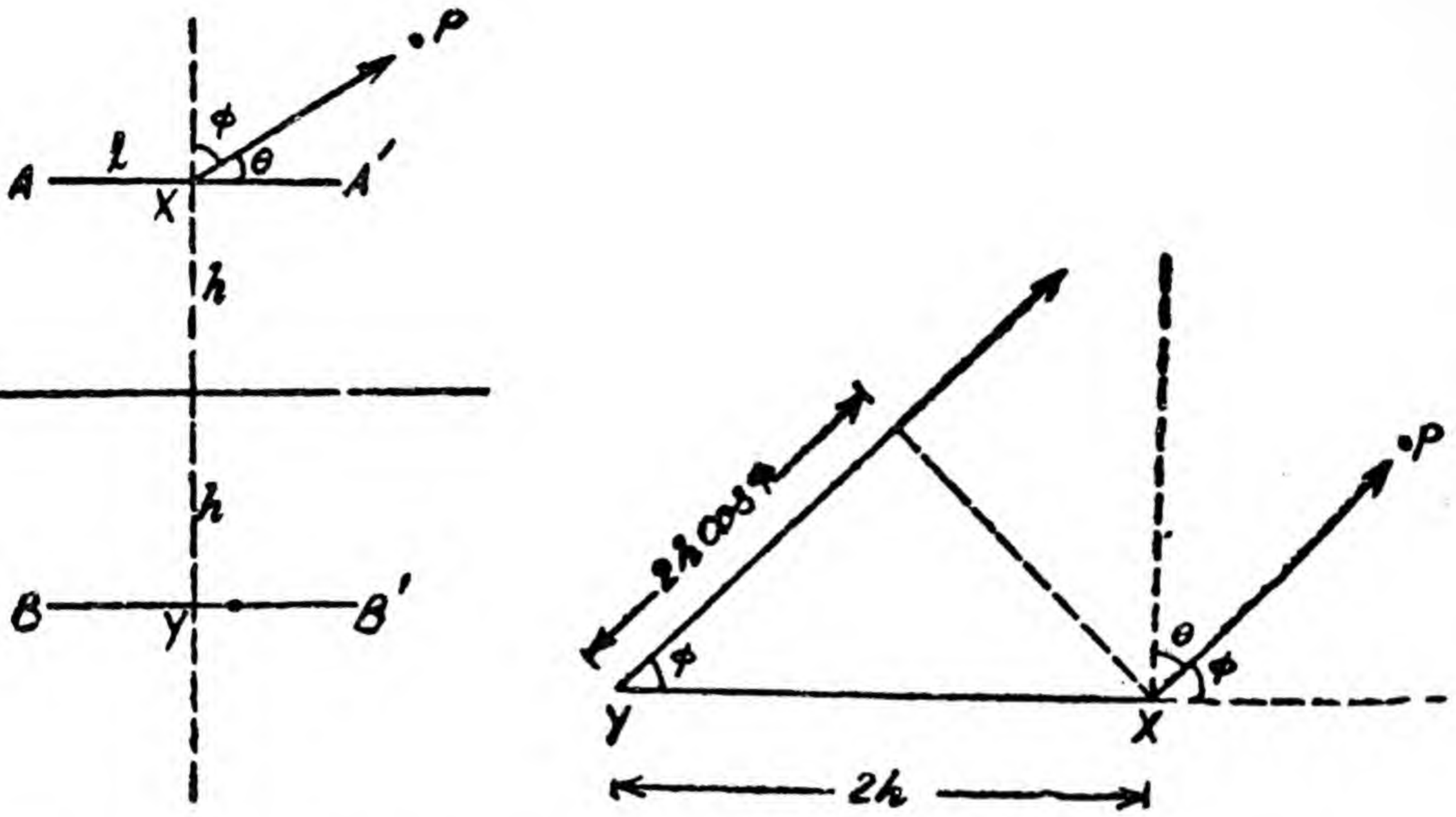


Fig. 16. Field calculations for ground affected horizontal antenna.

this case). Thus total phase difference is,

$$\alpha=\beta+\frac{2\pi}{\lambda}.\ 2h\cos\phi.$$

Also, if E_A represents the field at P due to antenna AA' and E_B that due to antenna BB', the phase difference between the two being α, the total field at P is

$$=\sqrt{\left\{ E_A^2+E_B^2 \right.}$$

$$\left.+2E_AE_B\cos\left(\beta+\frac{2\pi}{\lambda}.2h\cos\phi\right)\right\}.$$

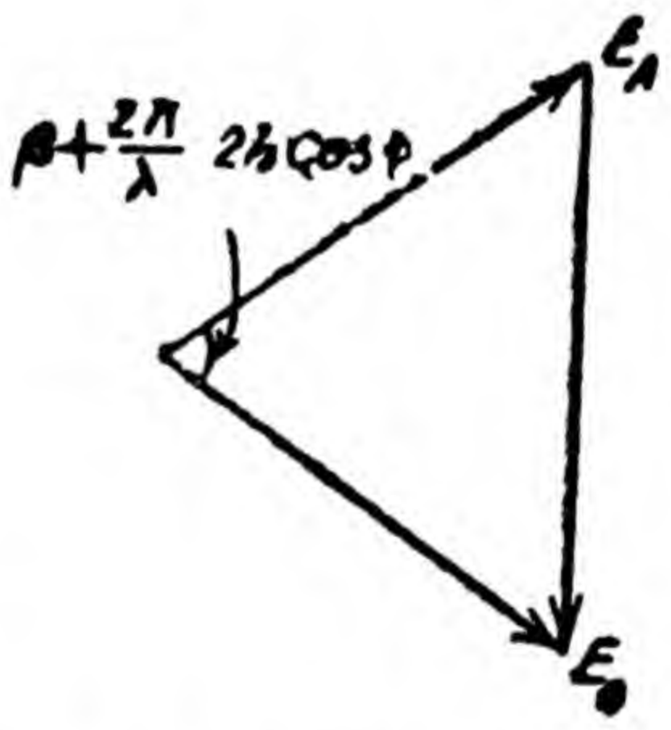


Fig. 17. (*a*) Resultant field due to E_A & E_B.

If $E_A=E_B$, resultant field at P will be

$$=E_A\sqrt{\left\{2+2\cos\left(\beta+\frac{4\pi h}{\lambda}\cos\phi\right)\right\}}$$

$$=\sqrt{2}.E_A\sqrt{\left\{1+\cos 2\left(\frac{\beta}{2}+\frac{2\pi}{\lambda}.\ h\cos\phi\right)\right\}}$$

$$=\sqrt{2}.E_A\sqrt{\left\{2\cos^2\left(\frac{\beta}{2}+\frac{2\pi}{\lambda}.\ h\cos\phi\right)\right\}}$$

$$=2E_A\cos\left(\frac{\beta}{2}+\frac{2\pi}{\lambda}.\ h\cos\phi\right).$$

Since $\beta=180°$, total field at P is

$$=2E_A\cos\left(90+\frac{2\pi}{\lambda}.\ h\cos\phi\right)$$

$$=-2E_A\sin\left(\frac{2\pi h}{\lambda}\cos\phi\right)$$

$$= -2E_A \sin \left[\frac{2\pi h}{\lambda} \cos (90 - \theta) \right]$$

$$= -2E_A \sin \left(\frac{2\pi h}{\lambda} \sin \theta \right),$$

which shows that the field is proportional to $2 \sin \left(\frac{2\pi h}{\lambda} \sin \theta \right)$.

Let us consider a half wave antenna, located half wavelength above the ground, then $l = \frac{1}{2}\lambda$,

$h = \frac{1}{2}\lambda$,

and take $\theta = 30°$.

Putting in the equation for field, we get the total field value equal to its maximum value $2E_A$. With different values of θ, we can easily calculate the corresponding magnitude of field strength. The data obtained are plotted in fig. 17 (*b*) to obtain the polar diagram of this antenna.

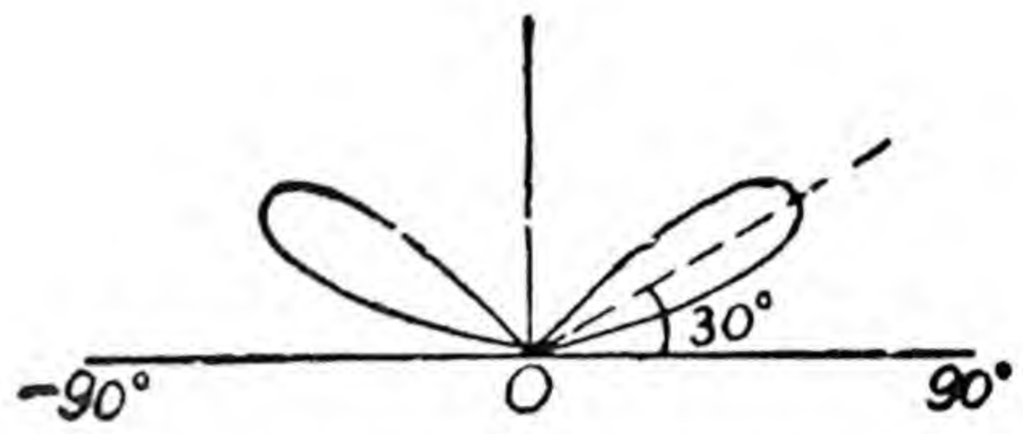


Fig. 17. (*b*) Polar diagram for half wave antenna.

(ii) Ground affected vertical antenna : In fig. 18, AA' represents the vertical aerial and its image is shown by BB'. The currents in the two antennas are in the same phase. We shall calculate the field at point P due to this vertical antenna.

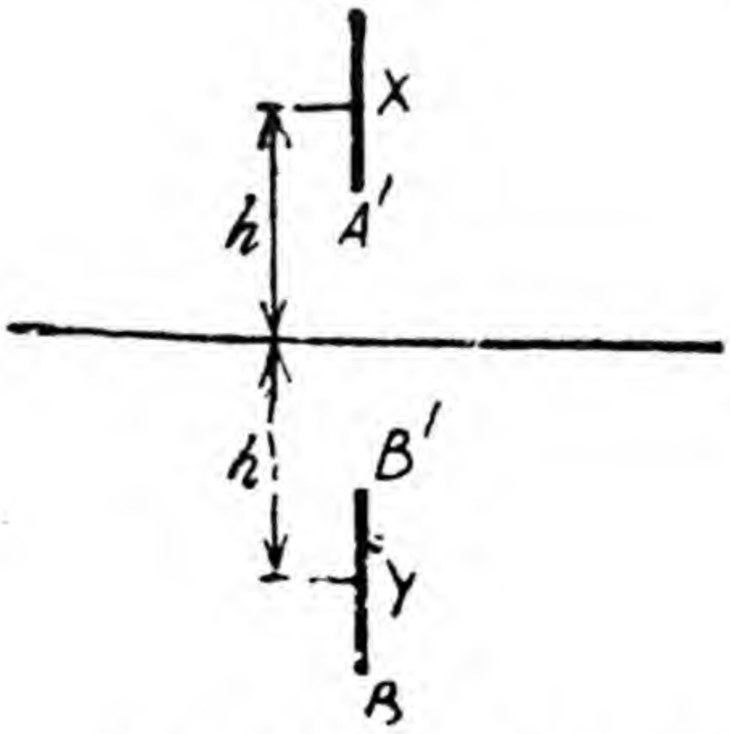
Total phase difference in this case

$$\alpha = \frac{2\pi}{\lambda} . \, 2h \cos \phi, \text{ since } \beta = 0,$$

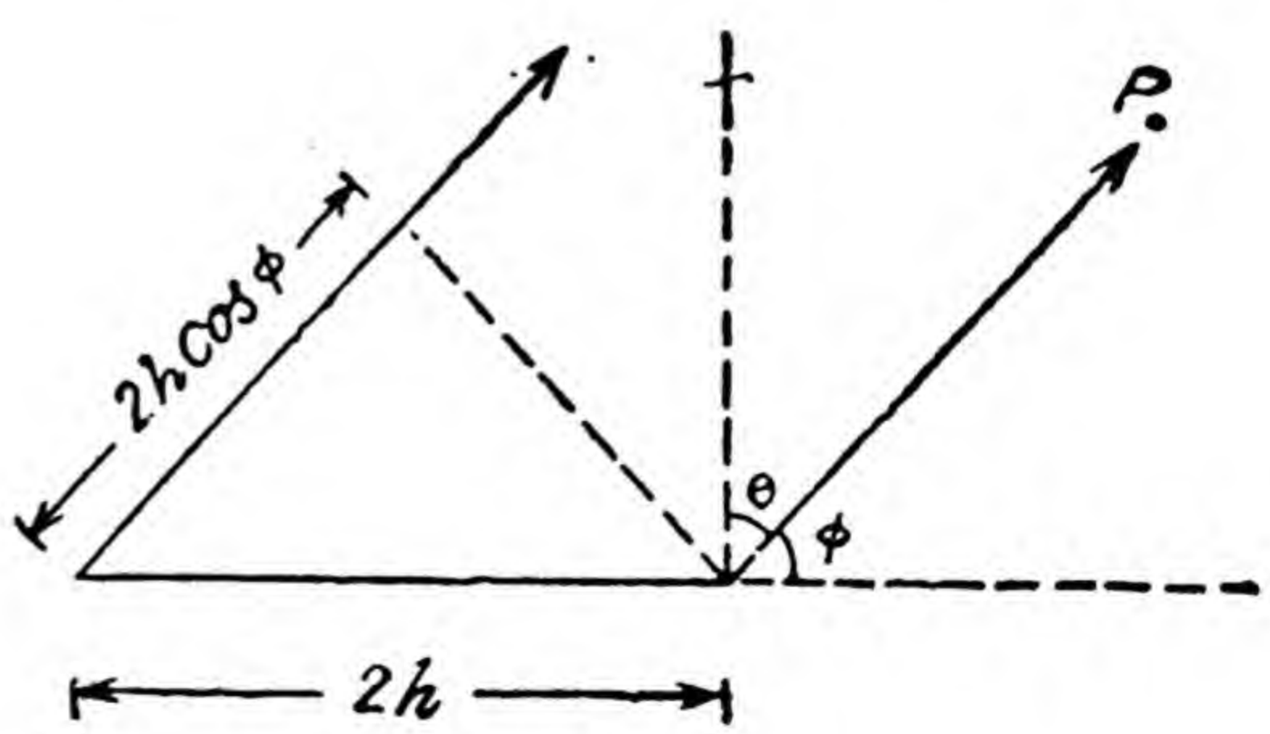
$\therefore$ Total field at P

$$= \sqrt{\left\{ E_A^2 + E_B^2 + 2E_A E_B \cos \left(\frac{2\pi}{\lambda} . \, 2h \cos \phi \right) \right\}}$$

$$= \sqrt{2}E_A \sqrt{\left\{ 1 + \cos \left(\frac{4\pi h}{\lambda} \cos \phi \right) \right\}} \qquad \text{if } E_A = E_B.$$



(*a*) Antenna representation.



(*b*) Field calculations.

Fig. 18. Ground affected vertical antenna representation and field calculation.

$$=2E_A \cos\left(\frac{2\pi h}{\lambda}\cos\phi\right)$$

$$=2E_A \cos\left(\frac{2\pi h}{\lambda}\sin\theta\right).$$

With different values of λ and θ, polar diagrams of the field radiated in space can be sketched. In next article, we shall discuss it in detail.

28·8. SHORT VERTICAL GROUNDED ANTENNA :

(a) Field Distribution : The field distribution around a vertical antenna may be obtained by determining the strength of the field at any point P due to the current in an elementary length dx and integrating over the length of the antenna. We assume the current distribution to be sinusodial and the antenna to be operated at a wavelength below its natural wavelength as shown in figure 19. Current can be expressed as :

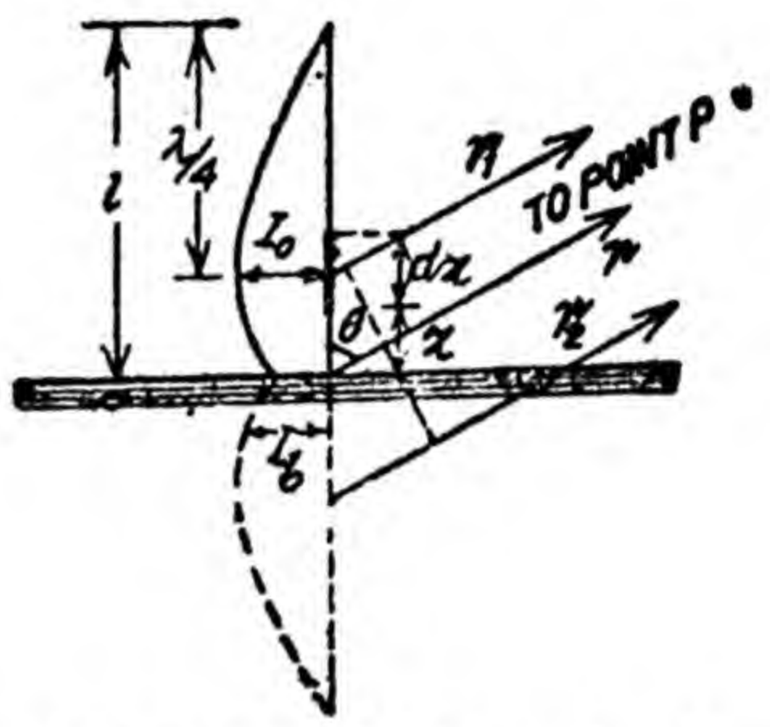


Fig. 19. Grounded vertical antenna and its current distribution.

$$I=|I|\, e^{j\omega(t-r/c)}$$

where the amplitude of the current $|I|$ at height x above the earth will be

$$|I|=I_0 \sin\frac{2\pi(l-z)}{\lambda}. \qquad ...(1a)$$

For $z=0$, $|I|=I_b$,

so that

$$I_b=I_0 \sin\frac{2\pi l}{\lambda}$$

or

$$I_0=I_b \Big/ \sin\frac{2\pi l}{\lambda}.$$

$$\text{Therefore } I=I_b \sin\frac{2\pi(l-z)}{\lambda}\Big/\sin\frac{2\pi l}{\lambda}, \qquad ...(1b)$$

where I_0 is maximum value of current or amplitude at an antinode and I_b is maximum value of current or amplitude at the base of the antenna. Equations (1a) and (1b) both are equally well for the purpose of analysis ; the latter merely expresses current at any length interms of base current of the antenna.

With some substitutions for current and distance values in the radiative magnetic field equation for short electric doublet, we can write that the strength of the radiative magnetic field, in m.k.s. system, due to this so chosen element of length dx will be

$$dH=\frac{I_0\, dz \sin\theta}{2\lambda r_1}\sin\frac{2\pi(l-z)}{\lambda}\, e^{j\omega(t-r_1/c)}. \qquad ...(2)$$

We assume that for distant point P, r, r_1 and r_2 are parallel lines. In equation (2) $1/r_1$ affects only the amplitude and hence at a large distance it may be regarded as constant. Also at a large distance the difference between r_1 and r can be neglected in its effect on the amplitude although its effect on the phase must be considered. Further from figure 19,

$$r_1 = r - z\cos\theta.$$

Therefore

$$dH = \frac{I_0\, dz \sin\theta}{2\lambda r} \sin\frac{2\pi(l-z)}{\lambda} \text{ exp. } j\omega\left(t - \frac{r - z\cos\theta}{c}\right) \quad \ldots(3)$$

Assuming the earth to be perfectly conducting, the reflection from the ground may be taken into account by adding to equation (3) the radiation of a similar element located below the earth surface at a distance z. The total field at P due to both these elements will be

$$dH = \frac{I_0\, dz \sin\theta}{2\lambda r}.\sin\frac{2\pi(l-z)}{\lambda}\left[\text{exp. } j\omega\left(t - \frac{r+z\cos\theta}{c}\right) + \text{exp. } j\omega\left(t - \frac{r-z\cos\theta}{c}\right)\right]$$

$$= \frac{I_0\, dz \sin\theta}{\lambda r}.\sin\frac{2\pi(l-z)}{\lambda}.\, e^{j\omega(t-r/c)}\cos\left(\frac{\omega z\cos\theta}{c}\right)$$

$$= \frac{I_0\, dz \sin\theta}{\lambda . r}\; e^{j\omega(t-r/c)}.\sin\frac{2\pi(l-z)}{\lambda}.\cos\frac{2\pi z\cos\theta}{\lambda} \quad \ldots(4)$$

Therefore, the field due to whole grounded antenna will be given by

$$H = \frac{I_0\sin\theta}{\lambda r}\, e^{j\omega(t-r/c)} \int_0^l \sin\beta(l-z).\cos(\beta z\cos\theta).dz$$

$$= \frac{I_0\sin\theta}{\lambda r}\, e^{j\omega(t-r/c)}.\; \frac{[\cos(\beta l\cos\theta) - \cos\beta l]}{\frac{2\pi}{\lambda}\sin^2\theta},$$

$$H = \frac{I_0}{2\pi r\sin\theta}[\cos(\beta l\cos\theta) - \cos\beta l]\, e^{j\omega(t-r/c)} \quad \ldots(5)$$

Therefore, the strength of electrostatic field in volt/meter will be given by

or $$E = \frac{60 I_0}{r\sin\theta}[\cos(\beta l\cos\theta) - \cos\beta l]\, e^{j\omega(t-r/c)}$$

or $$|E| = \frac{60 I_0}{r\sin\theta}[\cos(\beta l\cos\theta) - \cos\beta l] \quad \ldots(6)$$

From the above relation the value of electric field in different directions for different antenna lengths can be found. For example, the amplitude $|E|$ at a point on earth's surface for quarter wave

antenna will be

$$| E_{90} |_{\lambda/4} = \frac{60 I_0}{r} ; \qquad \text{also } | E_{90} |_{\lambda/2} = \frac{120 I_0}{r} . \qquad ...(7)$$

(b) Field patterns ; distribution of radiated energy in vertical plane : The distribution of radiated energy in the vertical plane will be proportional to the square of $| E |$ and is shown in fig. 20 for various ratios of the antenna length to the operating wavelength.

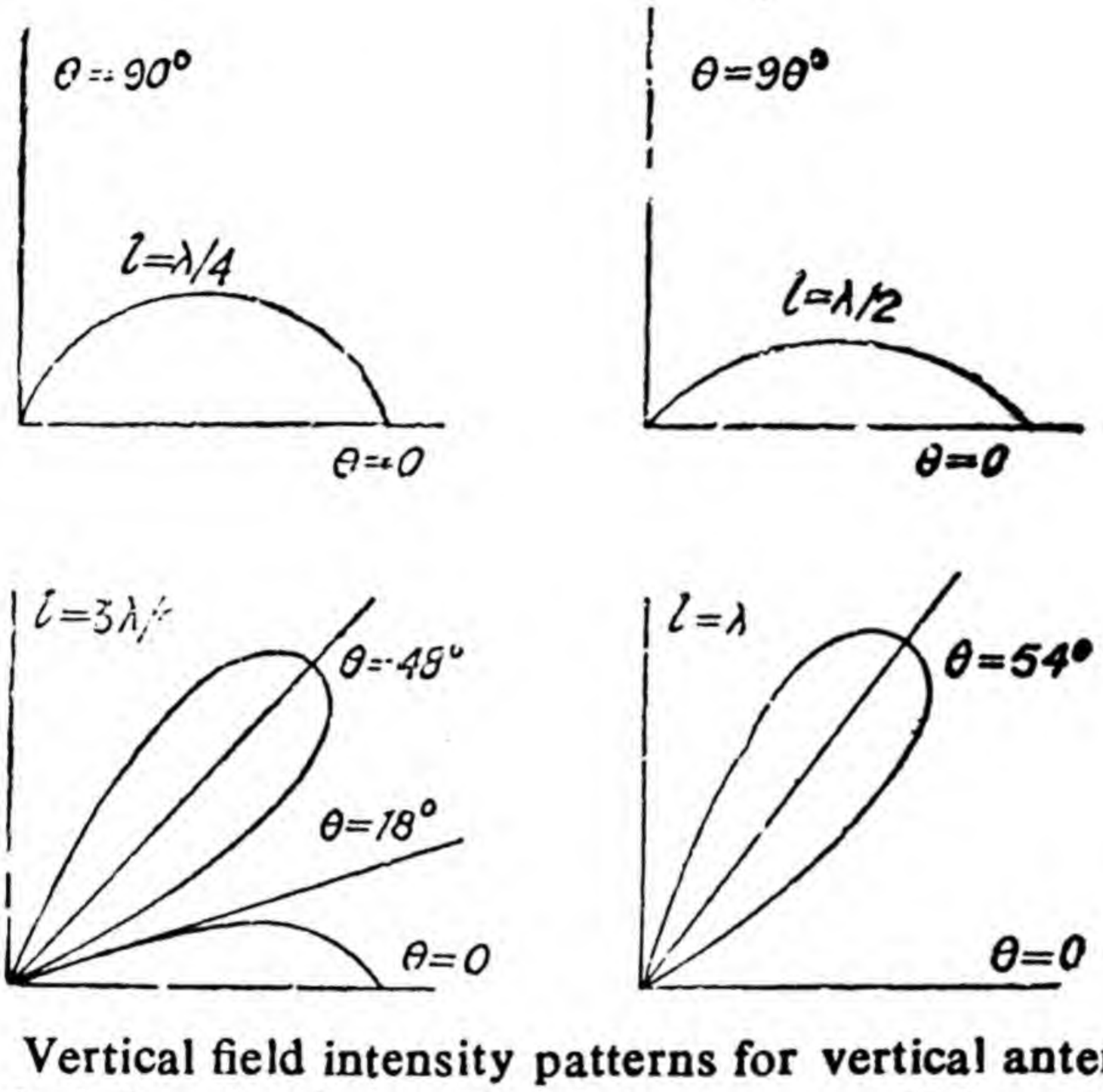


Fig. 20. Vertical field intensity patterns for vertical antenna.

When the height of the antenna is increased beyond quarter wavelength, the pattern tends to be flatten and a greater field intensity occurs along the ground.

If angle θ is measured from the ground then equation (6) for electric field will be

$$| E | = \frac{60 I_0}{r} \frac{[\cos (\beta l \sin \theta) - \cos \beta l]}{\cos \theta} . \qquad ...(8)$$

At various angles, the distribution of energy is represented in fig. 20. It is obvious that at $\frac{l}{\lambda} = 0{\cdot}75$, a secondary lobe appears at high angles and thus the energy splits in two portions. At $\frac{l}{\lambda} = 1{\cdot}0$, the primary lobe disappears completely and the maximum energy is now directed along a line at 54° above the horizontal. It can be seen that for $\frac{l}{\lambda} > 1$, the two lobes again appear.

(c) Calculation of Radiated Power and Radiation Resistance : As has been pointed out earlier, once the radiation fields are calculated, expressions for power radiated and radiation resistance

are atonce at hand. Quite obviously, grounded aerial will radiate through a *hemisphere*. By Poyinting's theorem, the power crossing unit area in M.K.S., system of units, is given by

$$= E.\ H_t \text{ watts/sec.}$$

Since $E = 120\pi H$,

power crossing per unit area is

$$= 120\pi H^2.$$

Therefore power crossing surface of ring assumed on the hemisphere will be

$$120\pi H^2 . 2\pi r^2 \sin\theta\, d\theta.$$

Thus the total power radiated through the hemisphere is

$$P = 240\pi^2 \int_0^{\pi/2} H^2 r^2 \sin\theta\, d\theta,$$

which is the actual power radiated by the antenna.

Now substituting the value of H from equation (5), the general expression for the power radiated bv an antenna of any length l is

$$P = 240\pi^2 \int_0^{\pi/2} \frac{I_0^2}{4\pi^2 r^2 \sin^2\theta} [\cos(\beta l \cos\theta) - \cos\beta l]^2 \times \cos^2\omega(t - r/c).r^2 \sin\theta\, d\theta$$

$$= 60 . \tfrac{1}{2} . \int_0^{\pi/2} I_0^2 \frac{[\cos(\beta l \cos\theta) - \cos\beta l]^2}{\sin\theta} d\theta,$$

where for $\cos^2\omega(t - r/c)$, the factor $\frac{1}{2}$ has been substituted for the reasons explained earlier. Therefore

$$P = 30\ I_0^2 \int_0^{\pi/2} \frac{[\cos(\beta l \cos\theta) - \cos\beta l]^2}{\sin\theta} d\theta,$$

$$P = 60\ I^2_{r.m.s.} \int_0^{\pi/2} \frac{[\cos(\beta l \cos\theta) - \cos\beta l]^2}{\sin\vartheta} . d\theta \text{ watts.} \quad ..(9)$$

The expression for radiation resistance thus becomes

$$R = \frac{P}{I^2_{r.m.s.}}$$

$$= 60 \int_0^{\pi/2} \frac{[\cos(\beta l \cos\theta) - \cos\beta l]^2}{\sin\theta} d\theta \text{ ohms.} \quad ...(10)$$

Equations (9) and (10) can easily be applied to the particular case. For quarter wave antenna $l = \lambda/4$, so that

$$R_{\lambda/4} = 60 \int_0^{\pi/2} \frac{\cos^2\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} d\theta$$

$$= 60 \times 0{\cdot}6095$$

$$= 36{\cdot}57 \text{ ohms.} \quad ...(11)$$

MARCONI AERIAL :

In order to provide a convenient erection, which is not possible

with dipole $\lambda/2$ aerial (Hertz aerial) if the wavelength were several hundred metres. Marconi introduced the $\lambda/4$ vertical aerial of which one end was earthed. Since earth is assumed to be perfect conductor, the lower end of Marconi aerial is thus connected to a large conductor. This aerial is, therefore, fed from the transmitter so as to provide a current anti-node and potential node at the base; and with current node and potential antinode at the open top as shown in fig. 21. If we compare it with half wave Hertz aerial, then the approximate resonant wavelength will be equal to $4l$, where l is the actual antenna length. The aerial is also called as '$\lambda/4$ aerial.'

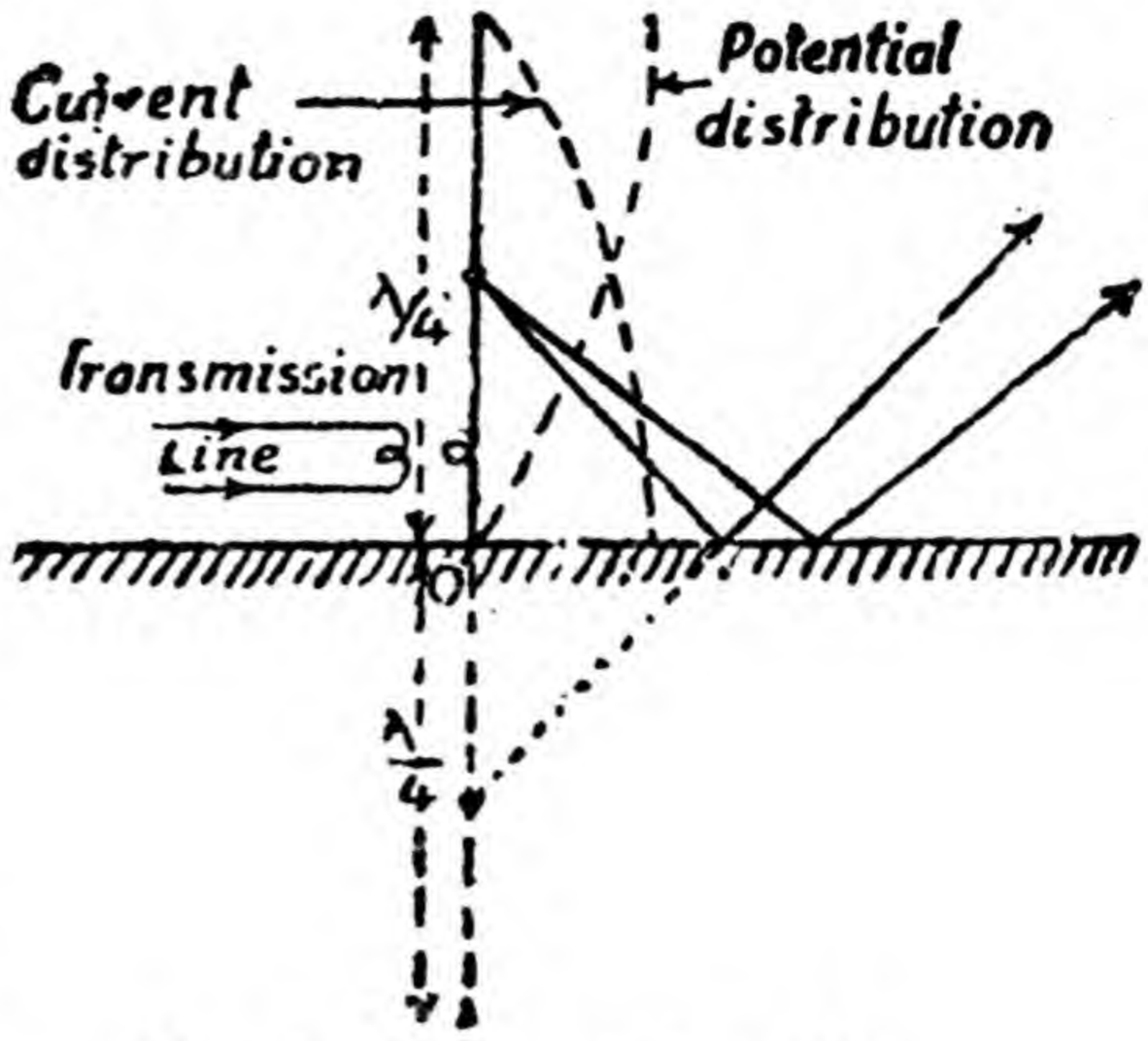


Fig. 21. Marconi aerial.

Due to ground effect, we consider the presence of image antenna below the surface of earth as shown in figure 21. The Marconi aerial, therefore, can be considered as equivalent to half wave dipole. This dipole, *i.e.*, the aerial and its image, radiates through a hemisphere with earth's surface as dimensional plane and not through a hemisphere as in the case of dipole proper.

Refer to equation (11) of art. 28·8, in which radiation resistance of quarter wave antenna is shown to be

$$R_{\lambda/4} = 36{\cdot}57 \text{ ohms}$$

and $$P_{\lambda/4} = 36{\cdot}57 . I^2_{r.m.s.} \text{ watts.}$$

It shows that Marconi aerial must be fed by a transmission line having characteristic impedance almost equal to 37 ohms. In practice, earth is not a perfect conductor, therefore the power radiated by the aerial will depend upon the nature of the soil around the aerial base.

28·9. FIELD STRENGTH AT A POINT CLOSE TO THE ANTENNA WITHIN THE RANGE OF DIRECT RAY :

We have shown previously that the radiative magnetic field due to antenna of length l is given by

$$H = \frac{I_{av}\, l \sin\theta}{2\lambda r} \cos\omega\left(t - \frac{r}{c}\right)$$

$$= \frac{\frac{2I_0}{\pi}\, l \sin\theta}{2\lambda r} \cos\omega\left(t - \frac{r}{c}\right),$$

$$H = \frac{I_0 l_e \sin\theta}{2\lambda r} \cos\omega\left(t - \frac{r}{c}\right), \text{ where } l_e = \frac{2l}{\pi}.$$

Considering only the amplitude part of this expression and

since the factor 'sin θ' expresses the distribution of energy in vertical plane, for direct ray $\theta = 90°$; we find

$$| H | = \frac{I_0 l_e}{2\lambda r}. \qquad \ldots(1)$$

In r.m.s. value the electric field E (amplitude) is related to $| H |$ by the expression

$$\mathcal{E} = \frac{E}{\sqrt{2}} = \frac{120\pi \mid H \mid}{\sqrt{2}}$$

or

$$\mathcal{E} = 120\pi \times \frac{I_{r\cdot m\cdot s}\, l_e}{2\lambda r}$$

$$= \frac{60\pi I_{r\cdot m\cdot s.} l_e}{\lambda r}, \qquad \ldots(2)$$

where $I_{r\cdot m\cdot s.}$ $(= I_0/\sqrt{2})$ denotes the current in the transmitting antenna.

We have introduced the concept of effective length over which the current is constant and if this is l_e for an earthed $\lambda/4$ aerial, the effect of the 'image' is to introduce a radiator of apparent length $2l_e$. The factor 2 is based on the assumption that earth is a perfect conductor, *i.e.* ground absorption is negligible. Therefore, putting $2l_e$ instead of l_e in above expression, we get

$$\mathcal{E} = \frac{120\pi\ I_{r\cdot m\cdot s.}\ l_e}{\lambda . r} \text{ volts/metre}, \qquad \ldots(3)$$

where λ, l_e and r are in metres.

Equation (3) gives an important relation which is widely used for the calculation of approximate field strength at distances extending upto 200 to 300 miles from the transmitting antenna. The above expression can be expressed as

$$\mathcal{E} = \frac{377\ I_{r\cdot m\cdot s.}\ l_e}{\lambda . r} \text{ volts/meter}. \qquad \ldots(4)$$

This field strength can be expressed in terms of power as follows :

For all the rays to be parallel to ground (sin $\theta = 1$ or $\theta = 90°$), antenna should be quite short. Therefore, we can take this to be a case of short vertical grounded antenna. The power due to short doublet is

$$dP = \frac{40\pi^2 I_0^2 l^2}{\lambda^2} \text{ watts}.$$

Applying this relation to short vertical aerial, we get

$$P = \frac{40\pi^2 I^2_{av} l^2}{\lambda^2}.$$

or

$$= \frac{40\pi^2 \left(\frac{2I_0}{\pi}\right)^2 l^2}{\lambda^2},$$

because current distribution on the antenna is sinusoidal. Thus

$$P=\frac{40\pi^2 I_0^2 l_e^2}{\lambda^2},$$

where $$l_e=\frac{2l}{\pi}$$

In case of ground antenna, the effective height will be twice of l_e, so that

$$P=\frac{40\pi^2 I_0^2(2l_e)^2}{\lambda^2}$$

$$=\frac{160\pi^2 I^2{}_0 l_e^2}{\lambda^2}$$

$$=\frac{320\pi^2 I^2{}_{r\cdot m\cdot s.}\ l_e^2}{\lambda^2}.$$

This is the power radiated through a sphere at the centre of which an antenna of length $2l_e$ is placed. But because a grounded antenna radiates through hemisphere, power radiated by it will be half of above value *i.e.*,

$$P=\frac{160\pi^2 I^2{}_{r\cdot m\cdot s.} l_e^2}{\lambda^2} \text{ watts.} \qquad ...(5)$$

Therefore,

$$\frac{\mathcal{E}^2}{P}=\frac{\left(\frac{120\pi I_{r\cdot m\cdot s.}\ l_e}{\lambda r}\right)^2}{\frac{160\pi^2 I^2{}_{r\cdot m\cdot s.}\ l_e^2}{\lambda^2}}$$

$$=\frac{90}{r^2}.$$

$$\therefore \quad \mathcal{E}=\frac{9\cdot 49\sqrt{P}}{r} \text{ volts/meter,} \qquad ...(6)$$

in which P is in watts and r is in meters.

NUMERICAL EXAMPLES :

Ex. 1. *What would be the field strength at* 30 *km. away from a transmitting station of* 25 *kilo watts power* ?

$$\mathcal{E}=\frac{9\cdot 49\sqrt{(25.10^3)}}{30.10^3}$$

$$=3\cdot 16.10^{-3}\sqrt{(250)}$$

$$=47\cdot 96.10^{-3} \text{ volts/meter}$$

$$=47\cdot 96 \text{ millivolts/meter.}$$

Ex. 2. *Find the effective height of an aerial given that a field strength of* 1·5 *mV/metre is produced at a distance of* 50 *kilometers by the transmitter operating at* 150 *kc./s. with an aerial current of* 25 *amperes.*

From equation (4), the field strength is given by

$$\mathcal{E}=\frac{377\ I_{r\cdot m\cdot s.}\ l_e}{\lambda r} \text{ volt/meter,}$$

were λ, r and l_e are in meters and current in amperes. It is given

that $$\mathcal{E}=1\cdot5.10^{-3} \text{ volt/meter}$$
$$=15.10^{-4} \text{ V./m.,}$$
$$\lambda=\frac{c}{n}=\frac{3\cdot10^{10}}{150.10^{7}}=2.10^{3} \text{ meters.}$$
$$r=50.10^{3} \text{ metres,}$$
$$I_{r\cdot m\cdot s.}=25 \text{ amperes.}$$

The effective length will be

$$l_e=\frac{\mathcal{E}.\lambda r}{377\ I_{r\cdot m\cdot s.}} \text{ metres}$$
$$=\frac{15.10^{-4}.2.10^{3}.50.10^{3}}{377\times25}$$
$$=\frac{6000}{377}=\mathbf{15\cdot9 \text{ metres.}}$$

Ex. 3. *A transmitting aerial operating at a wavelength of* 500 *meters has an effective height of* 75 *meter and the aerial current is* 100 *amperes. What is the field strength at a distance of* 200 *km* ?

The field strength is given by

$$\mathcal{E}=\frac{377\ I_{r\cdot m\cdot s.}\ l_e}{\lambda.r}\ V/m$$
$$=\frac{377\times100\times75}{500\times200\times10^{3}}=\cdot028\ V/\text{m.}$$
$$=\mathbf{28 \text{ V/m.}}$$

Ex. 4. *A short vertical grounded antenna is designed to radiate at* 1 *mc/s. If the effective height of the aerial is* 30 *metres, what would be its radiation resistance* ?

It has been shown that for short vertical grounded aerial,

$$P=\frac{160.\pi^2 I^2_{r\,m\cdot s.}\ l_e^2}{\lambda^2} \text{ watts}$$

and the radiation is expressed by the relation of the form

$$P=I^2_{r\cdot m\cdot s.}\ R.$$

Therefore, required radiation resistance is

$$R=\frac{160\pi^2 l_e^2}{\lambda^2},$$

where l_e and λ are in same units.

It has been given that $\lambda=\frac{c}{n}=\frac{3.10^{8}}{10^{6}}=3.10^{2}$ meters

and $$l_e=30 \text{ metres.}$$

Therefore $$R=\frac{160.(3\cdot14)^2.(30)^2}{(300)^2}=\mathbf{15\cdot8 \text{ ohms.}}$$

28·10. TOTAL EFFECTIVE RESISTANCE AND EFFICIENCY OF AN ANTENNA :

In addition to the radiated energy, energy may also be consumed by an antenna system as a result of loss resistance. Thus the effective total resistance of antenna is made up of two parts :

(1) radiation resistance (R),
(2) loss resistance (R_L).

Since we require better radiation characteristics of an antenna system, it is desired that loss resistance should be as small as possible. This loss resistance is made up of the following components :

(*i*) loss in inferior dielectrics,
(*ii*) brushing losses,
(*iii*) I^2R losses in antenna and ground circuits,
(*iv*) loss by eddy currents in neighbouring conductors,
(*v*) loss in earth connection,
(*vi*) loss by leakage over insulation.

From the study of all these losses, it has been found that with a given aerial and a given current :

(*a*) resistance corresponding to dielectric absorption losses, leakage and brushing decreases as frequency increases and can be considered as inversely proportional to the latter.

(*b*) resistance, corresponding to conductor losses in the aerial and earth system and to eddy current losses, increases as frequency increases and can be taken as proportional to the square root of frequency.

(*c*) It is obvious from the expression deduced previously for radiation resistance that it varies directly as the square of the frequency.

Fig. 22 represents the variation of total aerial resistance with frequency.

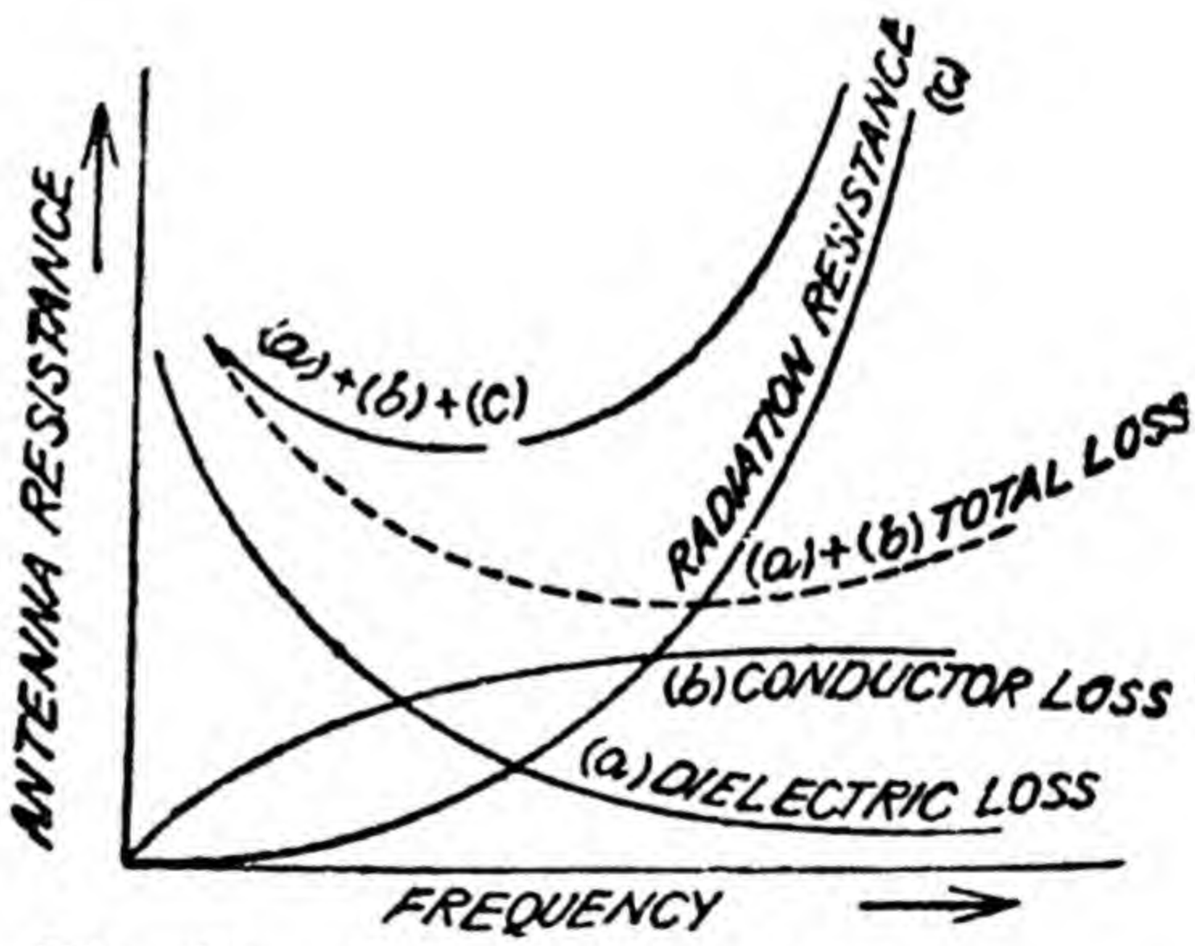


Fig. 22. Variation of total antenna resistance

Thus the total resistance component of antenna impedance is $(R+R_L)$. This total resistance determines the amount of energy that must be supplied to the antenna to produce a given current at the point to which the resistance is referred.

The *efficiency* of an aerial depends upon the ratio of the power radiated to the power supplied. The power radiated is determined interms of radiation resistance (R), while the total resistance $(R+R_L)$ determines the power supplied, with a given aerial. If η denotes the efficiency of the aerial, then

$$\eta = \frac{R}{R+R_L}.$$

Thus the efficiency of an aerial represents the fraction of the total energy supplied to the antenna which is converted into effective radiation from the antenna.

Ex. 1. ***Suppose we want to calculate the efficiency of a grounded aerial, operating at*** 500 *kc/s and having a total resistance of* 12 *ohms with effective height of* 30 *metres.*

Since for the calculation of frequency we require radiation resistance, it is given by

$$R=160\pi^2\left(\frac{l_e}{\lambda}\right)^2=1580\left(\frac{l_e}{\lambda}\right)^2$$

For short vertical grounded aerial, it is given that

$$l_e=30 \text{ metres}$$

$$\lambda=\frac{3.10^{10}}{500.10^3}=600 \text{ metres.}$$

Therefore $R=1580\times\frac{30\times30}{600\times600}$ ohms$=3{\cdot}95$ ohms.

Efficiency is thus $\eta=\frac{3{\cdot}95}{12}={\cdot}329=32{\cdot}9\%$.

Ex. 2. *A low frequency transmitting aerial has a radiation resistance of* 0·5 *ohm and total loss resistance of* 2·5 *ohms. If the current fed into the aerial is* 100 *amperes, calculate the radiated power, the power input and antenna efficiency.*

Power radiated $=RI^2=0{\cdot}5\times100^2$
$=5000W$

Power input $=I^2(R+R_L)=10^4\times3$
$=30{,}000W$

Aerial efficiency $=\frac{\text{Power radiated}}{\text{Power input}}=\frac{5000}{30{,}000}\times100$
$=16{\cdot}7\%$

or aerial efficiency $=\frac{R}{R+R_L}=\frac{0{\cdot}5}{0{\cdot}5+2{\cdot}5}=\frac{0{\cdot}5}{3}$
$=0{\cdot}167=16{\cdot}7\%$.

28·11. DIRECTIVE GAIN :

A radiating system composed of several spaced radiators is termed as an *antenna array*. These arrays, at high frequencies, are employed to achieve greater directive gain. The total field developed by an antenna array at a distant point will be the vector sum of the fields produced by the individual antenna of the array. In order to obtain the greater radiation in the desired direction, antennas of the array are excited in proper phase.

DIRECTIVE GAIN :

Case I : *When actual and reference antennas are fed with different powers, but developing the same field strength at the points under consideration.*

Let us first explain what we mean by the term gain' of an antenna array. The term 'gain' applied to an antenna system is a measure of directivity of the antenna field pattern as compared with some standard antenna which is normally taken as an isotropic

radiator, *i. e.*, an antenna radiating energy uniformly in all directions. Quantitatively, gain is the ratio of power that must be supplied to the standard antenna to deliver a particular field strength in the desired direction to the power that must be supplied to the actual antenna system so as to obtain the same field strength in the same direction. The directive gain can be expressed either as a power ratio or in terms of the equivalent number of decibels.

Assume n spaced aerials, each with current I and radiation resistance R, the phasing and spacing being such that in some direction reinforcement occurs, giving a field strength n times that from one aerial alone. If we neglect the effect of mutual coupling between aerials, then total power radiated by the array will be

$$P_n = n.RI^2.$$

Optimum field strength produced is

$$E_n = K\ (nI)\dagger,$$

where K is a constant, RI^2 and KI represent the power radiated and field produced by one array element respectively.

In order to produce a field strength of value E_n using one element only, it is necessary to use an aerial current I_0 which is n times that in one element of array ; but in that case the power radiated by a single equivalent aerial is given by

$$P_0 = I_0^2 R = (nI)^2\ R,$$

which is greater than P_n. Hence

$$\text{array improvement factor} = \frac{P_0}{P_n} = \frac{R\ (nI)^2}{n.RI^2} = n.$$

If the array improvement factor is expressed in decibels, it is usually called the gain of array, *i.e.*

$$\text{Array gain } G = 10 \log_{10} (P_0/P_n) \text{ dbs.} \qquad \ldots(1)$$

The maximum directive gain, G_{max}. is termed as the *directivity*. D, of an array. Since we specify particular direction in which the array radiates the maximum power, we need not mention the angles θ, ϕ with directivity or in other words it is a constant.

We can proceed to calculate the directive gain in case of short electric doublet. For short electric doublet, total power radiated into space is

$$P_n = \frac{40\pi^2\ I_0^2\ dl^2}{\lambda r} \text{ watts.} \qquad \ldots(2)$$

The electric field strength is given by

$$dE = \frac{60\pi I_0 \sin\theta\ dl}{\lambda r}.\cos\omega\ (t - r/c). \qquad \ldots(3)$$

---

† $|\ E\ | = \dfrac{60\pi I_0\ dl\ \sin\theta}{\lambda r} = KI_0,$

where $K = \dfrac{60\pi\ dl \sin\theta}{\lambda r}$.

If we take the isotropic radiator, E should not involve either θ or ϕ because radiation in every direction is uniform. Therefore, for isotropic radiator, we take

$$dE=\frac{60\pi I_0\, dl}{\lambda r}\cos\omega\,(t-r/c).$$

Obviously, electrostatic fields radiated by doublet and isotropic radiators have been taken to be equal to satisfy the definition of directive gain.

The power radiated by the isotropic radiator is then

$$dP=\tfrac{1}{2}dE.dH=\tfrac{1}{2}.\frac{dE^2}{120\pi}\text{ watts/sq. meter}$$

or total power radiated into space

$$*P_0=\tfrac{1}{2}\,\frac{dE^2}{120\pi}.4\pi r^2\text{ watts.}$$

or
$$P_0=\tfrac{1}{2}.\frac{(60\pi I_0\, dl)^2}{120\lambda^2\,\pi r^2}.4\pi r^2\text{ watts}$$

$$=\frac{60\pi^2\, I_0^2\, dl^2}{\lambda^2}\text{ watts.}\qquad\ldots(4)$$

From equations (2) and (4), we write

$$\frac{P_0}{P_n}=\frac{6}{4}=1{\cdot}5.$$

Therefore the directive gain of short doublet antenna, on using equation (1), is

$$G=10\log_{10}1{\cdot}5$$
$$=1{\cdot}75\ db.\qquad\ldots(5)$$

Since G is a constant, equal to 1·75, it can be termed as *directivity* of doublet antenna.

Similarly, to compute the directivity of half-wave dipole, we note that power radiated by it is

$$P_n=36{\cdot}5I_0^2\text{ watts,}\qquad\ldots(6)$$

and the field strength is

$$E=\frac{60I_0}{r}\left[\frac{\cos\left(\frac{\beta l\cos\theta}{2}\right)-\cos\frac{\beta l}{2}}{\sin\theta}\right]\cos\omega\,(t-r/c).$$

For isotropic radiator E must be independent of θ and consequently

$$E=\frac{60I_0}{r}\cos\omega\,(t-r/c).$$

so that total power radiated by an isotropic radiator is

$$P_0=\tfrac{1}{2}.\frac{E^2}{120\pi}.4\pi r^2\text{ watts}$$

---

*Factor $\frac{1}{2}$ appears for the average of $\cos^2\omega\,(t-r/c)$.

$$=60I_0^2 \text{ watts.} \qquad ..(7)$$

Applying equation (1), directive gain is found to be

$$G=10 \log_{10} \frac{60I_0^2}{36{\cdot}5I_0^2}$$

$$=10 \log_{10} 1{\cdot}6$$

$$=2{\cdot}15\ db. \qquad (8)$$

which can also be called as '*Directivity*' of a half-wave dipole.

From equations (5) and (8), we note that the maximum directive gain of a half wave dipole is only 0·4 *db* greater than for a short electric doublet.

Ex. 1. *A transmitter is fed with* 100*W of power and produces the same field strength at a given point as a* $\lambda/2$ *dipole fed with* 200*kW of power. Calculate the gain of the aerial (a) relative to* $\lambda/2$ *dipole, (b) relative to an isotropic radiator.*

(*a*) Gain of aerial relative to $\lambda/2$ dipole

$$=10 \log_{10} \frac{P_0}{P_n}\ db.$$

$$=10 \log_{10} \frac{200\times10^3}{100\times10^3}$$

$$=10\times0\ 301=3{\cdot}01\ db$$

(*b*) We have shown that $\lambda/2$ dipole is 2·15 *db* more directional than isotropic radiator. Therefore gain of the aerial relative to isotropic radiator

$$=2{\cdot}15+3{\cdot}01$$

$$=5{\cdot}16\ db$$

Ex. 2. *To produce a power density of* 1·2 *mW/m*^{2} *in a given direction, at a distance of* 2·0 *km, an antenna radiates a total of* 200*W. An isotropic antenna would have to radiate* 2500*W to produce the same power density at that distance. What, in decibels, is the directive gain of the practical antenna ?*

Directive gain of practical antenna can be calculated by using the ratio

$$=\frac{\text{Power radiated by isotropic antenna}}{\text{Power radiated by the practical antenna}}$$

$$=\frac{2500}{200}$$

Therefore directive gain of practical antenna is given by

$$G=10 \log_{10} \frac{2500}{200}$$

$$=10 \log_{10} 12{\cdot}5$$

$$=10{\cdot}97\ db$$

Case II. *When standard actual antennas are fed with the same power but developing different strengths at the point under consideration :*

Refer to art. 28·13 in which gain in field intensity $G_f(\phi)$ has been calculated by taking two antennas fed with same power.

Suppose the actual antenna system, consisting n elements and each element carrying a current I, is fed with power P, so that

$$P_n = n.RI^2,$$

where R is the radiation resistance of each element.

Then we take the standard antenna fed with the same power and having the same radiation resistance, so that current in standard antenna should be

$$I_0^2 = \frac{P_0}{R} = \frac{P_n}{R}$$

$$= \frac{n.RI^2}{R}$$

$$= nI^2.$$

Therefore $$I_0 = \sqrt{n}.I. \qquad \ldots(9)$$

Then the field developed at a point due to the actual antenna systems having n elements fed with power P_n, and each element carrying current I will be

$$E_n = n\,(KI). \qquad \ldots(10)$$

At the field due to standard antenna at the same point will be

$$E_0 = KI_0 = K\sqrt{n}.I. \qquad \ldots(11)$$

Therefore array improvement factor equations from (10) and (11) is

$$n = \frac{E_n^2}{E_0^2}.$$

so that power gain is

$$G_p(\phi) = 10 \log_{10} n = 10 \log_{10} \frac{E_n^2}{E_0^2}$$

$$= 10 \log_{10} \frac{P_n}{P_0}. \qquad \ldots(12)$$

for the system fed with same power. Compare equation (12) with equation (1). We can, for the present case, define the gain in field intensity as follows :

$$G_f(\phi) = \frac{\text{Field at a point due to actual antenna system fed with power } P \text{ (say)}}{\text{Field at the same point due to standard antenna system fed with the same power } P}$$

$$= 20 \log_{10} \frac{E_n}{E_0}.$$

POWER GAIN :

While defining directive gain, no account of losses due to ohmic heating, r. f. heating, mismatched antenna etc. has been taken. The only consideration was that of radiation pattern. Therefore to include these losses which lower the antenna efficiency, we define another term called power gain, denoted by A_p. It is defined as

$$A_p = \frac{\text{Maximum radiation intensity from actual antenna}}{\text{Radiation intensity from (lossless) isotropic source with same power input.}}$$

$$= \eta G,$$

where η is the radiation efficiency of the actual antenna and G is the directive gain.

The directive gain is always greater than the power gain and is significant for coverage, accuracy or resolution considerations.

Ex. 1. *An antenna has a radiation resistance of 72 ohms, a loss resistance of 10 ohms, and a power gain of 15. What is its directivity ?*

The efficiency, η, is given by

$$\eta = \frac{R}{R+R_L} = \frac{72}{72+10}$$

$$= 0{\cdot}878$$

Therefore directive gain is

$$G = \frac{A_p}{\eta} = \frac{15}{0{\cdot}878}$$

$$= 17{\cdot}08.$$

28·12. LINEAR ARRAYS OF n ISOTROPIC SOURCES OF EQUAL AMPLITUDE AND SPACING :

In the case of elementary dipole, we note that very little directivity is achieved. Maximum radiation takes place at right angles to the axis which falls off relatively slowly as the polar angle decreases towards zero. Further, although we find that half-wave dipole achieves a somewhat greater concentration of energy in the direction normal to the axis, its pattern does not differ substantially from that of the elementary dipole. Also, resonant wire antennas of lengths greater than $\lambda/2$, while producing more complicated patterns, do not result in highly directive patterns. Therefore one way in which greater directive patterns may be achieved is by the use of an array of dipole (or other) antennas. The resultant pattern due to such an array is the superposition of field contributions from each constituent element.

Fig. 2 shows a linear array composed of n isotropic point sources or equal amplitude and spacing, where n is an positive integer. The total field E, at a large distance, in the direction ϕ is given by

$$E=1+e^{j\psi}+e^{2j\psi}+e^{3j\psi}\ldots+e^{j(n-1)\psi}, \qquad \ldots(1)$$

where ψ is the total phase difference in the fields from any two *adjacent* sources and is given by

$$\psi=\frac{2\pi d}{\lambda}\cos\phi+\delta=dr\cos\phi+\delta, \text{ where } dr=\frac{2\pi d}{\lambda},$$

and δ is the phase difference of adjacent sources due to currents flowing in them. It has been assumed that the amplitudes of the fields from the sources are all equal and taken as unity. Source 1 has been taken as the phase centre so that the field from source 2 is advanced in phase by ψ, the field from source 3 is achieved by 2ψ, etc.

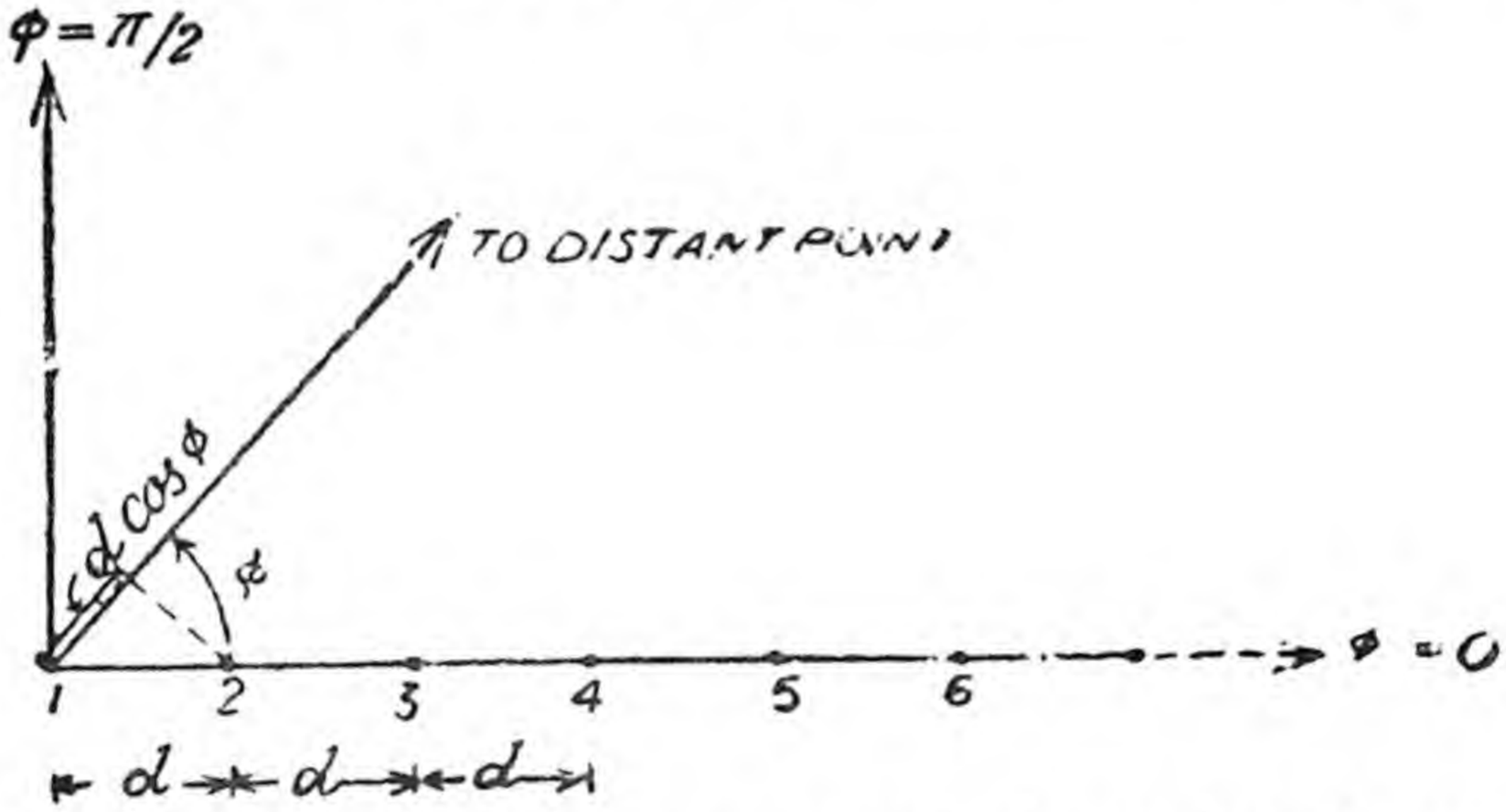


Fig. 23. Arrangement of linear array of *n* isotropic point sources.

Multiplying (1) by $e^{j\psi}$, we get

$$Ee^{j\psi}=e^{j\psi}+e^{2j\psi}+\ldots+e^{nj\psi}. \qquad \ldots(2)$$

Subtracting (2) from (1) and dividing by $(1-e^{j\psi})$, we arrive at

$$E=\frac{1-e^{jn\psi}}{1-e^{j\psi}}$$

$$=\frac{e^{jn\psi/2}\left[e^{jn\psi/2}-e^{-jn\psi/2}\right]}{e^{j\psi/2}\left[e^{j\psi/2}-e^{-j\psi/2}\right]}$$

$$=e^{j\psi}\frac{\sin n\psi/2}{\sin\psi/2}=\frac{\sin n\psi/2}{\sin\psi/2}\Big/\underline{\xi} \qquad \ldots(3)$$

where value of ξ is given by

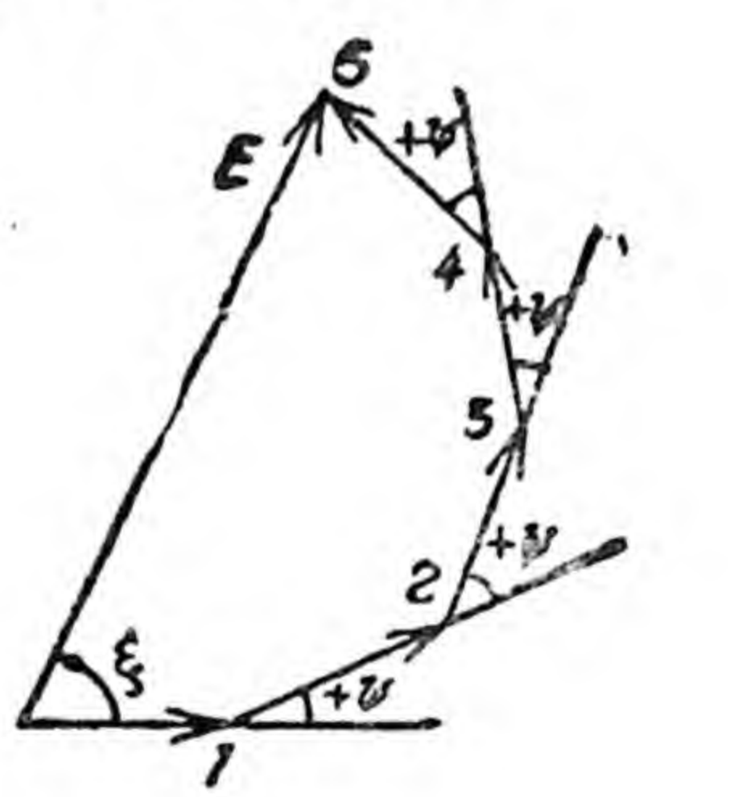
$$\xi=\frac{n-1}{2}\psi, \qquad \ldots(4)$$

and is referred to the field from source 1.

If the phase is referred to the centre point of the array, say at point 3 of the array as shown in fig. 24, then equation (3) becomes

$$E=\frac{\sin n\psi/2}{\sin\psi/2} \qquad \ldots(5)$$

since $\xi=0$.



Ftg. 24. The top figure refers to the case when phase is taken with reference to the origin at one end of the array ; whereas bottom figure refers to the case when phase is referred to the centre of the array.

When $\psi=0$ equation (5) becomes indeterminate, so that for this case E must be obtained as the limit of (5) as ψ approaches zero,

$$\text{i.e} \quad \lim_{\psi\to 0} E = \lim_{\psi\to 0} \frac{\cos\frac{n\psi}{2}\times\frac{n}{2}}{\cos\frac{\psi}{2}\times\frac{1}{2}} = n. \qquad \ldots(6)$$

This is the maximum value of the field which E can attain because when fields from all the sources meet in the same phase, phase difference between them is zero ($\psi=0$), and the resultant field is maximum in that direction, *i.e.*,

$$E_{max.}=n.$$

Hence, the normalised value of the total field would be

$$E_n=\frac{E}{E_{max.}}=\frac{1}{n}\,\frac{\sin(n\psi/2)}{\sin(\psi/2)}. \qquad \ldots(7)$$

The maximum value which E_n can attain is *unity* because then the field E will assume its maximum value.

To illustrate some of the properties of linear arrays equation (7) will now be applied to following special cases :

Case I. Broad side array (sources in phase) :

When sources are of same amplitude and phase, $\delta=0$, and

$$\psi=dr\cos\phi. \qquad \ldots(8)$$

For $\psi=0$, $\phi=(2K+1)\frac{\pi}{2}$, where $K=0, 1, 2, 3\ldots$;

The field is therefore maximum when

$$\phi=\frac{\pi}{2} \text{ and } \frac{3\pi}{2}$$

Thus the field is maximum in a direction *normal* to the array. This condition, which is characterised by *inphase sources* ($\delta=0$), results in a 'broad side' type of array.

Example : Now we shall sketch the pattern of a broadside array of four inphase isotropic point sources of equal amplitude. The spacing between sources is taken to be half wavelength, *i.e.*,

$$d=\lambda/2 \text{ and } n=4$$

Thus $$\psi=\frac{2\pi}{\lambda}\,d\cos\phi=\pi\cos\phi.$$

Now we shall determine the field strength variation :

(*a*) Maximum field direction

$$E_n=\frac{1}{n}\frac{\sin n\psi/2}{\sin\psi/2}=\frac{1}{4}\frac{\sin(2\pi\cos\phi)}{\sin(\frac{1}{2}\pi\cos\phi)}$$

$$=\cos\left(\frac{\pi}{2}\cos\phi\right)\cos(\pi\cos\phi)$$

For E_n to be maximum, E should be equal to E_{max}. That is $E_n=1$, or

$$\cos\left(\frac{\pi}{2}\cos\phi\right)\cos(\pi\cos\phi)=1$$

or $$\cos\left(\frac{\pi}{2}\cos\phi\right)=1$$

and $\cos(\pi\cos\phi)=1$.

Both conditions are satisfied when

$$\phi=(2K+1)\frac{\pi}{2},$$

where $K=0, 1, 2, 3,\ldots$

(*b*) To obtain secondary maximum field strength lobe, we put $\frac{dE_n}{d\phi}=0$. This results in following directions :

$$\phi=42°,\ (\pi\pm42°).$$

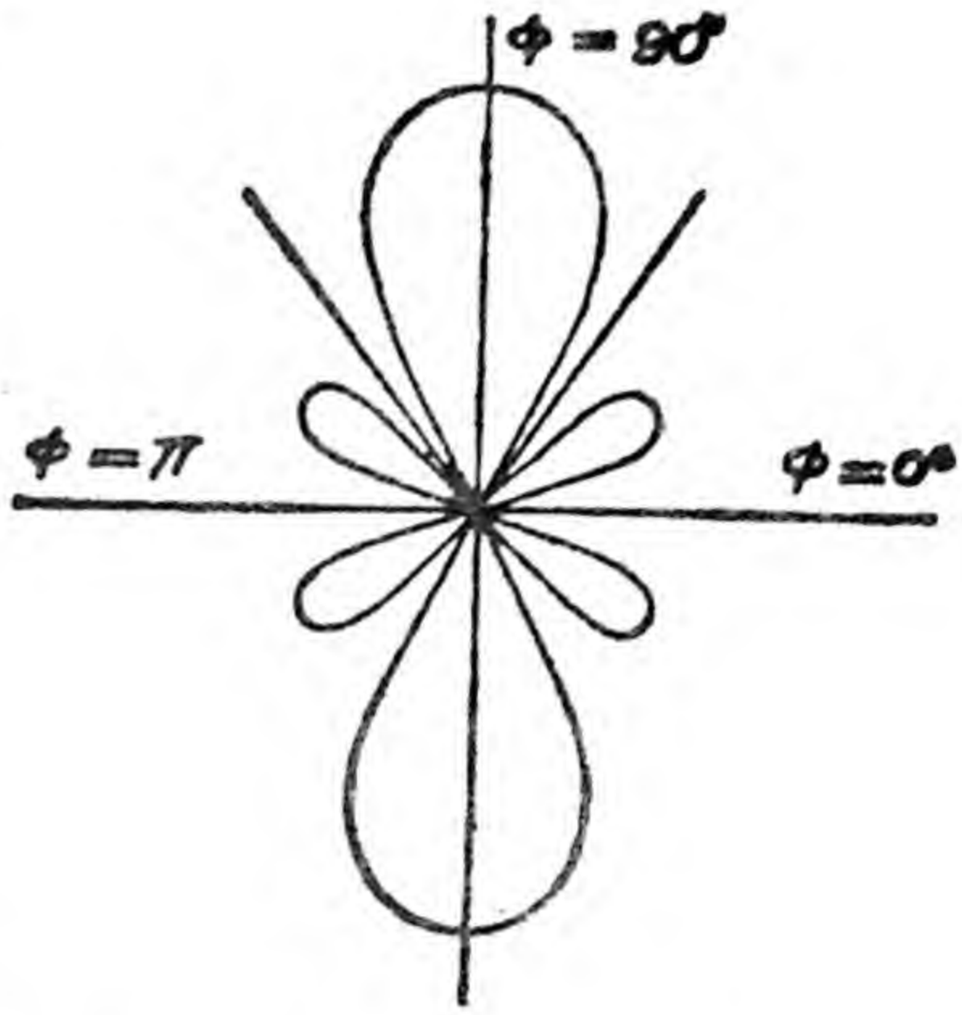


Fig. 25. Field pattern (Broad side array) for $n=4$, $d=\lambda/2$.

(c) For field strength to be minimum, we must put

$$E_n=0$$

or $$E_n=\cos\left(\frac{\pi}{2}\cos\phi\right)\cos(\pi\cos\phi)=0$$

$$\cos\left(\frac{\pi}{2}\cos\phi\right)=0, \qquad \text{or} \quad \cos(\pi\cos\phi)=0.$$

or $$\frac{\pi}{2}\cos\phi=\frac{\pi}{2}, \qquad \text{or} \quad \pi\cos\phi=\frac{\pi}{2},$$

giving $$\phi=K\pi, \qquad \text{giving } \phi=\left(K\frac{\pi}{2}-\frac{\pi}{3}\right),$$

where $K=0, 1, 2, 3,$ $\phi=0, 180°, 360°,$ where $K=0, 2, 4,$ $\phi=-60°, 120°, 240°.$

Now on the basis of above analysis of field strength, the pattern can be drawn as shown in the figure 25.

Case 2. Ordinary end-fire array :

An array in which the maximum field direction is *along* the array, *i.e.* in a direction of $\phi=0$, is called end fire array. For this we substitute the conditions $\psi=0$, $\phi=0$, into

$$\psi=\frac{2\pi d}{\lambda}\cos\phi+\delta,$$

i.e. $0=dr\cos 0°+\delta,$

$$\delta=-dr. \qquad \ldots (9)$$

We, thus, conclude

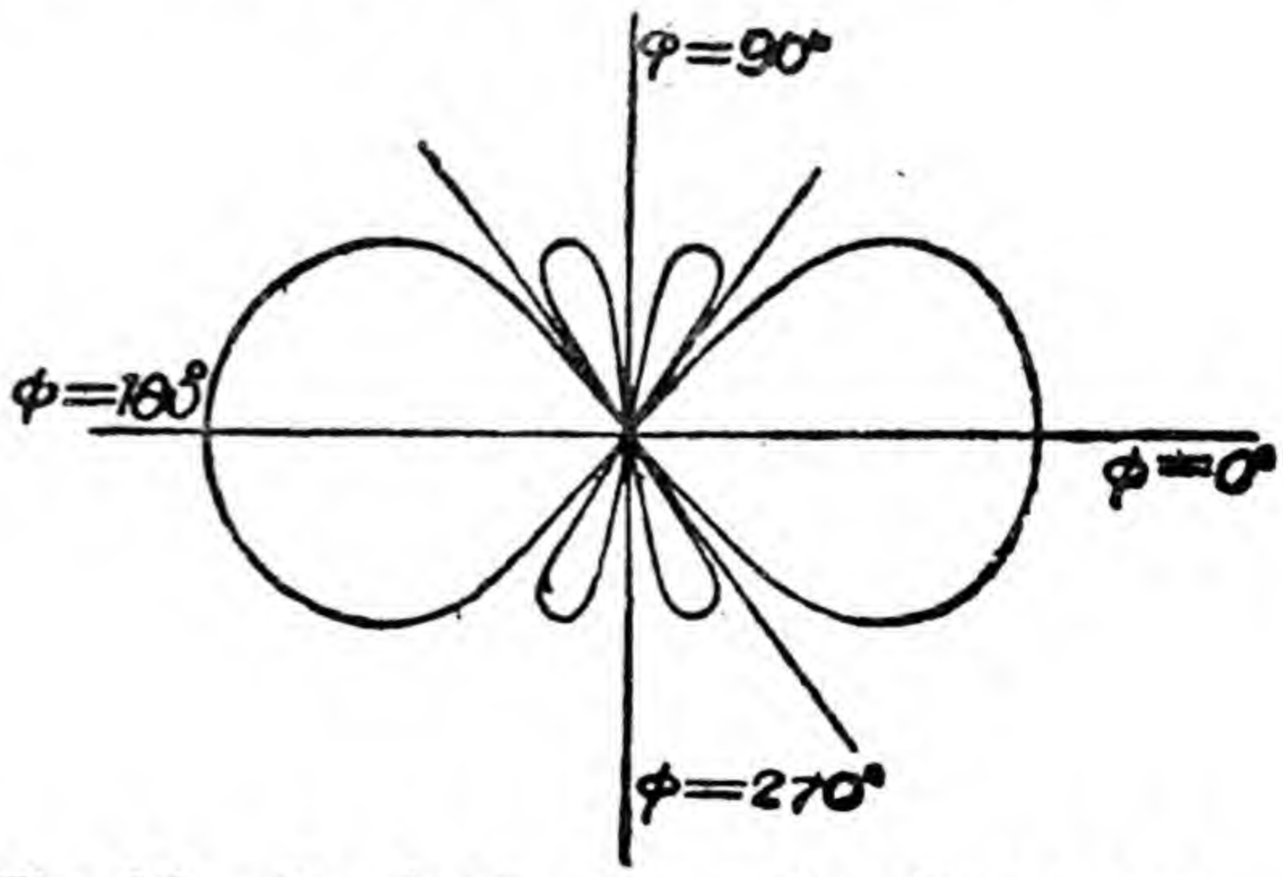


Fig. 26. The field pattern of ordinary end fire array for $n=4$, $d=\lambda/2$, $\delta=-\pi$.

that for an end fire array, the phase between sources is ***retarded progressively by the same amount as the spacing between sources in radians.***

Example : The field pattern of an end fire array of four isotropic point sources with spacing $\lambda/2$ and $\delta=-\pi$ is sketched in figure 26. The maximum and minimnm values of the field are calculated as in previous case *i.e.*

For $n=4$.

$d=\lambda/2$, So $\psi=\pi(\cos\phi-1)$.

$\delta=-\pi$.

$$\text{Therefore } E_n=\frac{1}{4}\frac{\sin 2\pi(\cos\phi-1)}{\sin\frac{\pi}{2}(\cos\phi-1)}$$

$$=\cos\frac{\pi}{2}(\cos\phi-1)\cos\pi(\cos\phi-1).$$

(*a*) For E_n maximum,

$$\cos\frac{\pi}{2}(\cos\phi-1)\cos\pi(\cos\phi-1)=1,$$

which gives $\phi=0, \pi, 2\pi,\ldots$

(*b*) For E_n minimum, $E_n=0$

$$\text{or} \quad \cos\frac{\pi}{2}(\cos\phi-1)\cos\pi(\cos\phi-1)=0,$$

$$\text{either} \quad \cos\frac{\pi}{2}(\cos\phi-1)=0$$

$$\text{which gives} \quad \phi=\frac{\pi}{2}, \frac{3\pi}{2},\ldots$$

$$\text{or} \quad \cos\pi(\cos\phi-1)=0,$$

which gives $\phi=60°, 120°, 240°$.

(*c*) The conditions for secondary lobes can be obtained on putting $\frac{dE_n}{d\phi}=0$.

The same shape of the field pattern is obtained in this case if

$\delta=+\pi$,

$d=\lambda/2$,

$n=4$.

This pattern as shown in figure 26 is bidirectional. If the spacing is less than $\lambda/2$, the maximum radiation is in the direction $\phi=0$, when $\delta=-dr$ and in the direction $\phi=180°$ when $\delta=+dr$.

Case 3. End fire array with increased directivity :

Hansen Woodyard predicted that a large directivity may be achieved by increasing the phase change between sources, so that

$$\delta=-(dr+\pi/n) \qquad \ldots(10)$$

and is referred as the condition for increased directivity. Thus for the phase differcnce of the fields at a large distance, we have

$$\psi=dr\ (\cos\phi-1)-\frac{\pi}{n}. \qquad \ldots(11)$$

Example : The field pattern of an endfire array of four isotropic point sources for this case is illustrated in fig. 27. The spacing between the sources is $\lambda/2$ and thus $\delta=-5\pi/4$. The conditions for the array pattern being, thus, the same as in previous case of ordinary end fire array except that the phase difference is increased by $\pi/4$. This additional phase difference yields considerably *sharper lobe* in the direction of $\phi=0$.

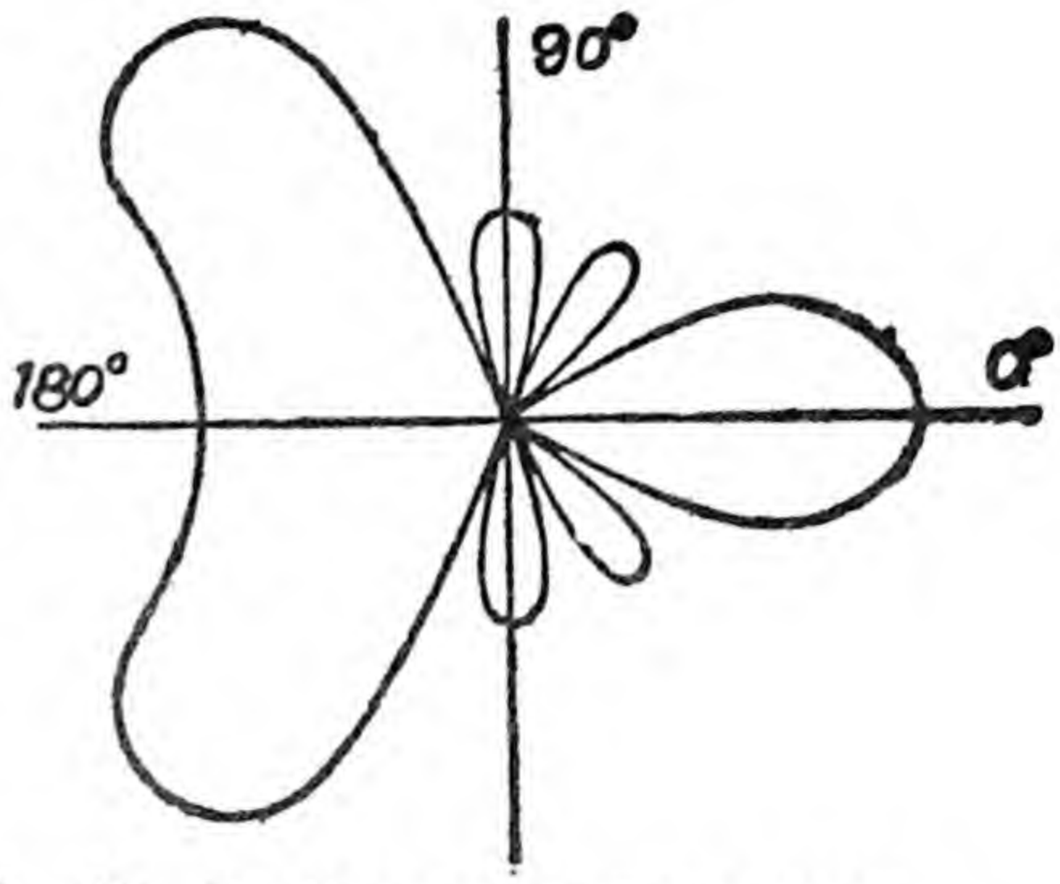


Fig. 27. Endfire array field pattern with increased directivity for $d=\lambda/2$, $n=4$, $\delta=-5\pi/4$.

The large value of spacing results in too great a range in ψ due to which back lobes in this case are excessively large. If the spacing is reduced then $|\psi|$ can be restricted in its range to a value of π/n at $\phi=0$ and a value in the vicinity of π at $\phi=180°$. Due to this restriction, increase in directivity because of additional phase difference, can be realised.

BEAMWIDTH OF MAIN LOBE :

(a) Broad side array :

Consider a lobe having first null direction defined by the angle ϕ_0. If θ is complementary angle of ϕ_0 which defines the null direction, then 2θ gives the beamwidth of lobe concerned (fig. 28).

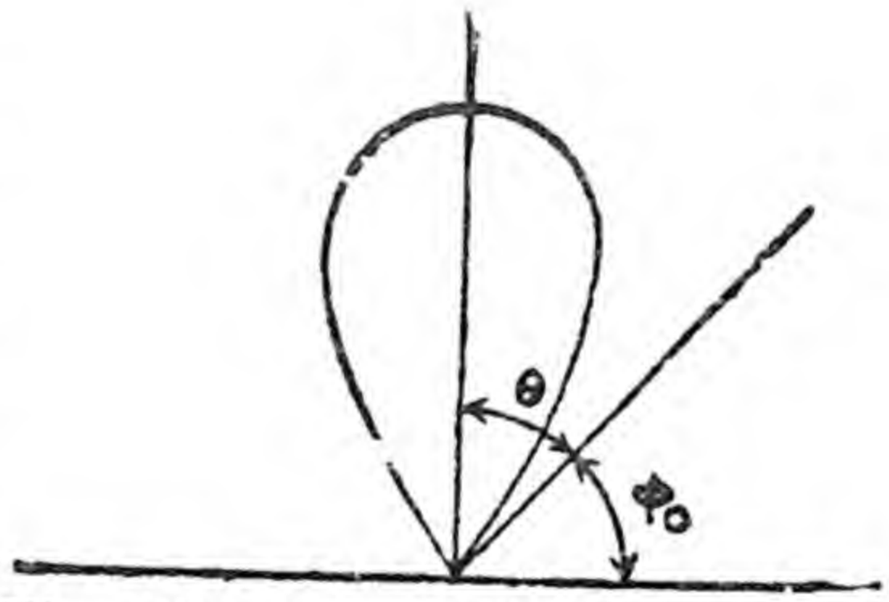


Fig. 28. Beamwidth of main lobe of broad side array.

Null direction of any array of n isotropic radiators of equal amplitude and phase occurs when

$$E_n=0 \qquad \text{that is } \frac{\sin(n\psi/2)}{\sin(\psi/2)}=0$$

or

$$\sin(n\psi/2)=0. \qquad \ldots(1)$$

since

$$\psi=\frac{2\pi d}{\lambda}\cos\phi_0=\frac{2\pi d}{\lambda}\sin\theta.$$

Putting in equation (1), we get

$$\sin\left(\frac{2\pi d}{\lambda}\cdot\frac{n}{2}\sin\theta\right)=0,$$

$$\frac{n}{2}\left(\frac{2\pi d\sin\theta}{\lambda}\right)=K\pi, \qquad \text{where } K=1, 2, 3.$$

For first null direction $K=1$; therefore

$$\sin\theta=\frac{\lambda}{nd}, \qquad \text{or} \quad \theta=\sin^{-1}\lambda/nd.$$

So the beamwidth will be given by $2\theta=2\sin^{-1}(\lambda/nd)$. ...(2)

When the spacing is large, λ/nd will be small and eq. (2) can be written as

$$2\theta=2\lambda/nd.$$

If it is given that $d=\lambda/2$, $n=10$, then the required beamwidth will be

$$2\theta=2\times\frac{\lambda}{10.\lambda/2}=\frac{2}{5}=0{\cdot}4.$$

(b) End fire array :

For end fire array the principal maximum is at $\phi=0$, and the first null is at $\phi=\phi_0$. Further for end fire array,

$$\psi=\frac{2\pi d}{\lambda}\cos\phi+\delta, \qquad ...(3)$$

with $$\delta=-\frac{2\pi d}{\lambda},$$

so that $$\psi=\frac{2\pi d}{\lambda}(\cos\phi-1). \qquad ..(4)$$

From equation (1), we note that for null direction,

$$\sin\left(\frac{n\psi}{2}\right)=0,$$

or $$\frac{n\psi}{2}=\pm K\pi, \quad \text{where } K=1, 2, 3... \qquad ...(5)$$

From equation (4) and (5), we find that

$$\frac{2\pi d}{\lambda}(\cos\phi-1)=-\frac{2K\pi}{n}$$

or $$\sin^2\frac{\phi}{2}=\frac{\lambda K}{2nd}$$

or $$\phi=2\sin^{-1}\sqrt{\left(\frac{\lambda K}{2nd}\right)}$$

For first null direction $\phi=\phi_0$ and $K=1$, we find

$$\phi_0=2\sin^{-1}\sqrt{\left(\frac{\lambda}{nd}\right)}$$

so that beamwidth of main lobe will be

$$2\phi_0=4\sin^{-1}\sqrt{\left(\frac{\lambda}{2nd}\right)}$$

When the spacing is large, $\frac{\lambda}{2nd}$ will be small, so that

$$2\phi_0=4\sqrt{\left(\frac{\lambda}{2nd}\right)}$$
$$\approx 2\sqrt{\left(\frac{2\lambda}{nd}\right)},$$

gives the beamwidth of main lobe for the long end fire array.

28·13. CONTINUOUS ARRAYS :

Continuous arrays of point sources consist of an infinite number of sources separated by infinitesimal distances. We shall derive

an expression for the radiation field for such an array consisting point sources of uniform amplitude and of the same phase (*i.e.* broadside case).

Let the array be of length a and is parallel to y-axis with its centre at the origin fig. 29 (a). The field developed by an infinitesimal length dy of the array at a distant point P in a direction θ will be

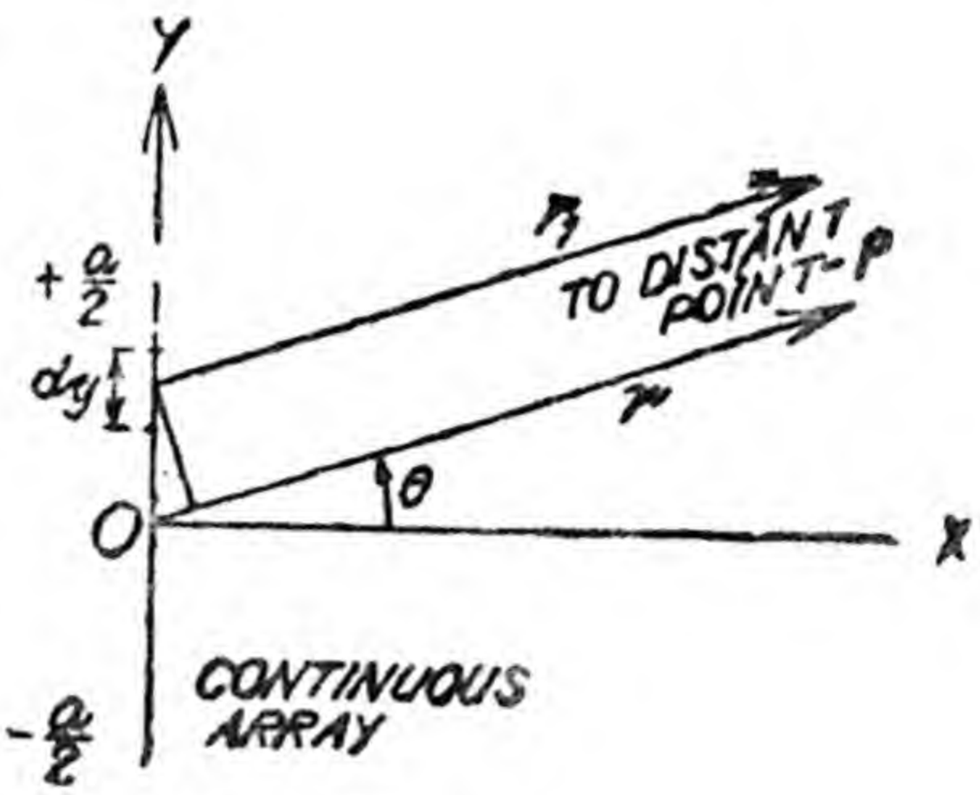


Fig. 29. (a)

$$*dE=\frac{A}{r_1}\,e^{j\omega(t-r_1/c)}\,dy.$$

$$=\frac{A}{r_1}\,e^{j(\omega t-\beta r_1)}.dy$$

on putting $\frac{\omega}{c}=\frac{2\pi}{\lambda}=\beta$; A is a constant involving amplitude.

Therefore, the total field at the point P will be an integral taken over the entire length of array, *i.e.*

$$E=\int_{-a/2}^{a/2}\frac{A}{r_1}\,e^{j(\omega t-\beta r_1)}\,dy$$

Referring the fig. 29 (a),

$$r_1=r-y\sin\theta,$$

so that

$$E=\int_{-a/2}^{+a/2}\frac{A}{(r-y\sin\theta)}\,e^{j\{\omega t-\beta(r-y\sin\theta)\}}\,dy.$$

When r is sufficiently large as compared to length a of the array, we can keep simple r in the amplitude term, *i.e.*,

$$E=\int_{-a/2}^{a/2}\frac{A}{r}\,e^{j(\omega t-\beta r)}\,e^{j\beta y\sin\theta}\,dy$$

$$=\frac{A}{r}\,e^{j(\omega t-\beta r)}\int_{-a/2}^{a/2}e^{j\beta y\sin\theta}\,dy.$$

Putting $A'=\frac{Ae^{j(\omega t-\beta r)}}{r}$ and integrating, we get

$$E=\frac{2A'}{\beta\sin\theta}\cdot\frac{e^{j\beta a\sin\theta/2}-e^{-j\beta a\sin\theta/2}}{j}$$

$$=\frac{2A'}{\beta\sin\theta}\sin\left(\frac{\beta a\sin\theta}{2}\right)$$

---

*We have seen that in the case of current element

$$dE=\frac{60\pi I_0\,dl\cos\theta}{\lambda r}\cos\omega(t-r/c).$$

cos θ instead of sin θ has been written because in this particular article θ is not being measured from the line of array, unlike to current element case.

So

Let $\psi' = \beta a \sin\theta = a_r \sin\theta$, where $a_r = \beta a = 2\pi a/\lambda$ which is array length in radians. Then

$$E = \frac{2A'}{\beta \sin\theta} \sin \psi'/2.$$

Putting ψ'/a for $\beta \sin\theta$, we get

$$E = aA' \frac{\sin \psi'/2}{\psi'/2}$$

For maximum value of E, fields from all the elements of the array must reach in the same phase at the point P *i.e.* for $E_{max.}$, ψ' should be taken as zero. Since expression for E assumes indeterminate form for $\psi' = 0$, we write that

$$\lim_{\psi \to 0} \frac{E}{} = E_{max.} = aA' \frac{\cos \psi'/2 . \frac{1}{2}}{\frac{1}{2}}$$

$$= aA'.$$

Therefore,

$$E_{normalized} = \frac{E}{E_{max}} = \frac{\sin \psi'/2}{\psi'/2}, \qquad \ldots(1)$$

which gives the radiation field of a continuous broad side array of length a, having uniform phase and amplitude.

For n discrete, equally spaced sources, we have shown previously that

$$E_n = \frac{1 \sin n\psi/2}{n \sin \psi/2}, \qquad \ldots(2)$$

where $\psi = dr \cos\phi + \delta$. (For inphase sources, $\delta = 0$.)

Figure of discrete point sources is shown in fig. 29 (*b*). If we compare fig. 29 (*b*) with fig. 29 (*a*), we find that

$$\phi = \pi/2 + \theta,$$

because ϕ is the angle measured from the line of the array. Therefore

$$\psi = dr \cos(\pi/2 + \theta)$$

$$= -dr \sin\theta$$

$$= -\beta d \sin\theta.$$

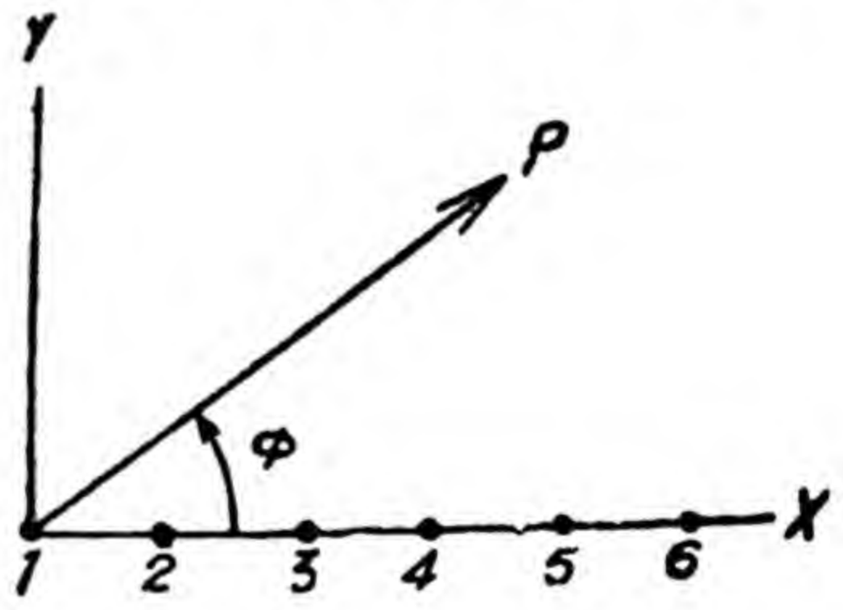


Fig. 29 (*b*)

When ψ is small enough (which is the case when θ, d are small) we can write equation (2) as

$$E_n = \frac{\sin n\psi/2}{n\psi/2} = \frac{\sin\left(\frac{\beta n d \sin\theta}{2}\right)}{\frac{\beta n d \sin\theta}{2}}. \qquad \ldots(3)$$

Now d is the distance between two adjacent sources and there are n sources in all, then the length of the array will be

$$a = (n-1)\, d$$

when $n \gg 1$ *i.e.* number of sources is *infinitely* large, so that

$$a \approx nd,$$

and then equation (3) becomes

$$E_n = \frac{\sin\left(\frac{\beta a \sin\theta}{2}\right)}{\frac{\beta a \sin\theta}{2}}$$

$$= \frac{\sin \psi'/2}{\psi'/2},$$

which is identical with the value of E_n obtained in equation (1) for the continuous array. Thus we have established that field pattern for an array of many discrete sources ($n \gg 1$) and for small values of ψ, is the same as the field pattern of a continuous array of the same length. Further we note that ψ' and ψ both assume small values for small values of θ. In broadside array case, beamwidth of main lobe $2\theta = \frac{2\lambda}{nd}$ will be small if array is sufficiently long, *i.e.* when $nd \gg \lambda$. This means when array is long, θ will be small which will result in small value of ψ and hence assumption made for equation (3) is justified, or in otherwords, the characteristics of the pattern of a long array are the same whether the array has many discrete sources or is a continuous distribution of sources.

The null direction θ_0 of the continuous array pattern is given by $E_n = 0$,

or $$\sin\frac{\psi'}{2} = 0$$

$$\frac{\psi'}{2} = \pm K\pi,$$

where $K = 1, 2, 3\ldots,$

Putting $\psi' = \beta a \sin\theta$, we get

$$\frac{\beta\pi a \sin\theta_0}{2} = \pm K\pi \quad (\theta = \theta_0 \text{ for null direction})$$

or $$\frac{2\pi a \sin\theta_0}{2\lambda} = \pm K\pi$$

or $$\theta_0 = \sin^{-1}\left(\pm\frac{K\lambda}{a}\right)$$

When $a \gg \lambda$, *i,e.* array is sufficiently long,

$$\theta_0 \simeq \pm\frac{K\lambda}{a} \text{ rad.}$$

$$= \pm\frac{57\cdot 3K\lambda}{a} \text{ deg.} \qquad \ldots (4)$$

The beamwidth between first nulls ($K = 1$) for a long array is then

$$2\theta_{01} = \frac{2\lambda}{a} \text{ rad.} = \frac{115\lambda}{a} \text{ deg.} \qquad \ldots (5)$$

It can be noted that if we put $a=nd$, equation (4) is quite identical with the expression for beamwidth of broad side array of n isotropic point sources. Thus for infinite number of sources $n \gg 1$, null locations for arrays of either discrete or continuous are the same.

In fig. 29 (c) field pattern of the main beam of continuous array of point sources of lengths 5λ, 10λ and 50λ are plotted. We note that the beamwidth between half power points θ_{HP}, of a long uniform broad side array is related to θ_{01} as

$$\theta_{HP}=0{\cdot}9\ \theta_{01}$$
$$=0{\cdot}9\ \lambda/a \text{ rad.}$$
$$=51\ \lambda/a \text{ deg.}$$

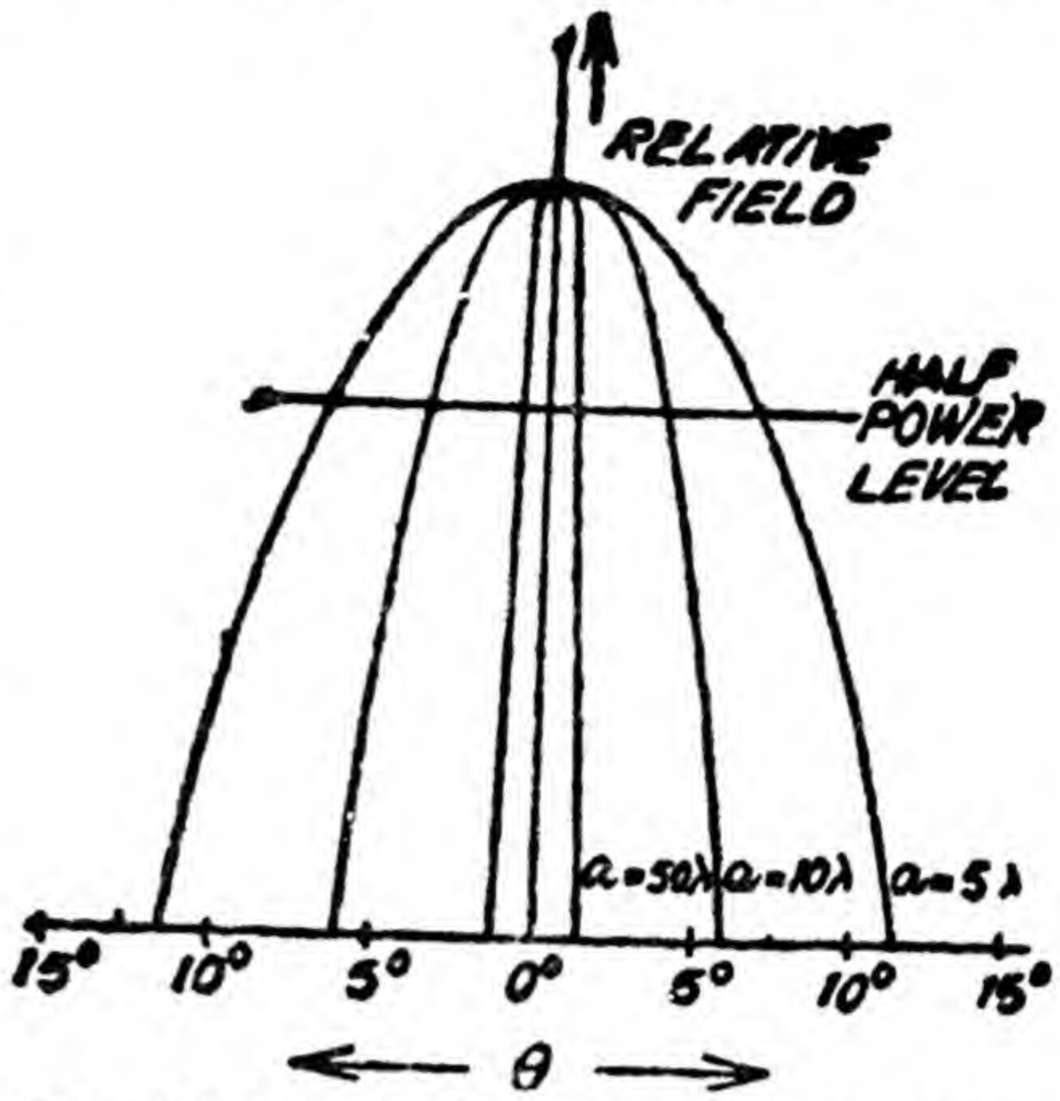


Fig. 29 (c) Main lobe field pattern of continuous uniform broad side array of length 5λ, 10λ and 50λ.

28·14. MULTIPLICATION OF PATTERNS :

In previous article we have discussed a direct method for the determination of radiation pattern of linear arrays. Multiplication of patterns provides another means for sketching such patterns. The chief advantage of the latter method lies in the fact that one can sketch rapidly, simply by inspection, the pattern of complicated arrays.

To illustrate this method, let us take up the following example:

Consider an array consisting of four elements, each spaced $\lambda/2$ apart and with currents in same phase. We shall treat antenna elements 1 and 2 as one unit, and 3 and 4 as other unit [fig 30(a)]. For the resultant pattern of one unit, antennas 1 and 2, or 3 and 4 can be replaced by a single antenna located at a point mid-way between them. The directional characteristic obtained is a 'figure of eight' [fig. 30 (b)].

Now the next step is to treat the two units together, *i.e.*, to determine the radiation pattern of two similar antennas that are spaced a wavelength apart and each of which has figure of eight directional pattern. The pattern of 2 non-directional radiators spaced λ apart and fed in phase is as shown in fig. 30 (c).

Thus the resultant pattern for the original four element array is obtained by multiplying the pattern (b) for a unit and pattern (c) for a group as shown in fig. 30 (d).

Thus if by simple calculations we could know the unit pattern and group pattern, it would be too simple to arrive atonce at the resultant approximate pattern by mere inspection.

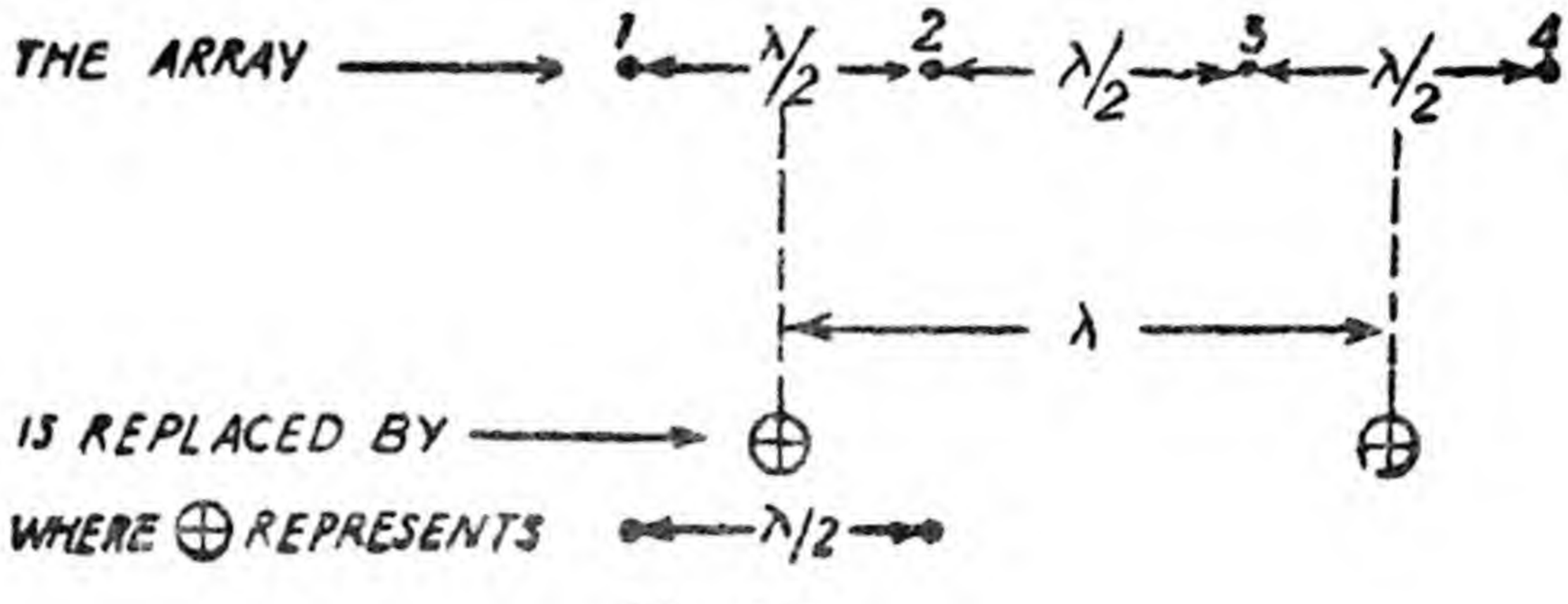


Fig. (*a*)

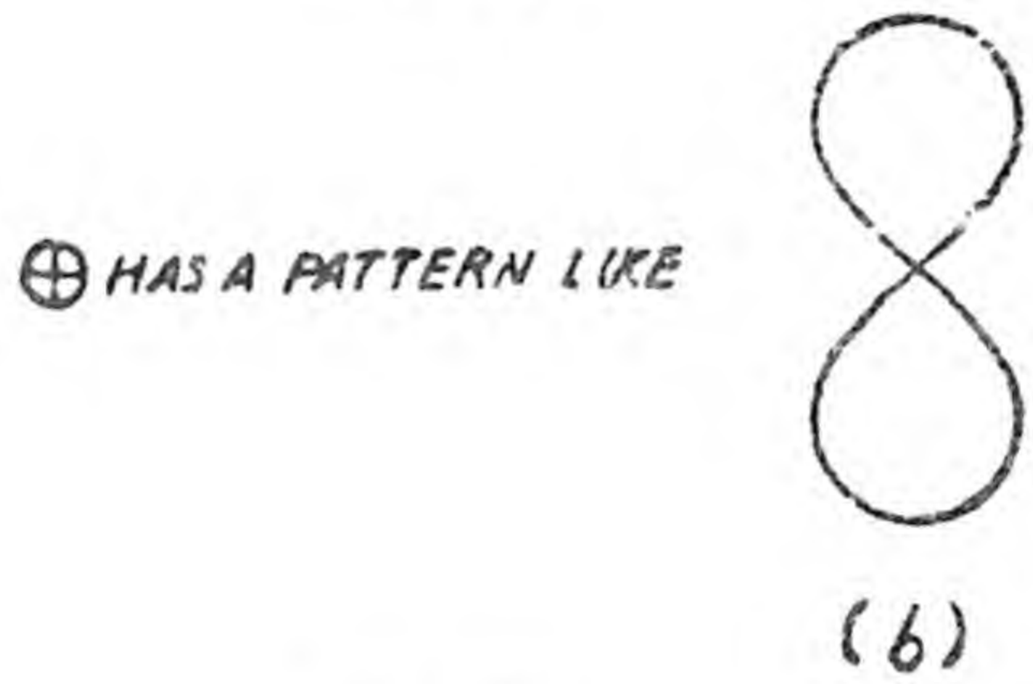


Fig. (*b*)

THE PATTERN OF 2 NON-DIRECTIONAL RADIATORS SPACED λ APART AND FED IN PHASE



THE RESULTANT PATTERN FOR ARRAY IS OBTAINED BY MULTIPLICATION AS SHOWN BELOW

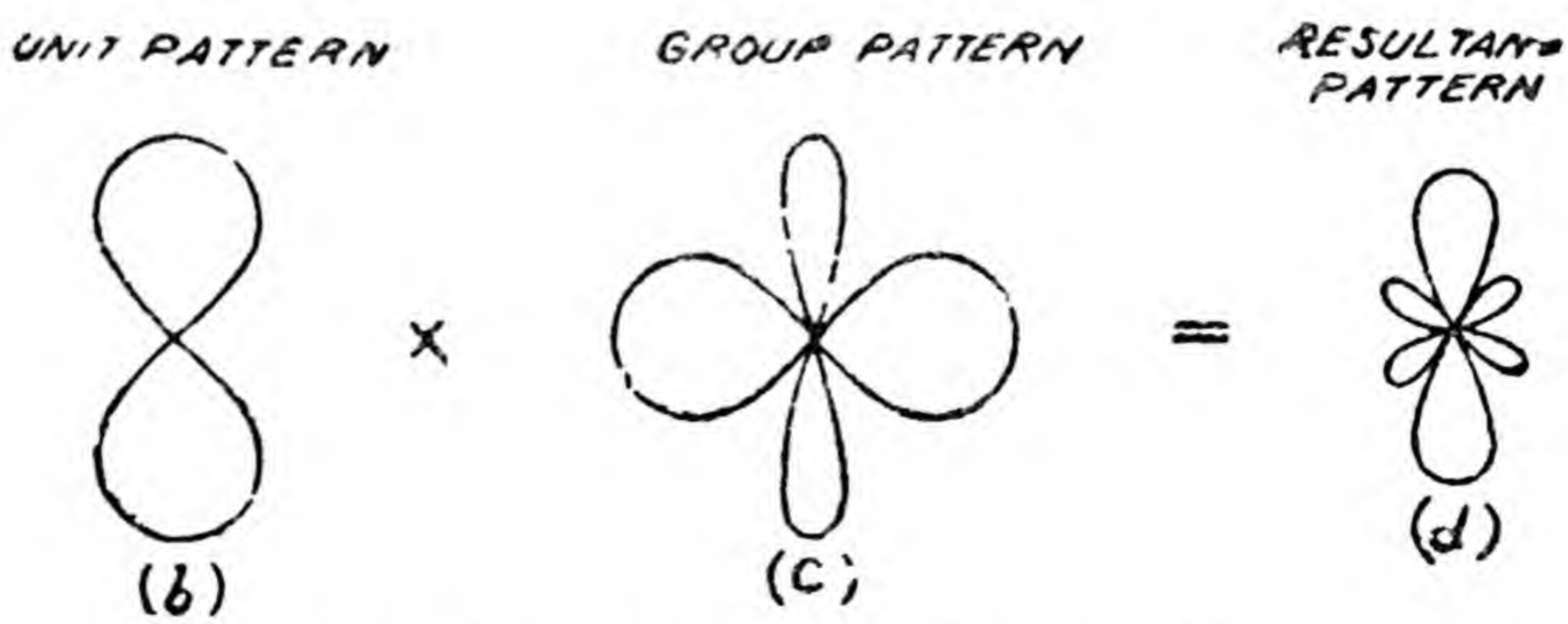


Fig. 30. Multiplication of patterns.

We note the following few points in the resultant pattern :

(*a*) the width of the principal lobe (between nulls) is the same as the width of the corresponding lobe of the group pattern :

(*b*) the sum of the nulls in the unit and group patterns, assuming none of the nulls are coincident, gives the number of nulls in the resultant pattern :

(*c*) the number of secondary lobes in the resultant pattern can be determined from the number of nulls in the resultant pattern.

28·15. ARRAY OF TWO DRIVEN $\lambda/2$ ELEMENTS :

The term element is taken to mean the basic unit antenna of which the array is constructed. We assume these elements to be linear and thin. Like to previous discussion of array of isotropic point sources, here too we shall study the two cases. One Broad side case in which the two elements are fed with equal inphase ($\delta = 0$) currents and the second end fire case in which the two elements are fed with equal currents in opposite phase ($\delta = \pm\pi$). In each case, expressions for field patterns and gain in field intensity will be derived.

We shall here stress on the calculation of form factor that is responsible for the directivity achievements of any type of array. The magnitude of the field can be estimated from the previous calculation of fields for the elemenes of the array.

Case I : BROAD SIDE CASE : (a) Field Patterns :

We assume that array is unaffected by ground or other objects, *i.e.*, array is in free space. We shall derive pattern expressions first for horizontal plane and then for vertical plane. Since angles θ and ϕ are confined to vertical and horizontal planes respectively. $E(\theta)$ and $E(\phi)$ will represent field intensities in vertical and horizontal planes. We consider a point at far distance from the array.

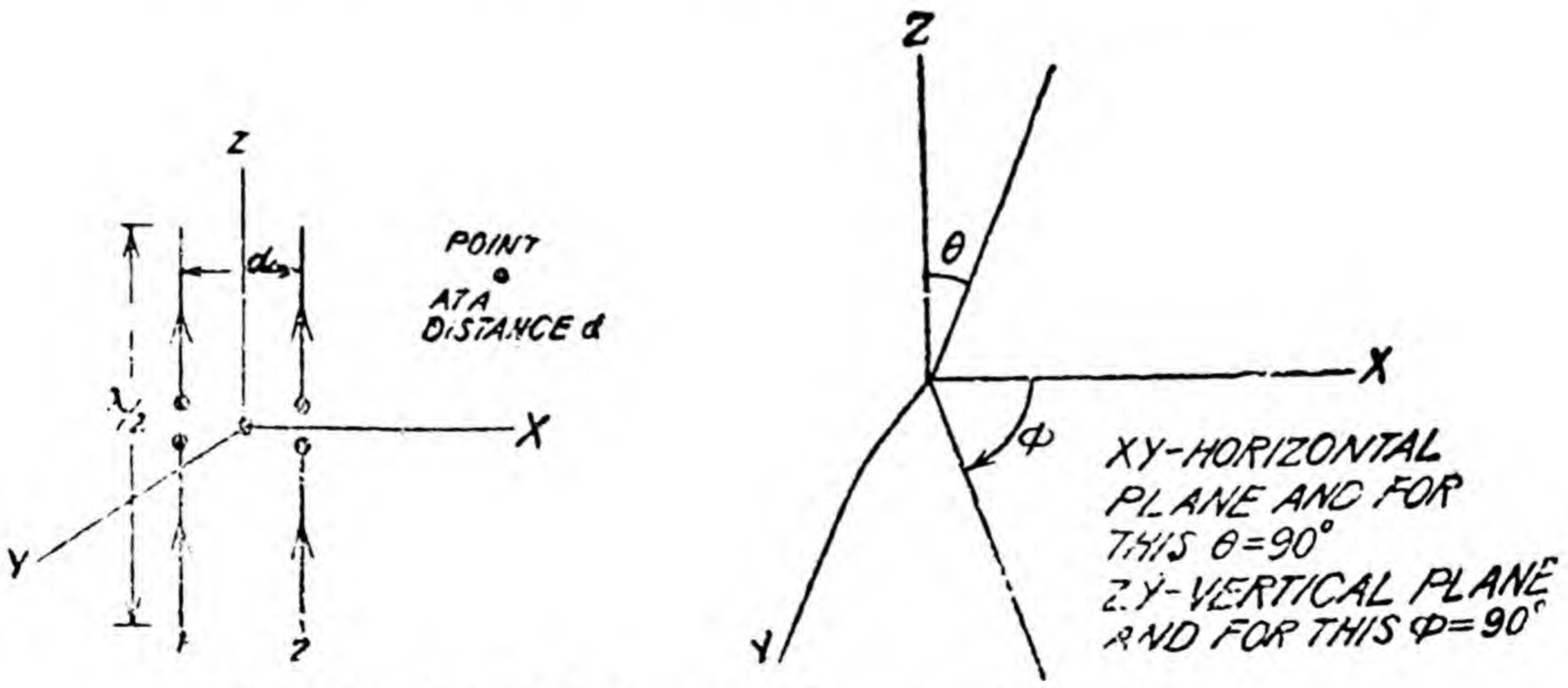


Fig. 31. Broad side array of two in phase $\lambda/2$ elements.

Since it is a broadside case, the maximum field will be directed in a direction normal to the line of array. Therefore as *X*-axis is

the line of array, maximum field will always be directed along Y axis in both the cases *i.e.*, whether we calculate field in horizontal XY plane or in vertical ZY plane.

Field intensity in horizontal plane :

At a large distance r,

$$(r \gg d),$$

the field intensity from a single element in the horizontal XY plane ($\theta=90°$) is given as

$$*E_1\ (\phi)=KI_1, \qquad \ldots (1)$$

where K is a constant involving the distance r and I_1 is the terminal current.

Since $E_1\ (\phi)$ is independent of ϕ, its pattern will be a circle in XY plane [fig. 32 (a)], which means element of $\lambda/2$ length radiates



(*a*) (*b*)

Fig. 32.

uniformly in horizontal plane or in other words, we can replace these elements by two isotropic point sources for the sake of intensity calculation in the horizontal plane. Now as is obvious from figure 32 (*b*), for two isotropic in phase point sources the resultant field intensity is

$$E_{iso}\ (\phi)=E_0\ e^{j\psi/2}+E_0\ e^{-j\psi/2}$$

$$=2E_0 \cos \psi/2, \qquad \ldots (1\ A)$$

where $\psi=\dfrac{2\pi d}{\lambda}\cos\phi=dr\cos\phi$, E_0 is the field intensity from a single element at a distance r (by the consideration of principle of pattern multiplication), so that

$$E_0=E_1\ (\phi)=KI_1.$$

---

*We note that for a current element

$$E=\frac{60\pi I\ dl\ \sin\theta}{\lambda r}=\frac{60\pi I\ dl}{\lambda r}\ (\text{for } \theta=90°)$$

$=KI$ at a fixed frequency of transmission.

Therefore resultant field intensity in the horizontal plane at a large distance by this array will be

$$E(\phi)=2E_1(\phi)\cos\psi/2$$
$$=2KI_1\cos\left(\frac{dr\cos\phi}{2}\right). \quad ...(2)$$

The pattern of $E(\phi)$ can be illustrated as shown in fig. 32 (*c*). The maximum field intensity is at $\phi=90°$ or broad side to the array, that is, along Y axis in XY plane.

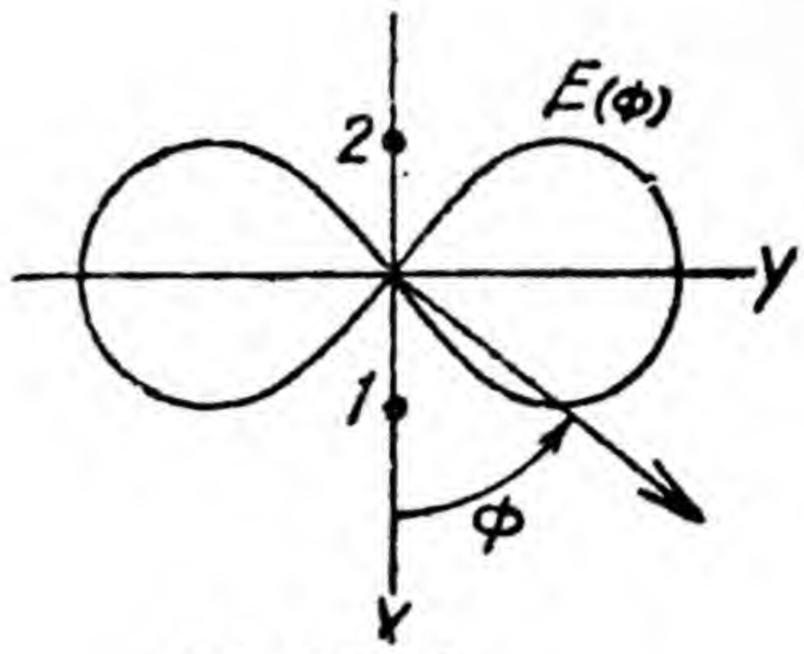


Fig 32 (*c*).

Field intensity in vertical plane :

When we consider the field in the vertical YZ plane, each element behaves as a half wave antenna and therefore from the theory of half wave antenna in space, we find that for a single element, the field intensity at a large distance r in the vertical place is given by

$$E_1(\theta)=KI_1\frac{\cos(\pi/2\cos\theta)}{\sin\theta}. \quad ...(3)$$

The pattern of two isotropic point sources in the vertical plane in place of two elements is

$$E_{iso}(\theta)=2E_0\cos\left(\frac{dr\cos\pi/2}{2}\right)^*$$
$$=2E_0$$
$$=2E_1(\theta), \quad ...(3A)$$

on applying the principle of pattern multiplication. Therefore, the resultant intensity in the vertical plane at a large distance r for the array of two elements will be

$$E(\theta)=2E_1(\theta)$$
$$=2.KI_1.\frac{\cos(\pi/2\cos\theta)}{\sin\theta}. \quad ..(4)$$

The space of the pattern is illustrated in fig. 32 (*d*). The field is maximum for $\theta=90°$ and $\phi=90°$ which is Y axis in ZY plane.

(*b*) **Gain in Field Intensity :**

Let us assume that R_{11}, R_{22} are the self resistances of element 1 and element 2 respectively and R_{12} is their mutual resistance. Let the total power input to the array be P. Then assuming no heat losses, power P_1 in element 1 is

$$\dagger P_1=I_1^2(R_{11}+R_{12}),$$

and the power P_2 in element 2 is

$$P_2=I_2^2(R_{22}+R_{12}).$$

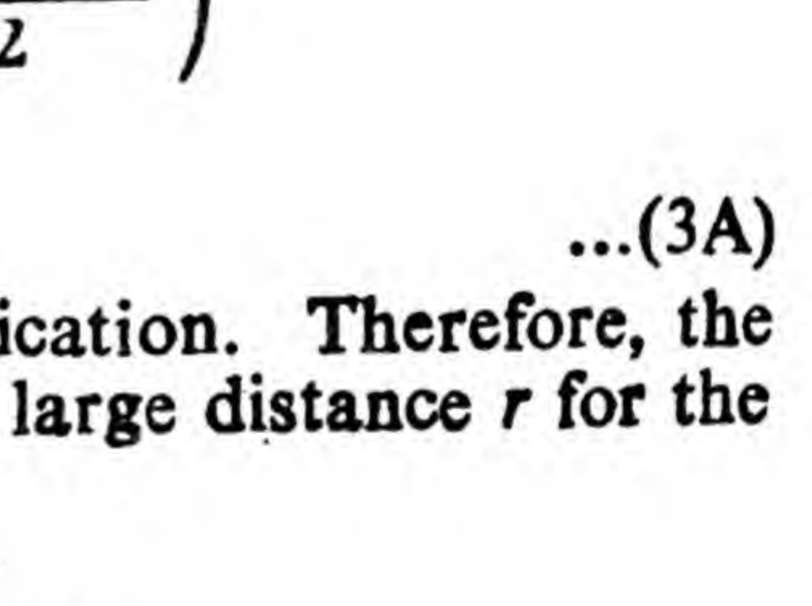


Fig. 32 (*d*)

*Because $\phi=\pi/2$ for ZY vertical plane.

$\dagger P_1=R_{11}I_1^2+I_1(R_{12}.I_2)$. Put $I_1=I_2$.

where I_1 and I_2 are r.m.s. currents. But $R_{22}=R_{11}$ and $I_2=I_1$, so that

$$P=P_1+P_2=2I_1^2\,(R_{11}+R_{12})$$

and thus $$I_1=\sqrt{\left\{\frac{P}{2\,(R_{11}+R_{12})}\right\}}.$$

Since we express gain of the array with reference to certain standard antenna, we take here a single $\lambda/2$ element as the reference antenna. Let the same power P be applied to this antenna. Then assuming no heat losses, the current I_0 at its terminals is

$$I_0=\sqrt{\left(\frac{P}{R_{00}}\right)},$$

where R_{00} is the self resistance of the reference antenna.

Now the gain in field intensity in the horizontal plane is defined as

$$G_f\,(\phi)=\frac{\text{Field intensity of the array in the horizontal plane (particular value of } \phi\text{) at a distance } r \text{ fed with power } P}{\text{Field intensity of the reference antenna in the horizontal plane (same } \phi\text{) at the same distance } r \text{ fed with the same power } P}$$

$$=\frac{\overline{E\,(\phi)}}{\overline{E_{H.W.}(\phi)}}=\frac{2KI_1\cos\left(\frac{dr\cos\phi}{2}\right)}{KI_0}$$

$$=\frac{2I_1}{I_0}\cos\left(\frac{dr\cos\phi}{2}\right)$$

$$=\sqrt{\left(\frac{2R_{00}}{R_{11}+R_{12}}\right)}\left|\cos\left(\frac{dr\cos\phi}{2}\right)\right|, \qquad \ldots(5)$$

on putting the values of I_1 and I_0. Bars indicate that gain will be confined to positive values or zero regardless of the values of dr and ϕ The maximum value of $G_f\,(\phi)$ is called directivity and is independent of ϕ, and can be referred as G_f simple.

For the case, when the spacing d is $\lambda/2$, and $R_{00}=R_{11}=73$ ohms ($\lambda/2$ antenna), $dr=\pi$ and $\dagger R_{12}=-13$ ohms, we find that

$$G_f\,(\phi)=1{\cdot}56\cos(\pi/2\cos\phi).$$

In the broadside direction $\phi=\pi/2$, the gain is

$$G_f\,(\phi)=1{\cdot}56.$$

We note that it is the maximum value of $G_f\,(\phi)$ and is thus G_f. The gain as a decibel ratio is given by

---

$$\dagger R_{12}=-\frac{V_{12}}{I_2} \text{ and } R_{21}=-\frac{V_{21}}{I_1}$$

where V_{21} is the e.m.f. that a current I_1 in antenna 1 induces at the open terminals of antenna 2.

$$*\text{Gain}=20\log_{10} G_f\ db$$
$$=3{\cdot}86\ db.$$

Now the gain in the vertical plane will be given by

$$G_f(\theta)=\frac{E(\theta)}{E_{H\cdot W.}(\theta)}=\frac{2KI_1\dfrac{\cos(\pi/2\cos\theta)}{\sin\theta}}{KI_0\dfrac{\cos(\pi/2\cos\theta)}{\sin\theta}}$$
$$=\sqrt{\left(\frac{2R_{00}}{R_{11}+R_{12}}\right)},\ \text{on putting the values of } I_1 \text{ and } I_0.$$

We note here that unlike to the gain in horizontal plane, gain in vertical plane is independent of θ. For case $d=\lambda/2$, and half wave antenna system, we have

$$G_f(\theta)=1{\cdot}56 \quad \text{or} \quad 3{\cdot}86\ db.$$

Case II. END FIRE CASE : (a) Field Patterns :

Array is shown in fig. 33 (*a*). We note that, unlike to case I, here the currents in the elements are in opposite phase instead of being in the same phase.

Since it is end fire case, the maximum field will always be directed along *X*-axis, the line of array, in both the cases of field calculation (in *XY* plane or *ZX* plane).

Field intensity in horizontal plane :

At a large distance r ($r \gg d$), the field intensity from a single element in the horizontal *XY* plane is given by

$$E_1(\phi)=KI_1,$$

where I_1 is the terminal current and K is constant involving the distance r.

Replacing the elements by isotropic point sources of equal amplitude, the field intensity in *XY* plane for two such out of phase source is given by

$$\dagger\dagger E_{iso}(\phi)=E_0e^{j\psi/2}-E_0e^{-j\psi/2}$$
$$=2E_0\sin\left(\frac{dr\cos\phi}{2}\right),\ \text{ignoring } j.$$

---

*Gain$=10\log_{10} G_f$, where G_f is power gain. When gain is expressed in field intensity, power P is proportional to square of field intensity and hence gain$=20\log_{10} G_f$.

††Since for end fire array $\psi=dr\cos\phi+\delta$, where $\delta=-\pi$, because currents are in opposite phase. Therefore

$$E_{iso}(\phi)=E_0e^{j\psi/2}+E_0e^{-j\psi/2}=2E_0\cos\left(\frac{dr\cos\phi-\pi}{2}\right)$$
$$=2E_0\sin\left(\frac{dr\cos\phi}{2}\right)$$

The account of phase $\delta=\pi$ has been taken by taking the difference of fields radiated by two elements.

By the principle of pattern multiplication, E_0 can be considered as the intensity from a single element at a large distance r. Therefore

$$E_0=E_1(\phi)=KI_1$$

and hence the field intensity in the horizontal plane at a large distance by this array will be

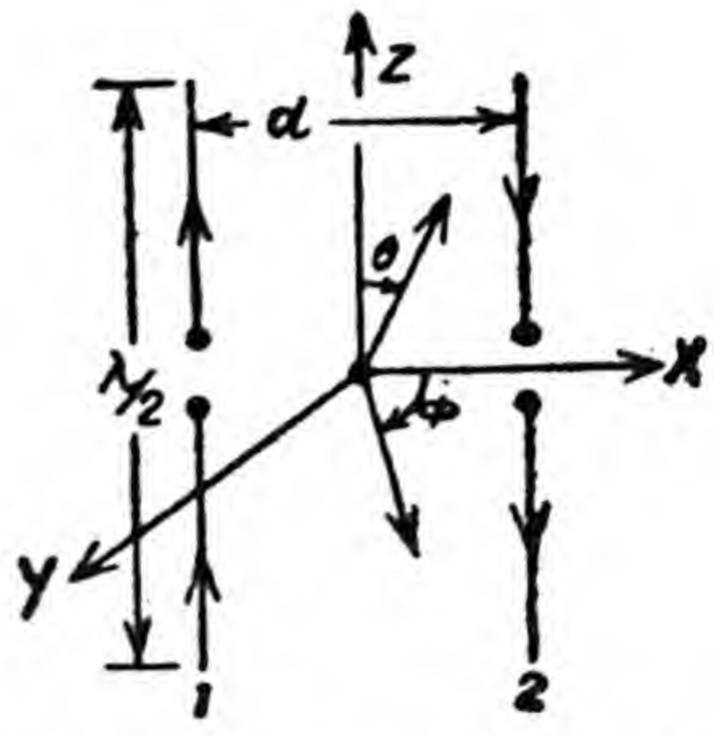


Fig. 33. (a) End fire array of two $\lambda/2$ elements.

$$E(\phi)=2KI_1 \sin\left(\frac{dr\cos\phi}{2}\right) \qquad ...(6)$$

The pattern of $E(\phi)$ can be plotted as shown in fig. 33 (*b*). The maximum field intensity at $\phi=0°$ and $\phi=180°$, *i.e.* along X axis in XY plane. Hence the array commonly referred to as an 'end fire' type.

Field Intensity in Vertical Plane :

The field intensity $E_1(\theta)$ from a single $\lambda/2$ element in the vertical plane at a large distance r is given by

$$E_1(\theta)=KI_1 \frac{\cos(\pi/2\cos\theta)}{\sin\theta}.$$

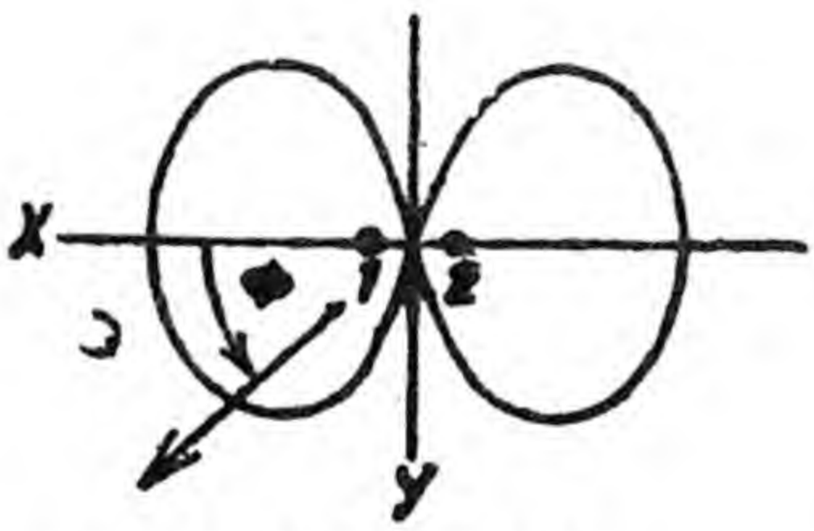


Fig. 33. (*b*) Pattern for $E(\phi)$.

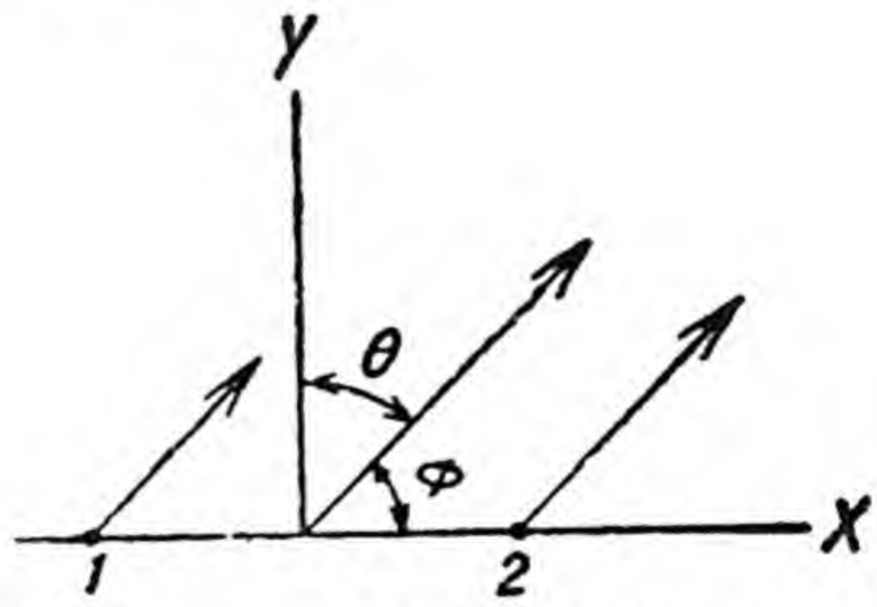


Fig. 33. (*c*) Array of two isotropic point sources.

The pattern of two isotropic point sources in the vertical plane in place of two elements

$$E_{iso}(\theta)=2E_0 \sin\left(\frac{dr\cos\phi}{2}\right).$$

But we note from fig. 33 (*c*) that in this case of two isotropic point† sources θ is complementary to ϕ and hence

$$E_{iso}(\theta)=2E_0 \sin\left(\frac{dr\sin\theta}{2}\right)^*.$$

---

†Since for isotropic radiator, vertical or horizontal planes are not distinguished as it radiates uniformly in all directions.

*Phase difference has again been accounted by taking sin $\psi/2$.

Putting $E_0=E_1(\theta)$ the field intensity in the vertical plane at a large distance r for the element of array is

$$E(\theta)=2KI_1\frac{\cos(\pi/2\cos\theta)}{\sin\theta}\sin\left(\frac{dr\sin\theta}{2}\right) \quad \ldots(7)$$

The maximum field intensity will be for $\phi=0$ or $\theta=90°$ *i.e.* along X-axis in ZX plane.

(b) Gain in Field Intensity :

Let us assume that R_{11}, R_{22} are the self resistances of element 1 and element 2 respectively and R_{12} is their mutual resistance. Let the total power input to the array be P. Then assuming no heat losses, power P_1 in element 1 is

$$P_1=I_1^2R_{11}+I_1(R_{12}I_2).$$

But $I_2=-I_1$, so that

$$P_1=I_1^2R_{11}-I_1^2R_{12},$$

since in this case, current between the two elements are in opposite phase.

Also, power P_2 in the element-2 is

$$P_2=I_2^2(R_{22}-R_{12})$$

The total power input is

$$P=P_1+P_2$$
$$=2I_1^2(R_{11}-R_{12}),$$

taking $R_1=R_{22}$ and $I_2=-I_1$. Thus

$$I_1=\sqrt{\left\{\frac{P}{2(R_{11}-R_{12})}\right\}}$$

Since we express gain of the array with reference to certain standard antenna, we take here a single $\lambda/2$ element as the reference antenna. Let the same power P be supplied to this antenna. Then assuming no heat losses the current I_0 at its terminals is

$$I_0=\sqrt{\left(\frac{P}{R_{00}}\right)}$$

where R_{00} is the self resistance of the reference antenna.

Now the gain in field intensity in the horizontal plane will be

$$G_f(\phi)=\frac{E(\phi)}{E_{H.W.}(\phi)}=\frac{2KI_1\sin\left(\frac{dr\cos\phi}{2}\right)}{KI_0}$$
$$=\sqrt{\left(\frac{2R_{00}}{R_{11}-R_{12}}\right)}\left|\sin\left(\frac{dr\cos\phi}{2}\right)\right| \ldots(8)$$

For half wave antenna array with a spacing of $\lambda/2$, the above expression assumes the form

$$G_f(\phi)=1{\cdot}3\mid\sin(\pi/2\cos\phi)\mid,$$

on putting $R_{00}=R_{11}=73$ ohms, $dr=\pi$, $R_{12}=-13$ ohms.

In end fire directions ($\phi=0°$ or or 180°), the pattern factor sin ($\pi/2$ cos ϕ) becomes unity and the gain is 1·3 or 2·3db. We call this maximum gain as directivity and is thus denoted as G_f, being independent of angle ϕ.

Now the gain in the vertical plane is given by

$$G_f(\phi)=\frac{E(\theta)}{E_{H\cdot W\cdot}(\theta)}=\frac{2KI_1\dfrac{\cos(\pi/2\cos\theta)}{\sin\theta}\sin\left(\dfrac{dr\sin\theta}{2}\right)}{KI_0\dfrac{\cos(\pi/2\cos\theta)}{\sin\theta}}$$

$$=\sqrt{\left(\frac{2R_{00}}{R_{11}-R_{12}}\right)}\left|\sin\left(\frac{dr\sin\theta}{2}\right)\right| \qquad ...(9)$$

The maximum radiation, as we find from equation (9) occurs in a direction $\theta=90°$, $\phi=0°$. Unlike to broad side case, here gain in the vertical plane is a function of angle θ. The directivity G_f will be $G_f=\sqrt{\left(\dfrac{2R_{00}}{R_{11}-R_{12}}\right)}\left|\sin\dfrac{dr}{2}\right|$ on putting $\theta=90°$.

For half wave elements array with a spacing of $\lambda/2$, $dr=\pi$. So that $G_f=\sqrt{\left(\dfrac{2R_{00}}{R_{11}-R_{12}}\right)}$

$=1·3$ or $2·3$ db.

Example : *An antenna array consists of four $\lambda/2$ elements arranged in a vertical plane ; each element being fed in phase from a single balanced transmission line. Draw the radiation pattern and calculate the directions of maxima if the spacing between the adjacent elements is $\lambda/2$.*

Refer to Case I Art. 28·15 :

Since all elements are fed in phase, $\delta=0$. This is a broad side case in which maximum of radiation is directed perpendicular to the line of array. Compared to Case I art. 28·15, here are four $\lambda/2$ elements instead of two.

Field intensity in horizontal plane : The discussion given in that article will remain as such for this case too except eq. (1A, will now become $E_{iso}(\phi)=E_0\,e^{j3\psi/2}+E_0\,e^{j\psi/2}+E_0\,e^{-j\psi/2}+E_0\,e^{-j3\psi/2}$

$$=2E_0\,[\cos 3\psi/2+\cos\psi/2]$$

$$=2E_0\left[\cos\left(\frac{3\,dr\cos\phi}{2}\right)+\cos\left(\frac{dr\cos\phi}{2}\right)\right]$$

Since $d=\dfrac{\lambda}{2}$, $dr=\dfrac{2\pi}{\lambda}\cdot\dfrac{\lambda}{2}=\pi$. Therefore eq. (2) will become

$$E(\phi)=2KI_1\left[\cos\left(\frac{3\pi}{2}\cos\phi\right)+\cos\left(\frac{\pi}{2}\cos\phi\right)\right]$$

$\phi=90°$, $E(\phi)=2(2KI_1)$ – maxima along *Y*-axis.

$\phi=0°, 180°$, $E(\ )=0$ – minima along *X*-axis.

Major lobe of the pattern will be same as shown in fig. 32 (*c*). The actual pattern is shown in fig. 30 (*d*).

Field intensity in vertical plane. The discussion will remain same as given in that article except eq. (3 A) will now become

$$E_{iso}(\theta)=2E_0\left[\cos\left(\frac{3\,dr\cos\frac{\pi}{2}}{2}\right)+\cos\left(\frac{dr\cos\frac{\pi}{2}}{2}\right)\right]$$

$$=2\,(2E_0)$$

so that eqn. (4) for the present case will be

$$E(\theta)=4\,KI_1\,\frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta}$$

Major lobe of pattern will be as shown in fig. 32 (*d*). The field is maximum for $\theta=90°$ and $\phi=90°$ which is *Y*-axis in *ZY* plane.

The readers may also opt for the method of multiplicication of patterns (Art. 28·14).

28·16. YAGI ANTENNA OR PARASITICALLY EXCITED END FIRE ARRAYS :

In previous article of 'array of two driven $\lambda/2$ elements', we have noted that both the elements were separately fed directly from transmission lines. Such directly fed elements are called 'driven' elements. It is, however, not necessary to feed each element of the array by direct connection to a transmission line. If only one dipole or element of the array is directly fed, the currents in the adjoining elements will flow due to the induction effect of the field set up by the driven element, or in other words the other elements of the array are excited due to induction effect of the field of the 'driven element'. Such a process of excitation is called parasitic excitation and the elements thus excited are '*parasitic elements*', End fire arrays employing this principle are known as Yagi antennas.

This principle of parasitic excitation can not be employed in broadside arrays. The reason can be understood as follows : for broadside arrays, currents in each element must be in the same phase ($\delta=0$) which requires that the distance between adjacent elements should be a full wavelength (because 'inphase' condition is possible only after the passage of a full wavelength) ; but with a full wavelength spacing between the elements, the pattern will

be of the shape shown in fig. 34 (*a*) which means there will be two end fire lobes of radiation as well as the two broad side lobes and this type of pattern is not a true broadside pattern. Due to this reason elements in broadside arrays are always driven rather than parasitic.

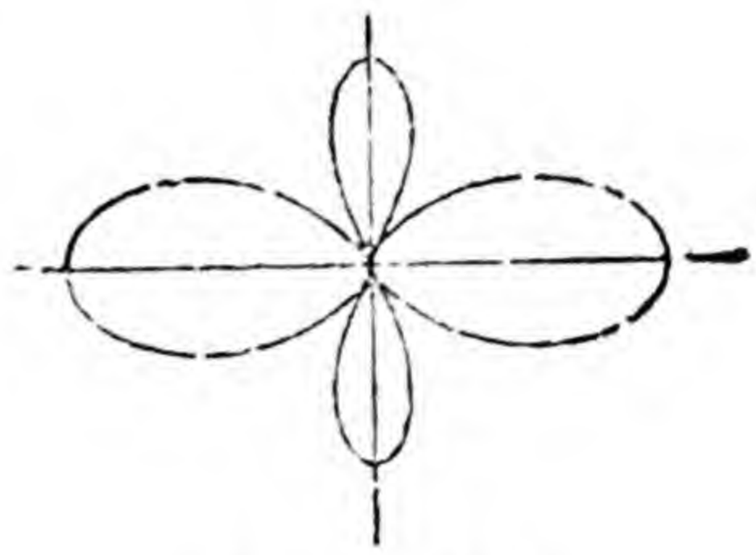


Fig. 34. (*a*)

Field strength at a point due to an array with parasitic elements :

We consider an array consisting one driven element $\lambda/2$ long (element 1) and one parasitic (element 2). We shall proceed to find gain in this case.

The circuit relations for the elements can be written as

$$V_1 = I_1 Z_{11} + I_2 Z_{12}$$
$$0 = I_2 Z_{22} + I_1 Z_{12},$$

where Z_{11}, Z_{22} are self impedances of elements 1, 2 and Z_{12} is the mutual impedance between them.

The current in element 2 will be

$$I_2 = -I_1 \frac{Z_{12}}{Z_{22}} = -I_1 \frac{|Z_{12}| \angle(\tau_m)}{|Z_{22}| \angle \tau_2}$$
$$= -I_1 \left| \frac{Z_{12}}{Z_{22}} \right| \angle \tau_m - \tau_2$$

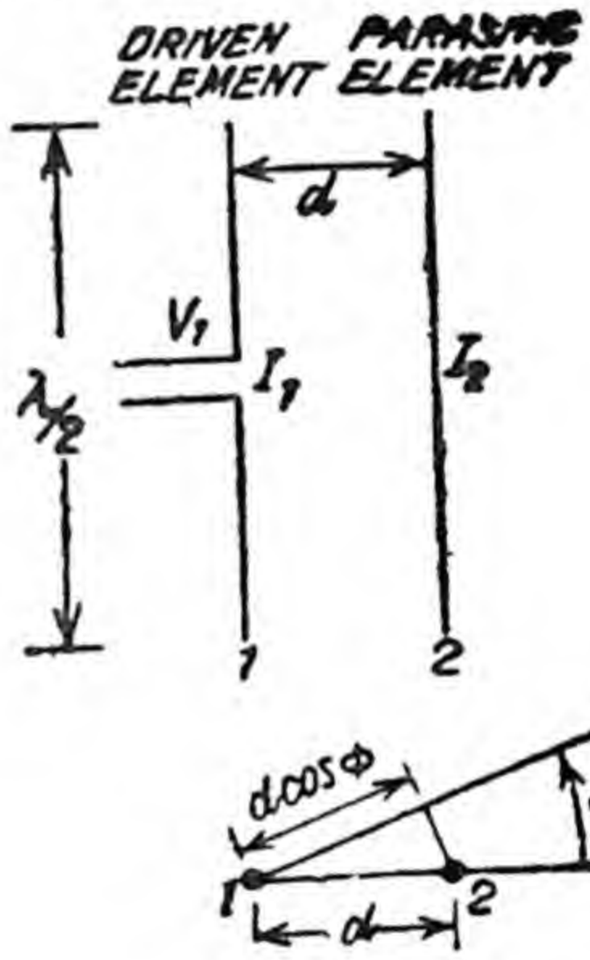


Fig. 34. (*b*)

or $$I_2 = I_1 \left| \frac{Z_{12}}{Z_{22}} \right| \angle \xi$$

where $\xi = \pi + \tau_m - \tau_2$ in which

$$\tau_m = \tan^{-1} \frac{X_{12}}{R_{12}}$$

$$\tau_2 = \tan^{-1} \frac{X_{22}}{R_{22}}$$

because $$Z_{12} = R_{12} + jX_{12}$$
$$Z_{22} = R_{22} + jX_{22}.$$

The electric field intensity at a large distance from the arrays as a function of ϕ is

$$E(\phi) = KI_1 + KI_2 \angle \left(\frac{2\pi}{\lambda} . d \cos \phi \right)$$
$$= K\{I_1 + I_2 \angle (dr \cos \phi)\},$$

where $dr \cos \phi$ is the path difference between the fields from etements 1 and 2 due to path difference $d \cos \phi$ between them.

Substituting for I_2, we get

$$E(\phi)=KI_1\left\{1+\left|\frac{Z_{11}}{Z_{22}}\right| \underline{/\ \xi+dr\cos\phi)}\right\} \quad \ldots(1)$$

The driving point impedance Z_1 of the driven element is

$$Z_1=\frac{V_1}{I_1}=\frac{I_1Z_{11}+I_2Z_{12}}{I_1}$$

$$=Z_{11}+Z_{12}\frac{I_2}{I_1}$$

But we have shown that

$$0=I_2Z_{22}+I_1Z_{12}$$

or
$$\frac{I_2}{I_1}=-\frac{Z_{12}}{Z_{22}}$$

so that driving point impedance becomes

$$Z_1=Z_{11}-\frac{Z_{12}^2}{Z_{22}}$$

$$=Z_{11}-\frac{|\ Z_{12}^2\ |\ \underline{/(2\tau_m)}}{|\ Z_{22}\ |\ \underline{/(\tau_2)}}$$

The real part of Z_1 is

$$R_1=R_{11}-\frac{|\ Z_{12}^2\ |}{|\ Z_{22}\ |}\cos(2\tau_m-\tau_2)$$

Adding for effective loss resistance, we can write

$$R_1=R_{1L}+R_{11}-\frac{|\ Z_{12}^2\ |}{|\ Z_{22}\ |}\cos(2\tau_m-\tau_2)$$

If power input to the driven element is P then

$$I_1=\sqrt{\left(\frac{P}{R_1}\right)}=\sqrt{\left(\frac{P}{R_{1L}+R_{11}-|\ Z_{12}^2/Z_{22}\ |\cos(2\tau_m-\tau_2)}\right)}$$

Putting this value of I_1 in equation (1), we get the expression of electric field at a far point due to array. That is

$$E(\phi)=K\left[\frac{P}{R_{1L}+R_{11}-|\ Z_{12}^2\ /\ Z_{22}\ |\cos(2\tau_m-\tau_2)}\right]^{1/2}\left[1+\frac{|Z_{12}|}{|Z_{22}|}\underline{/\ (\xi+dr\cos\phi)}\right]$$

Directivity : The directivity is defined as

$$G_f(\phi)=\frac{\text{Field intensity due to the array at a point fed with power } P}{\text{Field intensity due to reference antenna at the same point fed with same power } P}$$

We choose a single $\lambda/2$ element as the reference antenna. The field intensity at the same distance fed with same power P is

$$E_{H.W.}(\phi)=KI_0=K\sqrt{\left(\frac{P}{R_{00}+R_{0L}}\right)}$$

where R_{00} is self resistance of single $\lambda/2$ element.
R_{0L} is loss resistance of single $\lambda/2$ element.

Therefore

$$G_f(\phi)=\frac{E(\phi)}{E_{H.W.}(\phi)}$$

$$=\frac{K\left[\dfrac{P}{R_{11}+R_{1L}-|Z_{12}^2/Z_{22}|\cos(2\tau_m-\tau_2)}\right]^{1/2}\times\left[1+\left|\dfrac{Z_{12}}{Z_{22}}\right|/(\xi+dr\cos\phi)\right]}{K\left[\dfrac{P}{R_{00}+R_{0L}}\right]^{1/2}}$$

Since $R_{00}=R_{11}$ and letting $R_{0L}=R_{1L}$, we have

$$G_f(\phi)=\sqrt{\left(\frac{R_{11}+R_{1L}}{R_{11}+R_{1L}-|Z_{12}^2/Z_{22}|\cos(2\tau_m-\tau_2)}\right)}\times\left(1+\left|\frac{Z_{12}}{Z_{22}}\right|/\xi+dr\cos\phi)\right)$$

If by detuning X_{22} is made large then Z_{22} will become large so we neglect the terms containing Z_{22} in the denominator, giving

$$G_f(\phi)=\sqrt{\left(\frac{R_{11}+R_{1L}}{R_{11}+R_{1L}}\right)}=1.$$

That is, field of the array becomes equal to the field of single $\lambda/2$ antenna.

Reflectors and Directors : The magnitude of the current in the parasitic element and its phase relation to the current in the driven element depends on its tuning which can be accomplished in two ways :

(*i*) by having fixed length of element but by inserting lumped reactance in series with antenna.

(*ii*) by adjusting the length of the element.

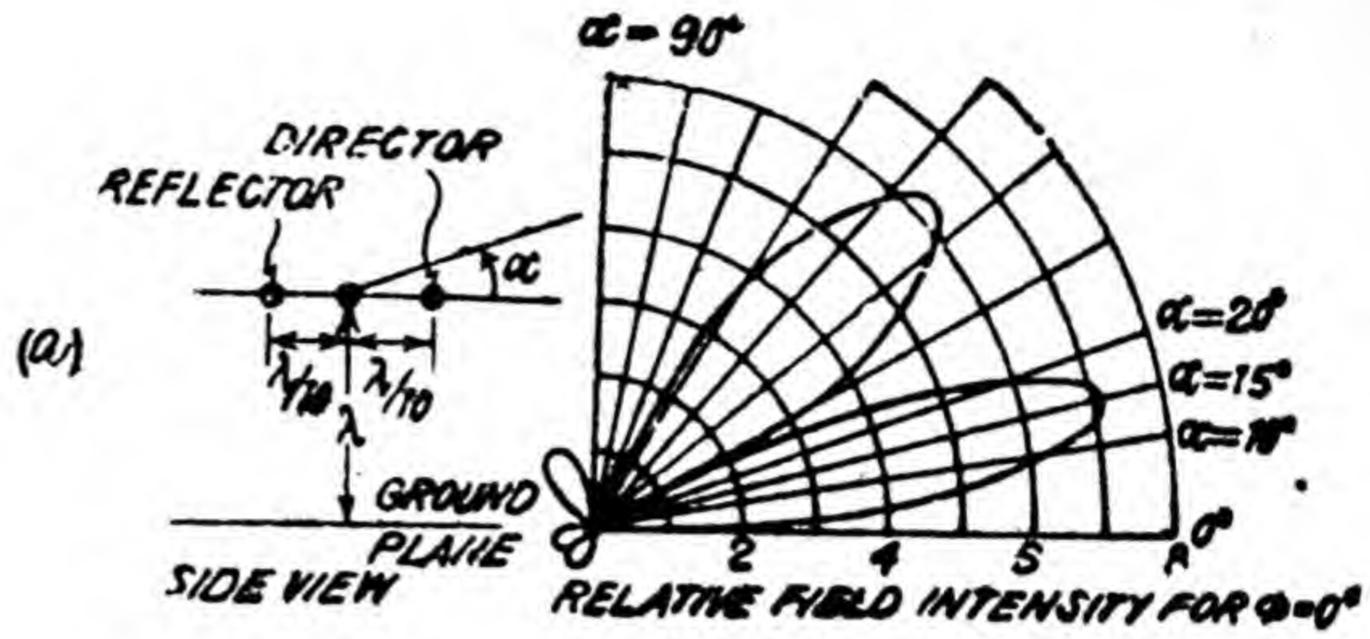
By changing the tuning of the parasitic element, it can act as a reflector sending the maximum radiation in $\phi=180°$ direction or as a director sending the maximum radiation in $\phi=0°$ direction.

Field Patterns :

Field patterns for a three element array located one wavelength above a large ground plane are shown in fig. 35. In fig. 35 (*a*) vertical plane pattern is shown in which we note that radiation appears at negative elevation angles due to the finite size of the ground plane. In fig. 35 (*b*) horizontal field patterns are shown at elevation angles $\beta=10°$, 15° and 20°.

Array may be constructed with a large number of parasitic elements. Yagi array is an example of such an array consisting a number of parasitic director elements arranged in a row in front of the driven element. It uses one element as reflector also.

Vertical plane pattern at three elevation angles for a three element array located λ above a large ground plane.



Horizontal plane pattern for the array for three elevation angles.

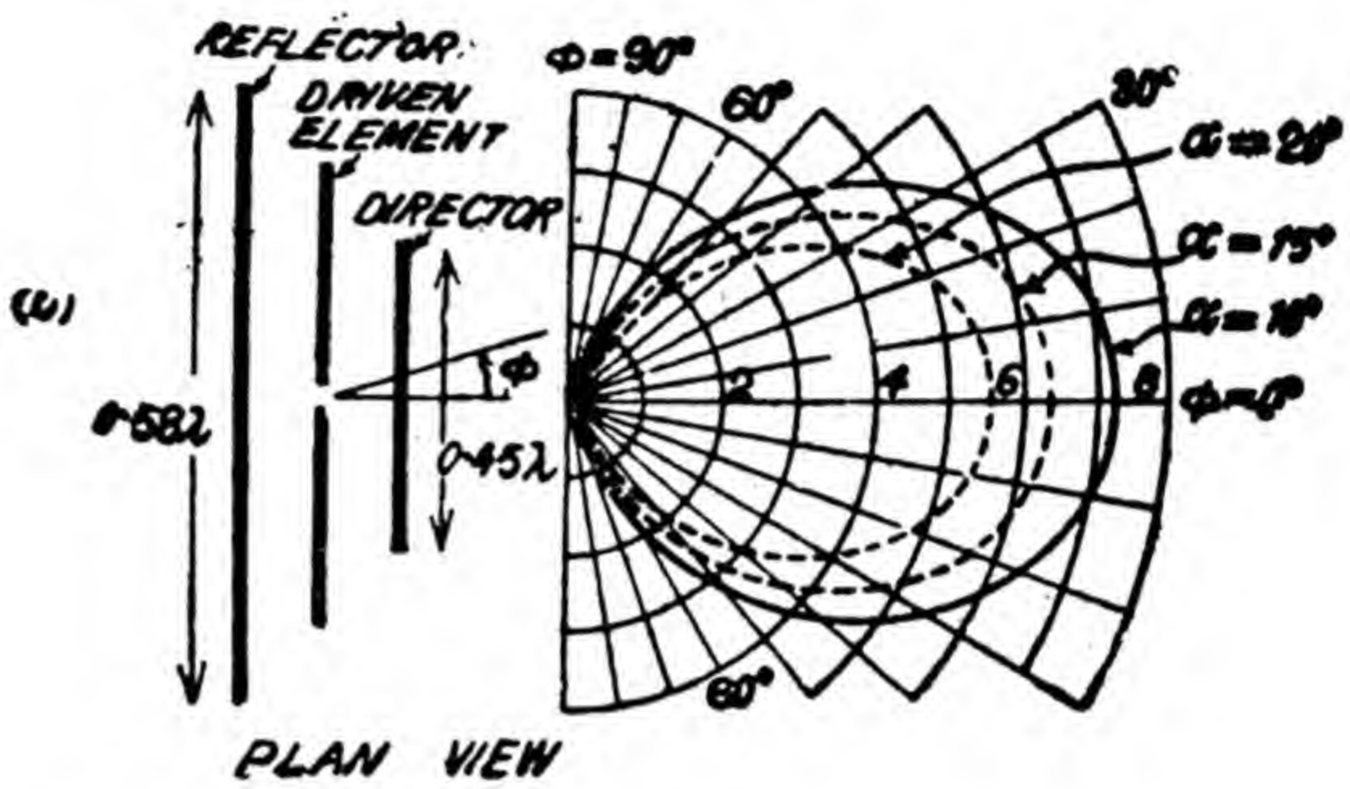


Fig. 35. Patterns of an array consisting three elements located λ distance above a large ground plane.

Yagi Antennas : For the field calculation at a certain distant point developed by such parasitic end fire array, the phases of currents in parasitic dipole elements are of utmost concern. Phases of currents in parasitic elements depend upon the following two factors :

(i) *Length of the parasitic elements* :

A parasitic dipole cut exactly a half wavelength or slightly longer will be inductive* and thus the phase of its current will lag the induced e.m.f., while a parasitic dipole cut shorter than a half wavelength will be capacitive and thus the phase of the current in it will lead the induced e.m.f.

(ii) *Spacing from the adjacent element* :

Close spacing between the elements is used in parasitic arrays to obtain good excitation and in the process of excitation, induction field of the elements play the major role.

The exact analysis of such arrays is complicated. It is known, however, that when the elements are shorter than $\lambda/2$ and are properly spaced, they reinforce the field of the driven element in a direction away from the driven element. Such elements are called as *director*. Thus if a director is placed infront of a driven element, it will direct the field in a direction away from the driven

*A half wave antenna can be approximated with quarter wave open circuited transmission line. Refer to impedance curves of such a line.

element [to right, in fig. 36 (*a*)]. Thus if there is a line of directors, each one will excite the next one.

On the other hand, when the elements are slightly longer than $\lambda/2$ or exactly $\lambda/2$, and are correctly spaced, they reinforce the field of the driven element in a direction towards the driven element or more precisely they reflect the field towards the driven element. Such elements are called *reflectors*. Thus if a reflector is placed adjacent to the driven element, any element placed beyond the reflector will receive negligible field and will not be appreciably excited [to the left of reflector in fig. 36 (*b*)].

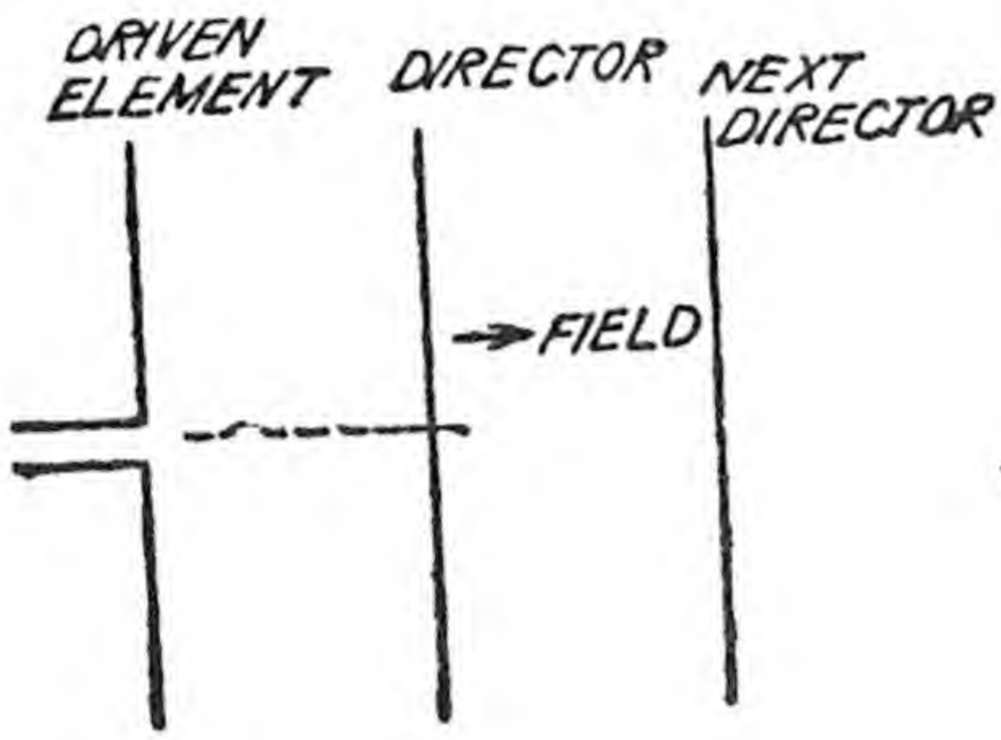


Fig. 36. (*a*)

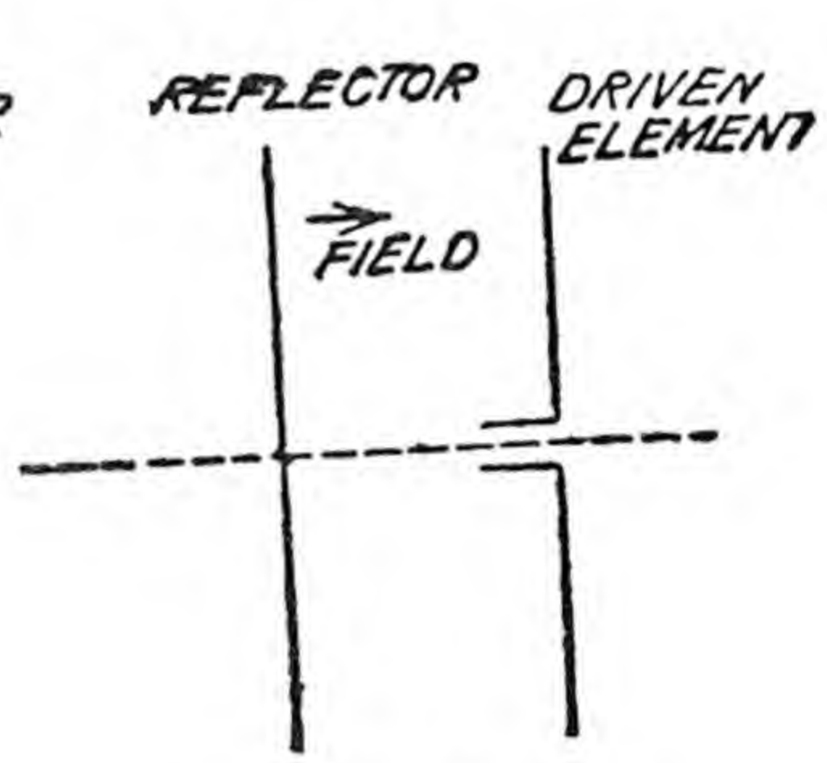


Fig. 36. (*b*)

A Yagi end fire array usually consists of one driven element, one reflector on one side of it, and a number of directors on the other side of it as shown in fig. 36(*c*). Directive gain of around 10 *db* is readily achieved with a moderate number of elements (five or six). A Yagi array has the following advantages :

(*a*) *they produce a uniform directional beam of moderate directivity, and*

(*b*) *they employ an easy way of feeding the system because only one element is fed, i.e., simplicity of feed system design and low cost.*

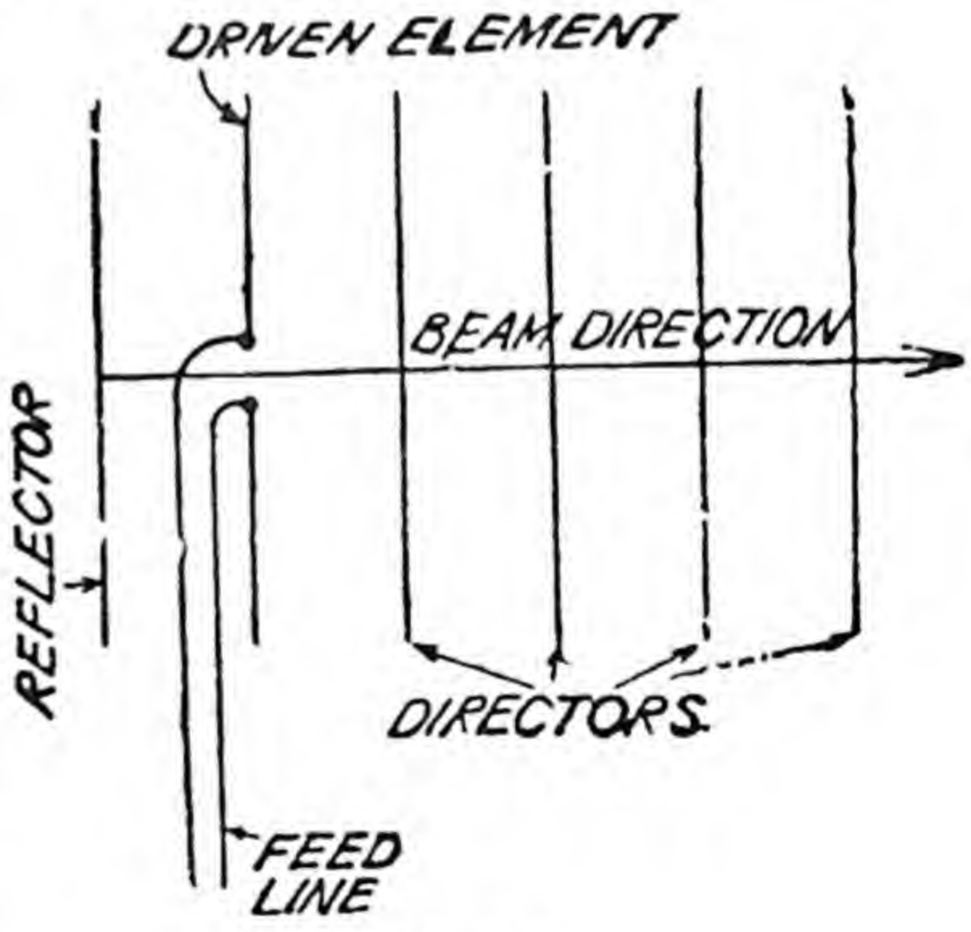


Fig. 36. (*c*)

The input impedance of Yagi array tends to be low and the beamwidth is limited to around 2% typically. To obtain higher gains, broadside array, consisting many Yagi arrays as constituent element is used.

Dimensions of reflectors, directors and folded dipole units spacing for Delhi TV channel 4 and 7 are shown in fig. 36 (*d*) and 36 (*e*), respectively.

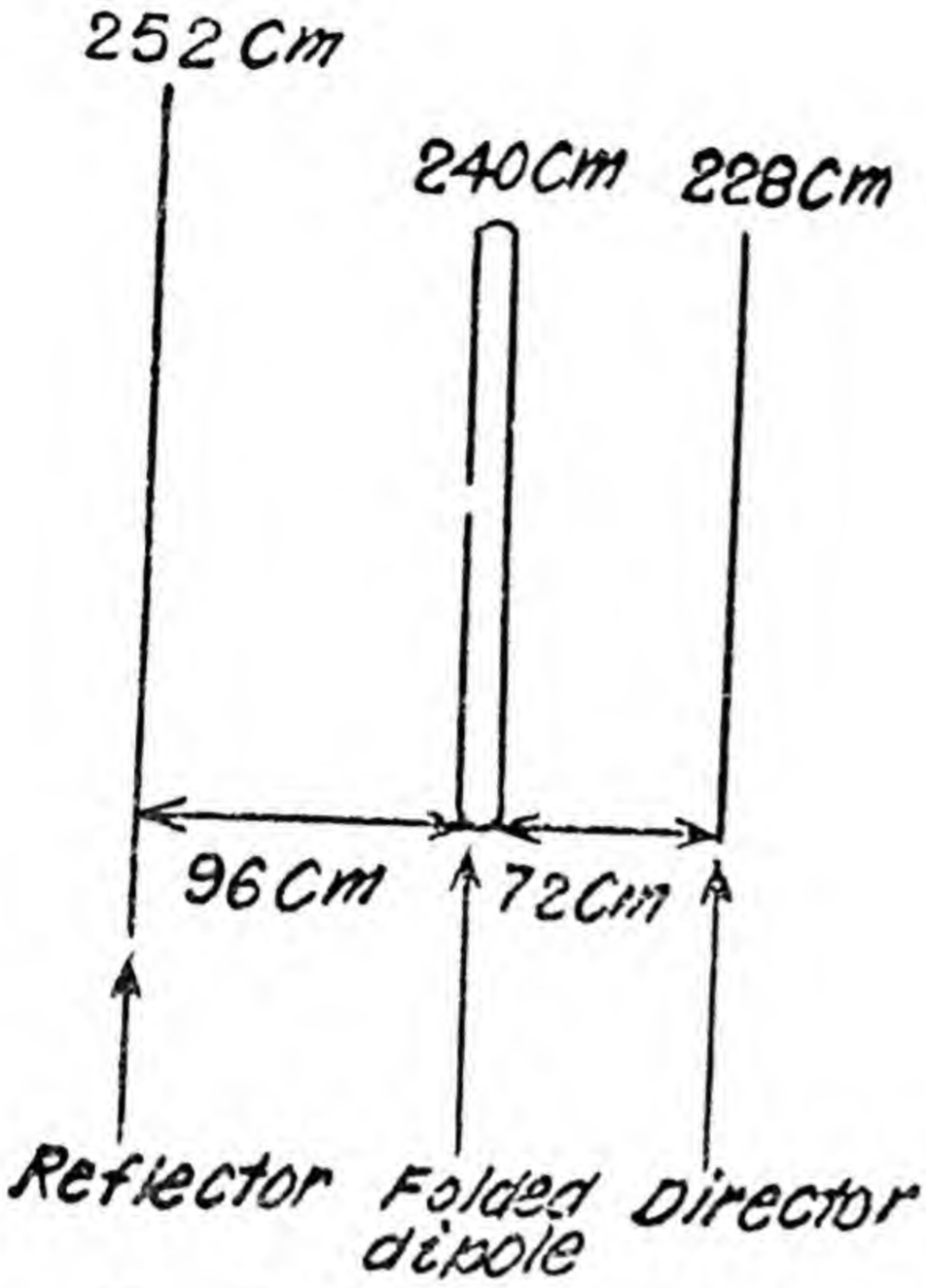


Fig. 36. (*d*) Channel-4 aerial

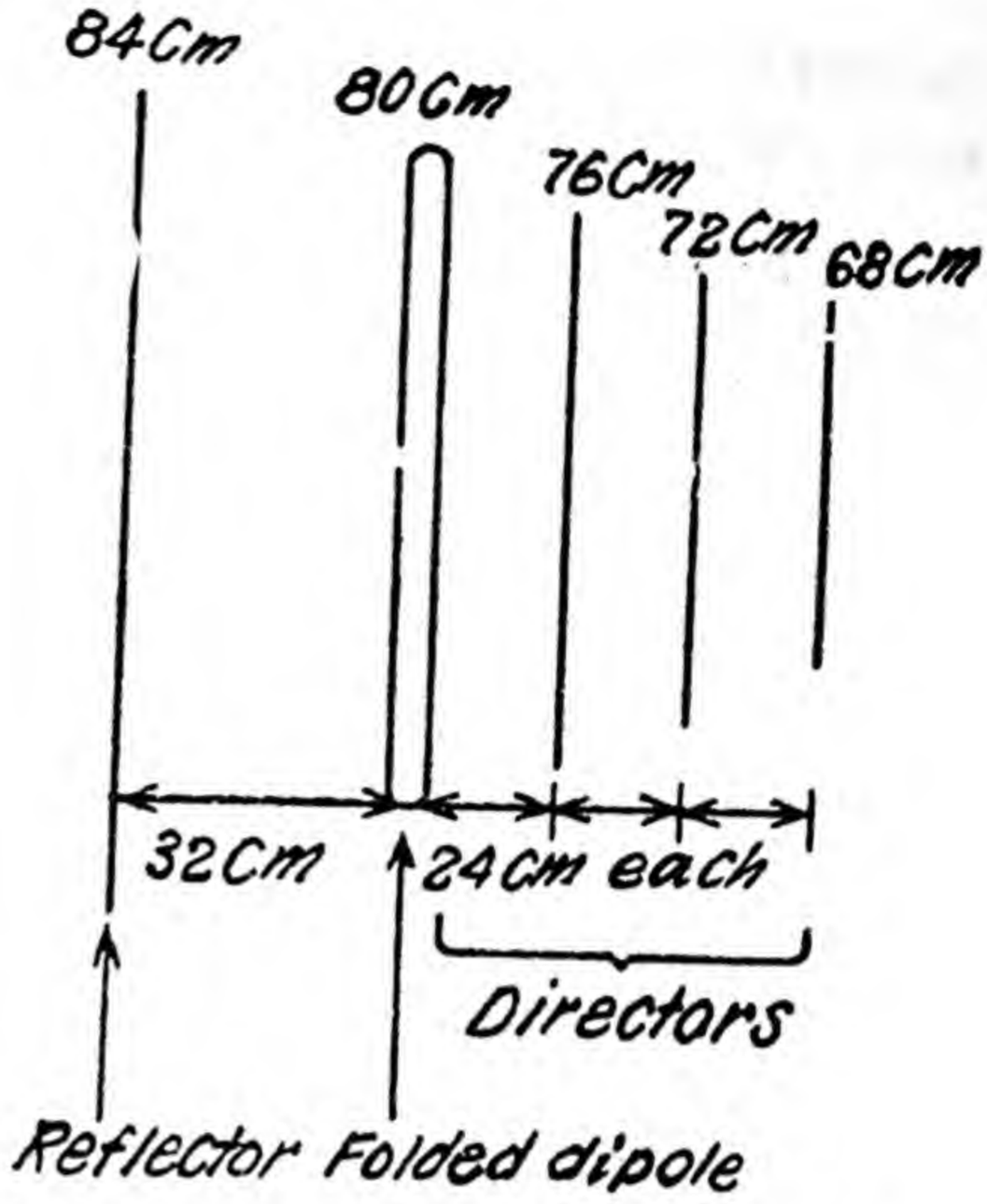


Fig 36. (*e*) Channel-7 aerial.

28·17. LOOP ANTENNA :

The loop antenna is a coil of any convenient cross-section. The radiation efficiency of closed loops is insufficient as regards the transmission purposes, unless their dimensions are made comparable to the wavelength employed. The ordinary loops are so designed that their dimensions are small compared with wavelength. This causes the currents to be of the same magnitude and phase throughout the loop. However, it must be pointed out here that if dimensions of loop are small compared to wavelength employed, the distributed capacity and leakage inductance between the turns of the loop may cause the current to vary in magnitude, phase being substantially the same. Due to this effect, the directional pattern of such loop antenna may differ from that of an idealized loop.

Closed loops are generally used for reception purposes. Now we shall calculate the voltage induced in a loop by a wave polarised* in the plane of a loop.

Consider that a rectangular loop remote from ground, of wire of height h, width W, is acted upon by electromagnetic wave of E volts/metre (fig. 37). Let the field strength E at any instant be given by

$$E = E_m \sin \omega (t - r/c),$$

where every notation bears its usual specification.

The e.m.f. induced in the loop will be the line integral of field strength E.

---

*It means electric vector E is in a plane parallel to arms 1 and 2. Since E is perpendicular to width W, the e.m f will be induced in arms 1 and 2.

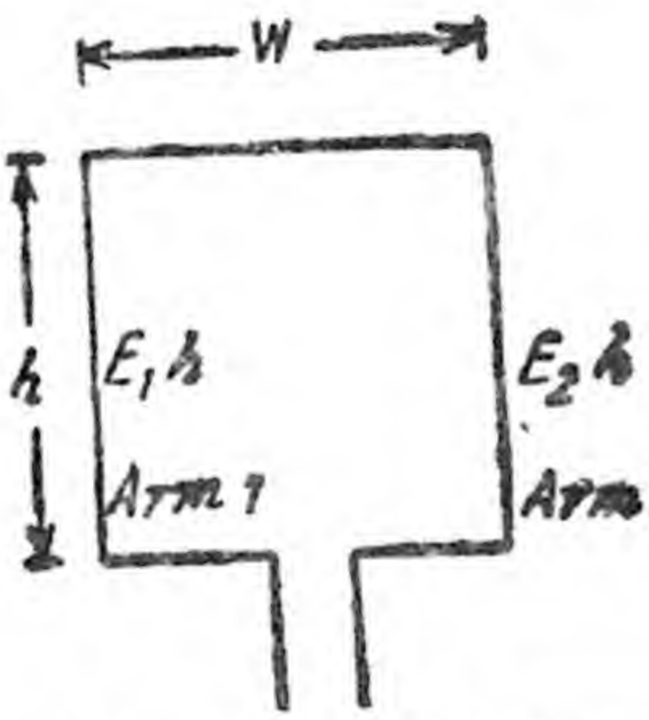
The e.m.f. induced in one arm will be in the same direction as in other arm, but differ in phase, the magnitude of which depends upon W.

The field strength in arm 1 will be

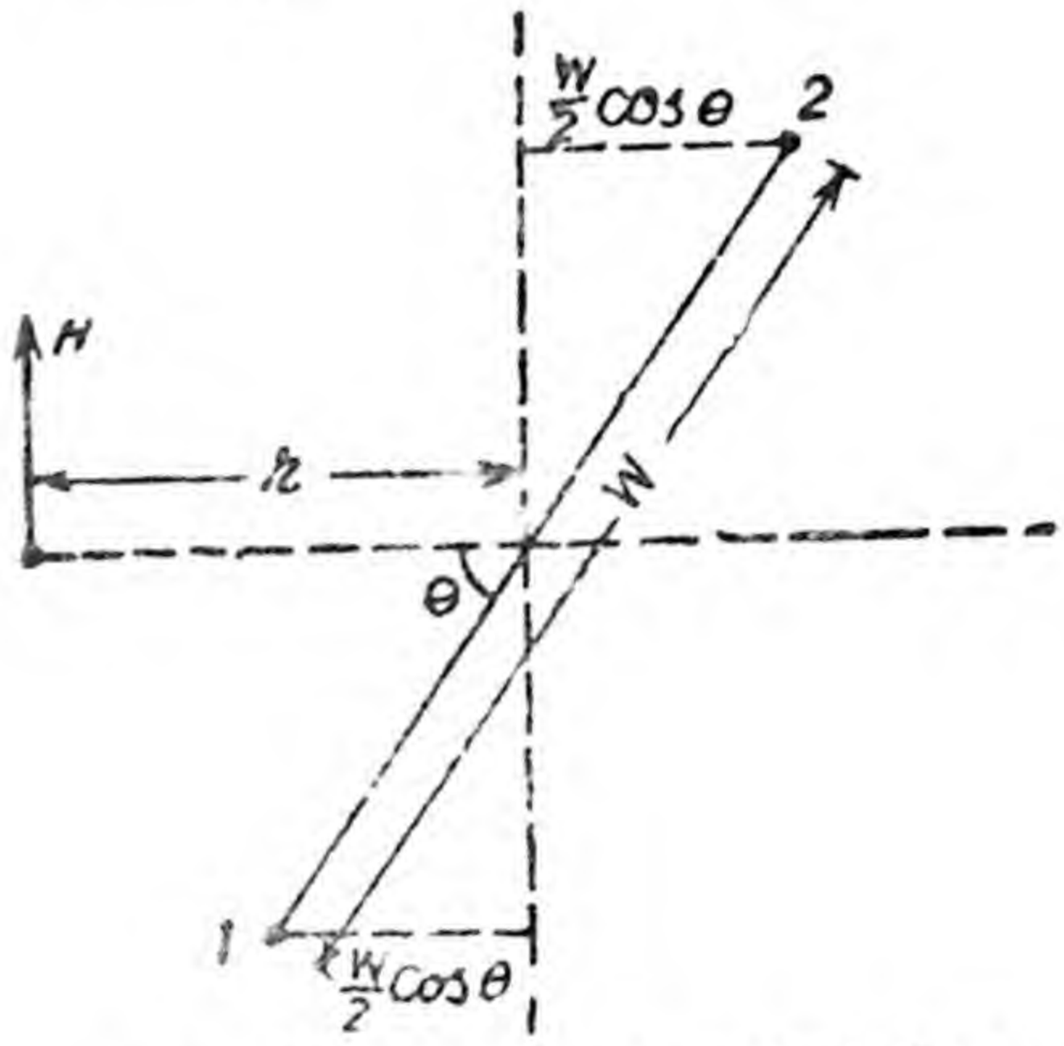
$$E_1 = E_m \sin \omega \left(t - \frac{r - \frac{1}{2} W \cos \theta}{c} \right)$$

and in arm 2,

$$E_2 = E_m \sin \omega \left(t - \frac{r + \frac{1}{2} W \cos \theta}{c} \right)$$



(*a*) Rectangular loop antenna.



(*b*) Induced e.m.f. calculations.

Fig. 37. Rectangular loop antenna.

The resultant e.m.f. will be

$$* = E_1 h - E_2 h$$

$$= E_m . h \left[\sin \omega \left(t - \frac{r - \frac{1}{2} W \cos \theta}{c} \right) - \sin \omega \left(t - \frac{r + \frac{1}{2} W \cos \theta}{c} \right) \right]$$

$$= 2 E_m . h \left[\cos \omega \left(t - \frac{r}{c} \right) \sin \omega \left(\frac{W \cos \theta}{2c} \right) \right].$$

If there are N turns in the loop, the induced e.m.f. will be

$$e = 2 N E_m h . \sin \omega \left(\frac{W \cos \theta}{2c} \right) . \cos \omega \left(t - \frac{r}{c} \right)$$

Since the dimensions are small compared with the wavelength, we can assume

---

*Quite in analogy w th

$$E = - \frac{dV}{dx}.$$

$$\sin \omega \left(\frac{W \cos \theta}{2c}\right) = \omega . \frac{W \cos \theta}{2c}$$

Therefore, $e = 2NE_m.h.\omega . \frac{W \cos \theta}{2c} . \cos \omega \left(t - \frac{r}{c}\right)$

$$= N.Wh.E_m . \frac{\omega}{c} \cos \theta . \cos \omega \left(t - \frac{r}{c}\right)$$

Also $Wh = A$, area of the loop, and

$$\frac{\omega}{c} = \frac{2\pi}{\lambda},$$

Therefore $$e = \frac{2\pi NA \cos \theta}{\lambda} E_m \cos \omega \left(t - \frac{r}{c}\right)$$

or $$| e | = \frac{2\pi NA \cos \theta}{\lambda} E_m \text{ volts,} \qquad \text{...(1)}$$

where E_m is in volts/metre, A the area of the loop in square metres, and θ is the angle between the plane of the loop and the direction of the transmitting station.

The maximum value of $| e |$ will occur when $\theta = 0, \pi, 2\pi, \ldots,$ *i.e. when the plane of the coil is along the direction of transmitting signal.* If the plane of coil is at right angles to the direction of transmitting signal, $\theta = 90°$ and hence induced e.m.f. is zero.

This latter property leads to its use in direction finding since by connecting a loop to the receiver and then rotating it, a position is obtained at which received signal is zero. A direction perpendicular to it will be the direction of transmitting station. If at two receiving stations, separated by a known distance, these measurements are taken, the location of transmitter can be found easily by triangulation.

Since the voltage induced is the multiplication of the field strength with height of the loop, the factor $\frac{2\pi NA \cos \theta}{\lambda}$, being a multiple of field strength E_m in the expression for $| e |$, is termed as effective height of the loop.

If the loop antenna is used for the transmission purposes, the field radiated by an ordinary loop in a plane perpendicular to its own, assuming the current to be constant throughout the loop, can be given by the formula, provided loop is small compared with a wavelength :

$$E = \frac{120\pi^2 N}{r} \left(\frac{A}{\lambda^2}\right) I_s \cos \theta, \qquad \text{.. (2)}$$

where r is distance in metres, and

I_s is loop current in amp,

rest notations being the same.

To derive equation (2), ***we proceed as follows*** :

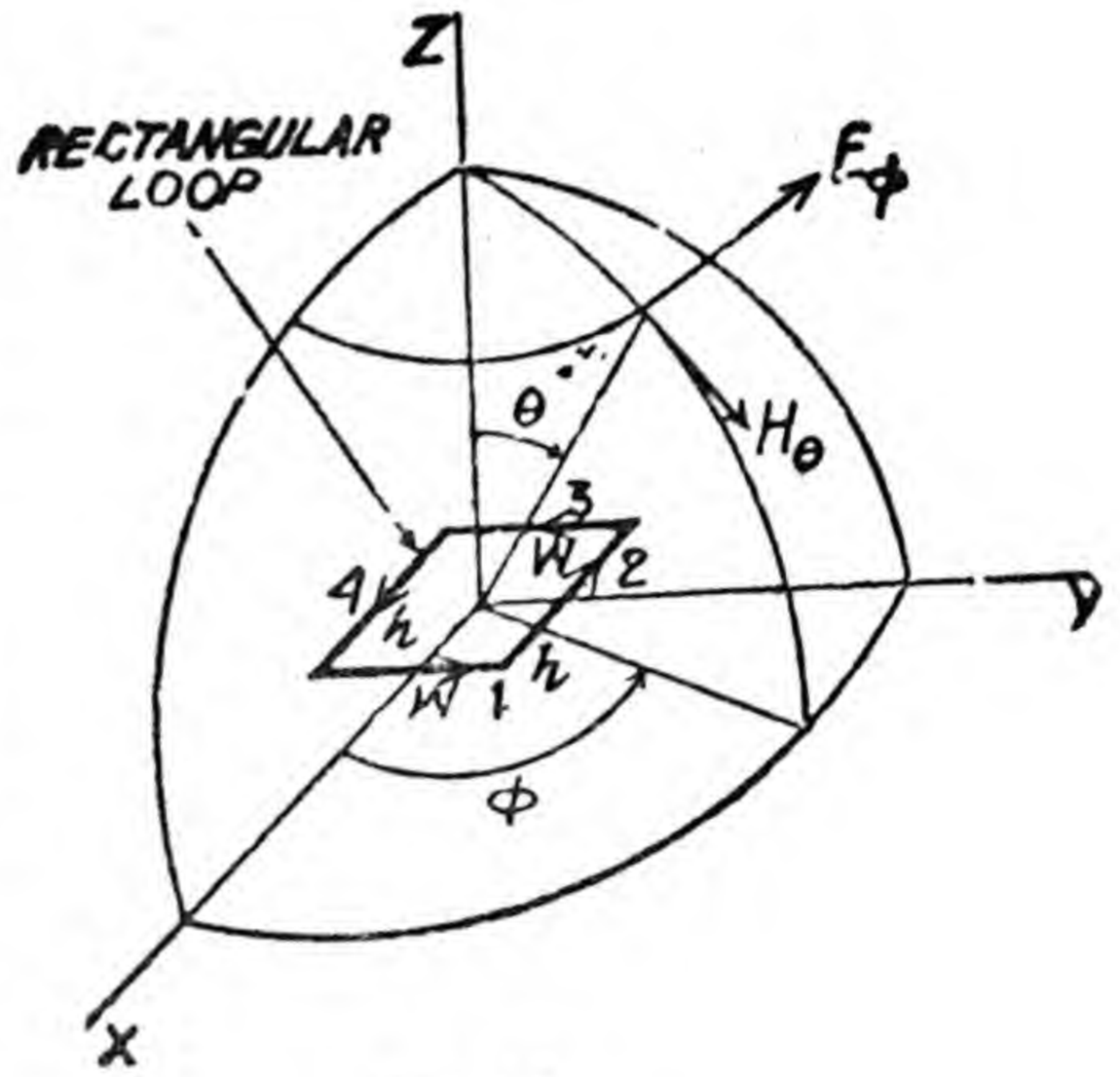


Fig. 38

Let us consider that the loop is placed in the coordinate system as shown in the fig. 38. If the loop is oriented as in fig. 38, its radiation field will have only E_ϕ component. If the dimensions of the loop are small compared to wavelength, the current throughout the whole loop can be assumed to be uniform in phase. We can now treat the loops as four short linear dipoles (1, 2, 3 and 4).

To find the radiation pattern in the *yz* plane, we should consider only 2 and 4 dipoles, *i.e.* two out of four small linear dipoles. Since 2 and 4 arms of the loop are being treated as short linear dipoles, their radiation pattern in *xy* plane, in which they are situated, will be of the form shown in fig. 39 (*a*), while a plane perpendicular to *xy*-plane, *i.e.*, in *yz* plane pattern is circular as shown in fig. 39 (*b*).



Fig. 39

Since the field pattern of 2, 4 dipoles is circular in *yz*-plane, they radiate uniformly in all directions, *i.e.* individual small dipoles 2 and 4 are non-directional in the *yz*-plane and hence the radiation pattern of this loop in *yz* plane will be the same as that due to two isotropic point sources. A cross-section through the loop in the *yz*-plane is shown in fig. 39 (*c*).

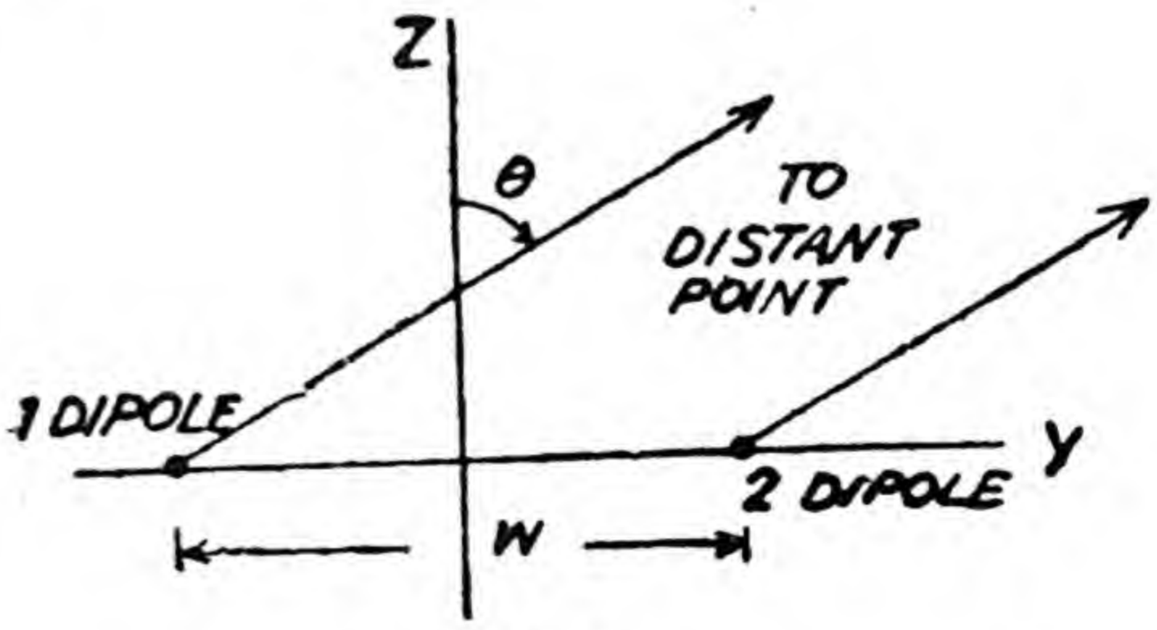


Fig. 39 (*c*)

Since field due to 3 and 1 dipoles in yz-direction (for these two, circular patterns will be in xz-plane) is negligible, we ignore these dipoles.

Thus treating dipoles 2, 4, as two isotropic point sources, the total field at a far point due to the two dipoles will be

$$E_\phi = -E_{\phi 0} e^{j\psi/2} + E_{\phi 0} e^{-j\psi/2}$$
$$= -2jE_{\phi 0} \sin \psi/2,$$

where $E_{\phi 0}$ is electric field from individual dipole which is equal to $\frac{j60.\pi I_0 l}{\lambda r}$ from dipole theory. (In this expression $\sin\theta$ is 1 because here dipole is along x-axis and therefore, if θ from x-axis in any direction in yz-plane is measured, it will come out to be 90°).

In this problem, length l is the height h of the loop and let retarded current on the loop be I_s, then

$$E_{\phi 0} = \frac{j60.\pi.I_s h}{\lambda r}.$$

Thus expression for total field

$$E_{\phi 0} = -2j \frac{j.60.\pi I_s h}{\lambda r} \sin \frac{\psi}{2}$$
$$= \frac{120\pi I_s h}{\lambda r} \sin \frac{\psi}{2}.$$

Also from previous theory of arrays,

$$\psi = \frac{2\pi\, d \sin\theta}{\lambda}.$$

But here instead of d, we should write W, *i.e.*

$$\psi = \frac{2\pi W \sin\theta}{\lambda}$$

Further if $W \ll \lambda$, we can write $\sin \psi/2$ equal to $\psi/2$.

Hence $E_\phi = \frac{120\pi I_s h}{\lambda r} \cdot \frac{\psi}{2} = \frac{120\pi I_s h}{\lambda r} \cdot \frac{2\pi W \sin\theta}{\lambda} \cdot \frac{1}{2}$

$$= \frac{120\pi^2 I_s}{r} \cdot \frac{Wh \sin\theta}{\lambda^2}$$

$$= \frac{120\pi^2 I_s}{r} \cdot \frac{A}{\lambda^2} \sin\theta,$$

where A is the area of the loop.

If there are N turns in the loop, the total electrostatic field radiated in yz-plane will be

$$|E_\phi|_{total} = \frac{120\pi^2 N A}{r \lambda^2} \cdot I_s \sin\theta.$$

If θ be measured from y-axis, then

$$| E_\phi |_{total} = \frac{120\pi^2 N}{r}\left[\frac{A}{\lambda^2}\right]. I_s \cos\theta,$$

gives the instantaneous value of electric component of the radiation field.

The corresponding magnetic field is

$$(H_\phi)_{total} = \frac{(E_\phi)_{total}}{377} = \frac{(E_\phi)_{total}}{120\pi}$$

$$= \frac{\pi N}{r}\left[\frac{A}{\lambda^2}\right] I_s \cos\theta.$$

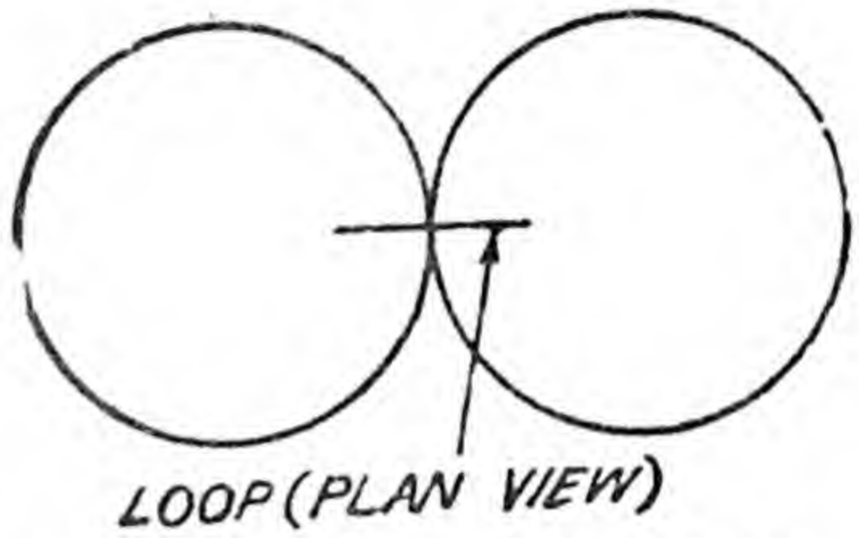


Fig. 39. (*d*) Directional characteris ic of loop anter na.

The field pattern represented by equations (1) and (2) is shown in figure 39 (*d*). Pattern appears as the cross-section of the figure of revolution about an axis that is perpendicular to the plane of the loop. This pattern due to loop antenna remote from ground is exactly the same as that of a doublet antenna placed at the centre of the loop with its axis perpendicular to the plane of loop. The main difference in two patterns is that the field radiated from a loop is every where polarised in a direction parallel to the plane of the loop.

NUMERICAL EXAMPLES

Ex. 1. *Find the voltage induced by a plane wave of field strength* ·01 *volt/meter and wavelength* 300 *meters in a loop of* 1 *square metre of* 12 *turns. The plane of the loop being in the direction of wave propagation.*

The voltage induced is given by

$$e = \frac{2\pi NA \cos\theta}{\lambda} E_m.$$

Given $E_m = \cdot 01$ volt/metre,

$\lambda = 300$ metres,

$A = 1$ square metre,

$\theta = 0°$.

$$\therefore \quad e = \frac{2\pi.12.1.\cos 0°}{300} \times \cdot 01 = \frac{24 \times 3\cdot 14 \times \cdot 1}{300}$$

$$= \frac{0\cdot 0536}{300} \text{volts}$$

$$= \mathbf{2\cdot 5 \text{ milli volts}}$$

Ex. 2. *In a field strength measuring set of* 500 *kc/s. using a loop aerial composed of* 70 *turns* 1 *metre square area, it was found that a maximum potential difference of* 2 *millivolts was developed across this loop aerial when the latter was tuned to resonance. Calculate what was field strength, given that Q of loop aerial is* 30.

The maximum e.m.f. is induced only when the plane of the coil coincides with the direction of transmitted signal *i.e.* for $\theta=0$, π, 2π. Thus

$$|e|_{max}=\frac{2\pi AN}{\lambda}E_m.$$

Given $|e|_{max}=2.10^{-3}$ volts,

$N=10$,

$A=1$ sq. metre,

$$\lambda=\frac{c}{n}=\frac{3.10^8}{500.10^3}=\frac{3.10^3}{5}$$

$=600$ metres.

Therefore, $2.10^{-3}=\frac{2\pi.10.1}{600}E_m.$

$$\therefore \quad E_m=\frac{60.10^{-3}}{3{\cdot}14}\text{ volts/metre}$$

$$=\frac{60}{3{\cdot}14}\text{ mV/metre}$$

$=19{\cdot}107$ mV/metre.

28·18. PARABOLIC REFLECTORS :

We know that an array of linear antennas can be employed to achieve a directional radiation pattern in which the radiated power is concentrated in a narrow beam. The same directivity objectives can also be achieved by the use of reflectors and lenses. Parabolic reflectors are employed when it is convenient to build antennas with aperatures of many wavelengths.

Arrays are commonly employed at lower frequencies. They are used upto about 1000 mc/s although in special case they may be used upto atleast 3,000 mc/s. Reflectors and lenses are most common above 1,000 mc/s, although parabolic reflectors are fairly common down to atleast 100 mc/s. Lenses are basically microwave devices, not ordinarily used below about 3000 mc/s.

In frequency region around 1000 mc/s, the choice between an array and a paraboloidal reflector may sometime be difficult. Arrays are employed when scanning by array phasing is desired, whereas reflectors are indicated when broad band operation or low noise temperature is desired. Sometimes a combination—a linear array feed and a parabolic cylinder reflector is employed.

The principal advantage of lenses over reflector is that the feed and feed support structure do not block the aperture, for the rays are transmitted through the lense rather than returned towards the feed. Since feed for lenses can be placed farther off the optical axis, they can also be employed in applications requiring a beam that can be moved angularly with respect to the axis. Further permissible mechanical tolerances are somewhat greater for lenses than for reflectors. On the other hand, lenses are somewhat bulkier expensive for the same gain and bandwidth than are reflectors. But

these factors are less significant at very short wavelengths, above 10,000 mc/s-a region in which lenses are most commonly used.

Focussing by a Parabolic Reflector : A parabolic reflector antenna is shown in figure 40 (*a*). The parabola reflects the waves originating from a source at the focus into a parallel beam and consequently these spherical waves originated by the source at the focus of the reflector are converted into *plane waves of uniform phase* across the mouth or aperture of the parabola. This can be understood as follows :

Parabolic curve is defined such that the distance from a fixed point F, called focus, to any point, say A_3, on the curve is always equal to the perpendicular distance between that point A_3 on the curve and fixed line called directrix, *i.e.*

$$FA_3 = A_3C_3.$$

$$FA_2 = A_2C_2 \text{ etc.}$$

Now we see from fig. 40 (*b*) that

$$C_3B_3 = C_3A_3 + A_3B_3 = FA_3 + A_3B_3$$

$$= \text{path } FA_3B_3.$$

Also

$$C_2B_2 = C_2A_2 + A_2B_2 = FA_2 + A_2B_2$$

$$= \text{path } FA_2B_2.$$

But $$C_3B_3 = C_2B_2,$$

so that $$\text{path } FA_3B_3 = \text{path } FA_2B_2.$$

Or, in otherwords the length of this path is equal to the distance between the directrix and plane BB' irrespective of the angle at which the ray is incident on the curve. This means that the *phaser* of all waves thus arriving at the aperture are the *same* and *a plane wavefront* is created in the aperture plane.

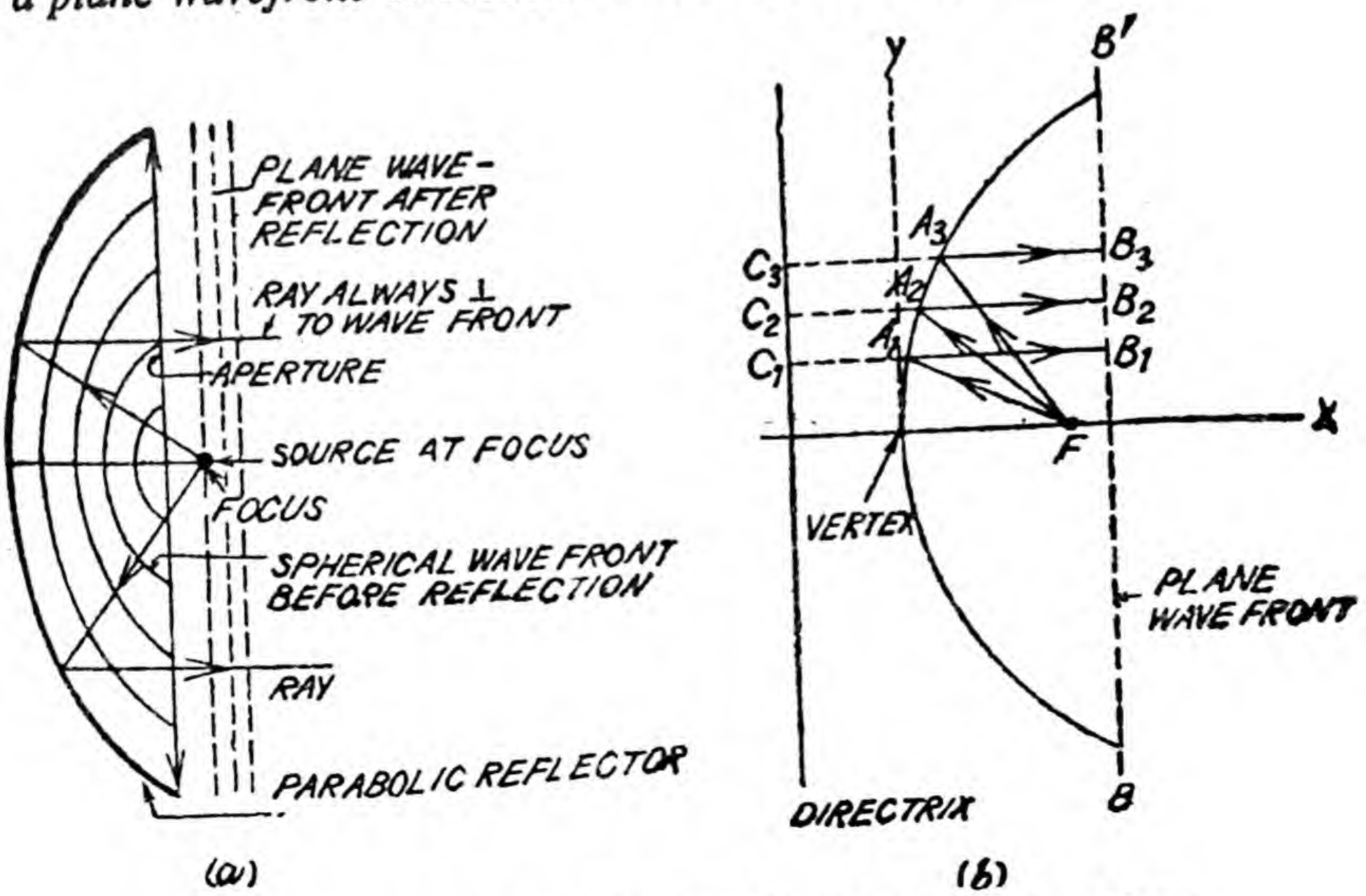
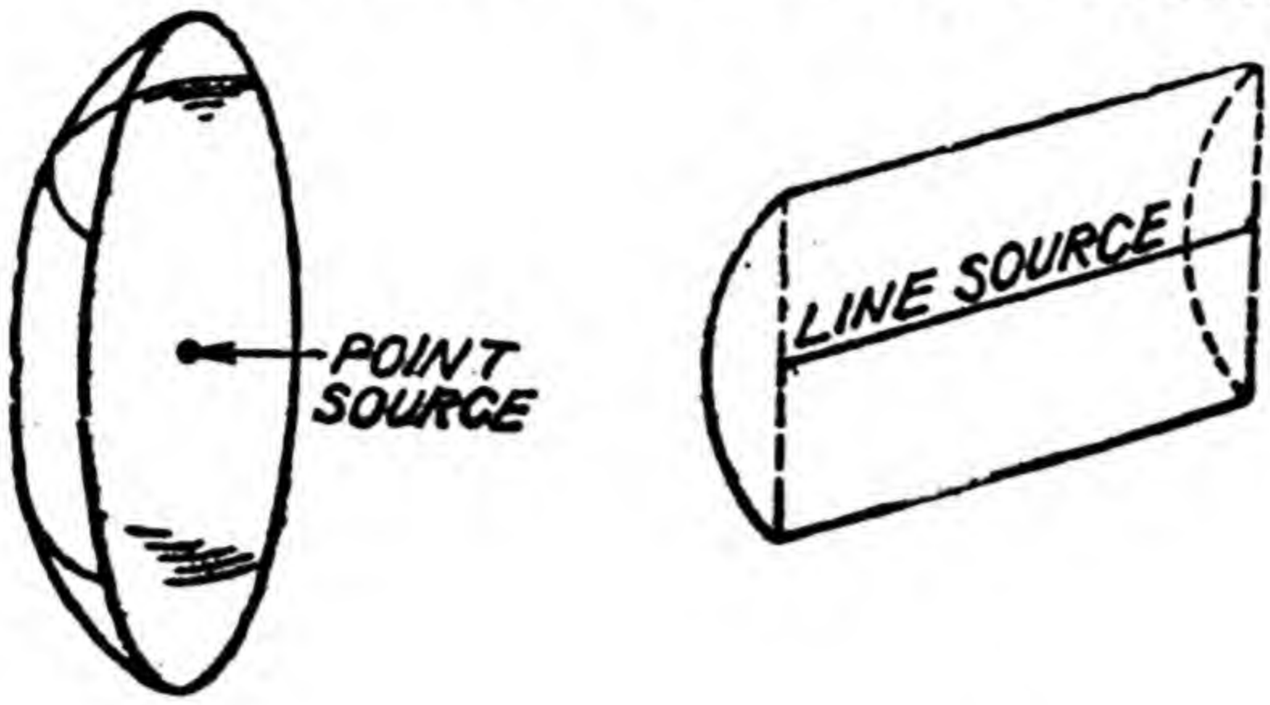


Fig. 40. Parabolic reflector focussing action.

Paraboloids and Parabolic Cylinders as reflectors : A parabola being a plane curve, is two dimensional whereas a parabolic reflector, being a curved surface, is a three dimensional object.



(*a*) (*b*)

Fig 1. Paraboloid and Parabolic cylinder reflector.

There are two types of such a curved surface; one is paraboloid formed by rotating the parabola about its axis shown in fig. 41 (*a*) and the other is a parabolic cylinder formed by translating the parabola of fig. 40 (*b*) in z-direction (perpendicular to the plane of paper) shown in fig. 41 (*b*). The aperture diameter of a paraboloid may range from few inches to thousand feet.

Field distribution across the aperture of Parabolic Reflector :

(a) For Cylindrical Parabola : Let a line source be placed at the focus along *Z*-axis.

Let this be isotropic in *XY* plane.

Suppose p_y is the power density at point *A* distant *y* from *X*-axis. Then power *P'* in a strip of length *AB* ($\approx dy$) and breadth *dz* will be

$$P' = dz\, dy\, p_y.$$

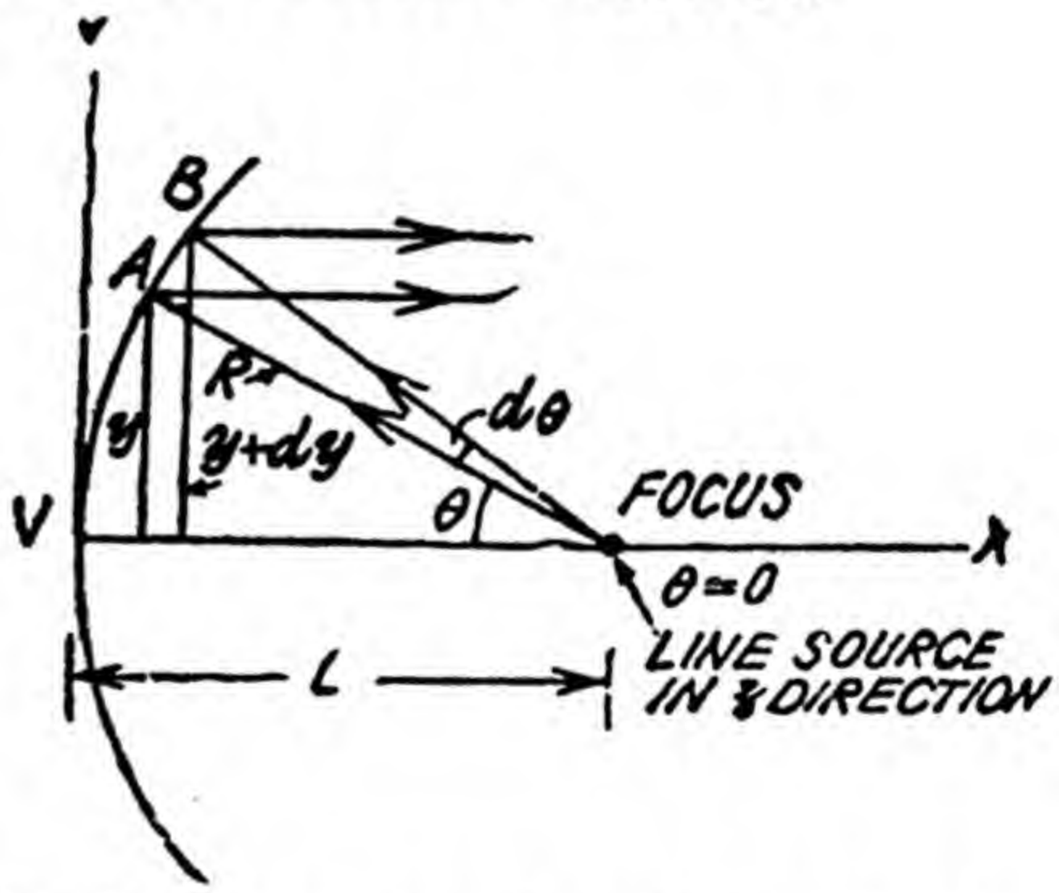


Fig. 42. Cylinrdical parabola; *Z*-axis is perpendicular to plane of paper.

If breadth is taken as unity *i.e.* $dz = 1$ or for a unit distance in *z*-direction, the power in the strip *AB* is

$$P = p_y\, dy. \qquad \ldots(1)$$

Suppose u' is the power per unit angle per unit length in the *z*-direction, then

$$P = u'\, d\theta. \qquad \ldots(2)$$

So that from equations (1) and (2), we get

$$\frac{p_y}{u'} = \frac{1}{dy/d\theta} = \frac{1}{\frac{d}{d\theta}(R \sin \theta)}, \qquad \ldots(3)$$

where $y = R \sin \theta.$

Further we know that

path length 2 = path length 1

$$R + R\cos\theta = L + L$$

or

$$R = \frac{2L}{1+\cos\theta}.$$

or

Therefore equation (3) assumes the form

$$\frac{p_v}{u'} = \frac{1}{\frac{d}{d\theta}\left(\frac{2L\sin\theta}{1+\cos\theta}\right)}$$

$$= \frac{1+\cos\theta}{2L}. \qquad \ldots(4)$$

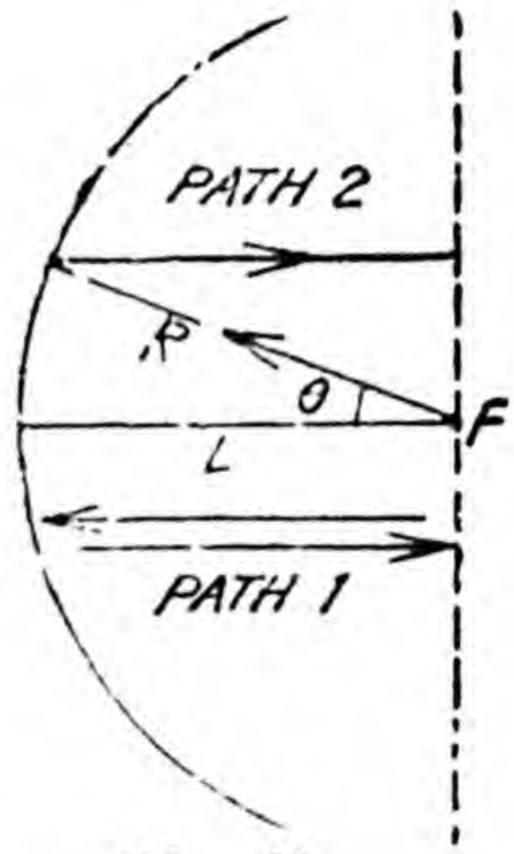


Fig. 43.

The ratio of power density p_θ at θ to the power density p_0 at $\theta = 0$ is

$$\frac{p_\theta}{p_0} = \frac{\left(\frac{1+\cos\theta}{2L}\right)u'}{\left(\frac{1+\cos 0}{2L}\right)u'} = \frac{1+\cos\theta}{2}. \qquad \ldots(5)$$

Therefore, the relation in corresponding field intensities would be

$$\frac{E_\theta}{E_0} = \sqrt{\left(\frac{p_\theta}{p_0}\right)}$$

$$= \sqrt{\left(\frac{1+\cos\theta}{2}\right)}$$

or

$$E_\theta = E_0\sqrt{\left(\frac{1+\cos\theta}{2}\right)}. \qquad \ldots(6)$$

The expression gives the field intensity at a distance y from the axis as a function of E_θ — the field intensity at the vertex.

Equation (6) furnishes completely the information to sketch the secondary field pattern because E_0, the field at the vertex, is merely the field due to the primary antenna placed at the focus. This means we can easily compute E_0 from the earlier methods once the type of the feed radiator or primary antenna is known to us and consequently, we can obtain E_θ. The form factor of field E_θ then decides the shape of secondary field pattern. The calculation of E_θ easily leads to the computation of power and directivity.

(b) For Paraboloid. A paraboloidal reflector is shown in fig. 44. An isotropic point source is located at the focus. The total power P through the annular section of radius r and width dr is

$$P = 2\pi r\, dr\, p_r, \qquad \ldots(7)$$

where p_r is the power density at a distance r from the axis.

Suppose u is the power per unit solid angle, then the total power through the same annular section will be

$$P = u\, d\omega,$$

where $d\omega$ is the solid angle subtended by the annular section of radius r and width dr. It will be

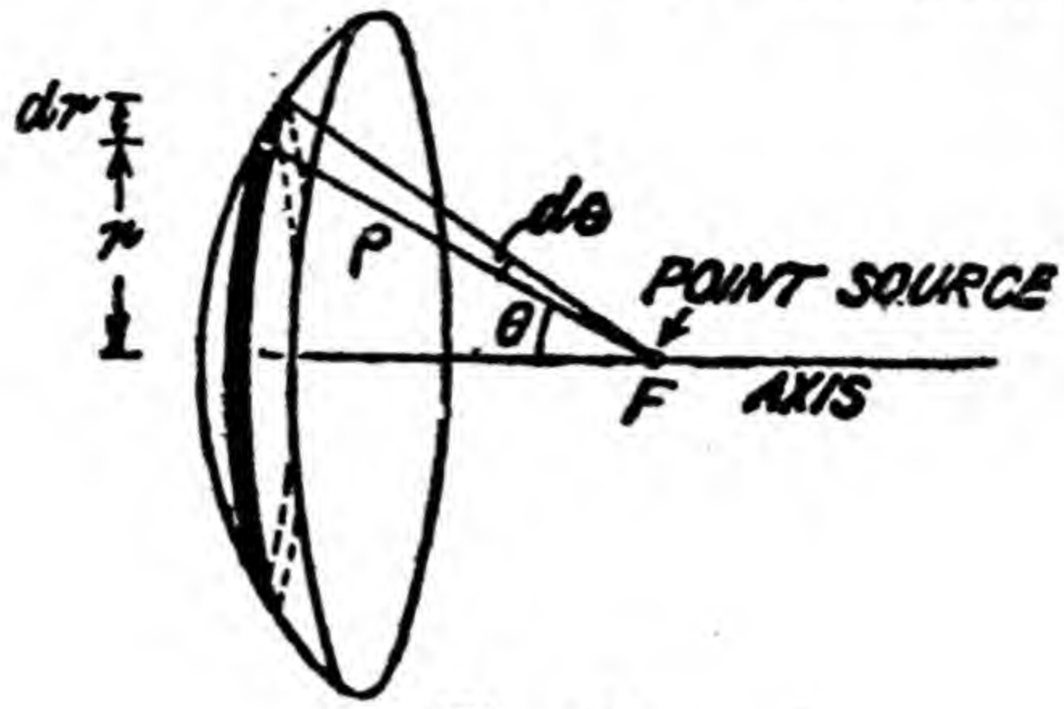


Fig. 44.

$$d\omega = \frac{dA \cos 0°}{\rho^2}$$

$$= \frac{2\pi\rho \sin\theta\ \rho\, d\theta}{\rho^2} = 2\pi \sin\theta\, d\theta.$$

Therefore,

$$P = 2\pi \sin\theta\, d\theta . u. \qquad \ldots(8)$$

From equations (7) and (8), we get

$$\frac{p_r}{u} = \frac{\sin\theta}{r\,(dr/d\theta)},$$

but

$$r = \rho \sin\theta.$$

$$= \frac{2L \sin\theta}{1+\cos\theta}$$

we find

$$\frac{p_r}{u} = \frac{\sin\theta}{\left(\frac{2L\sin\theta}{1+\cos\theta}\right)\frac{d}{d\theta}\left(\frac{2L\sin\theta}{1+\cos\theta}\right)}$$

$$= \left(\frac{1+\cos\theta}{2L}\right)\frac{1}{\frac{d}{d\theta}\left(\frac{2L\sin\theta}{1+\cos\theta}\right)}$$

$$= \frac{(1+\cos\theta)^2}{4L^2}. \qquad \ldots(9)$$

From equation (9), we atonce arrive at

$$\frac{p_\theta}{p_0} = \frac{\frac{(1+\cos\theta)^2}{4L^2}\,u}{\frac{(1+\cos 0)^2}{4L^2}\,u} = \frac{(1+\cos\theta)^2}{4}$$

where p_θ and p_0 are the power densities at $\theta=\theta$ and $\theta=0°$ respectively. Therefore relation in corresponding field intensities would be

$$\frac{E_\theta}{E_0} = \sqrt{\left(\frac{p_\theta}{p_0}\right)}$$

$$= \frac{1+\cos\theta}{2}$$

or $$E_\theta = E_0 \frac{(1+\cos\theta)}{2} \quad \ldots(10)$$

The expression gives the field intensity at a distance r from the axis as a function of E_0—the field intensity at the axial point on the surface.

Equation (10) furnishes the complete information to sketch secondary field pattern. E_0, the field at the vertex due to the primary antenna being known, E_θ is easily calculated and hence the calculations of power and directivity.

Primary and Secondary Patterns : Field Radiators or Source :

The field pattern of the feed radiator or source, called primary radiator, is termed as primary pattern whereas the field pattern of the entire antenna including source and reflector is called the secondary pattern.

(*a*) *Isotropic Source* : Suppose isotropic source is placed at the focus of a paraboloid to feed the latter. Then as is obvious from the figure 45 only shaded portion will be reflected back by paraboloid as a plane wave of circular cross section.

If L is even multiple of $\lambda/4$ then directed radiation from source will be in phase opposition to reflected radiation because distance travelled by direct radiation will be

$$FV = L = n\lambda/4,$$

where $$n = 2, 4, 6, \ldots$$
$$= \lambda/2, \lambda, 3\lambda/2, \ldots$$

so that at point V phase will be reversed and thus direct radiation

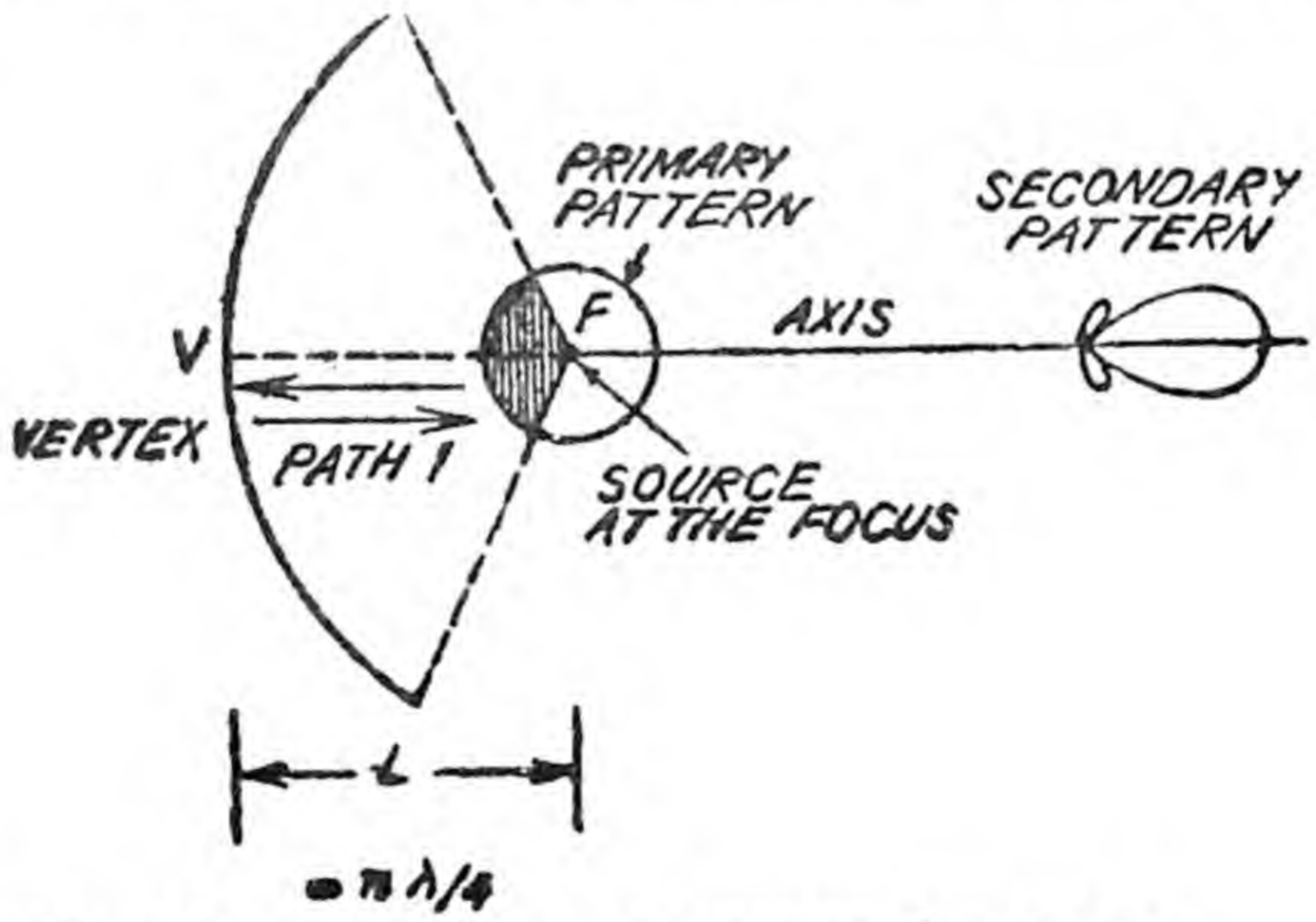


Fig. 40. Reflectors with isotropic source.

will cancel the central region of the reflected wave. If, however, L is odd multiple of $\lambda/4$ then

$$FV = L = n\lambda/4, \quad \text{where} \quad n = 1, 3, 5\ldots$$
$$= \lambda/4, 3\lambda/4, 5\lambda/4, \ldots$$

so that at point V phase is not reversed and then the direct radiation in this case will reinforce the central region of the reflected wave. We observe that primary antenna lies in the path of the

reflected wave. This situation gives rise to two main disadvantages :

(*i*) primary antenna acts as an obstruction and thus blocks the central portion of the aperture,

(*it*) reflected waves from central portion of the reflector interact with the primary antenna and result in mismatching.

To avoid these effects, primary antenna may be displaced to one side and only a portion of reflector is used as shown in fig 46.

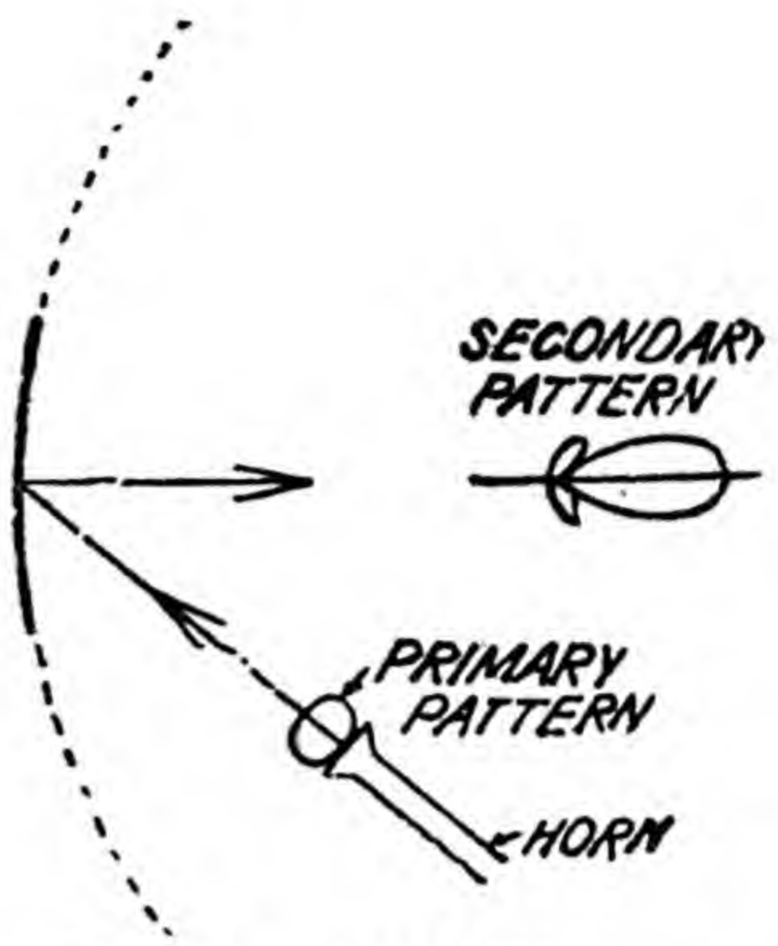


Fig. 46. Reflector with horn antenna as feeder.

(*b*) *Other Feed Radiators* : Suitable directional patterns may be achieved by using various types of primary radiators. One such radiator, horn antenna, is shown in fig. 46. We shall not enter into the the details.

Illumination Tapper : The radiation from the feed is called the '*illumination*' of the reflector. The variation of the intensity of radiation over the aperture is called '*illumination taper*.' The ratio of the power density at the edge of the aperture to that at the centre is called '*edge taper*' :

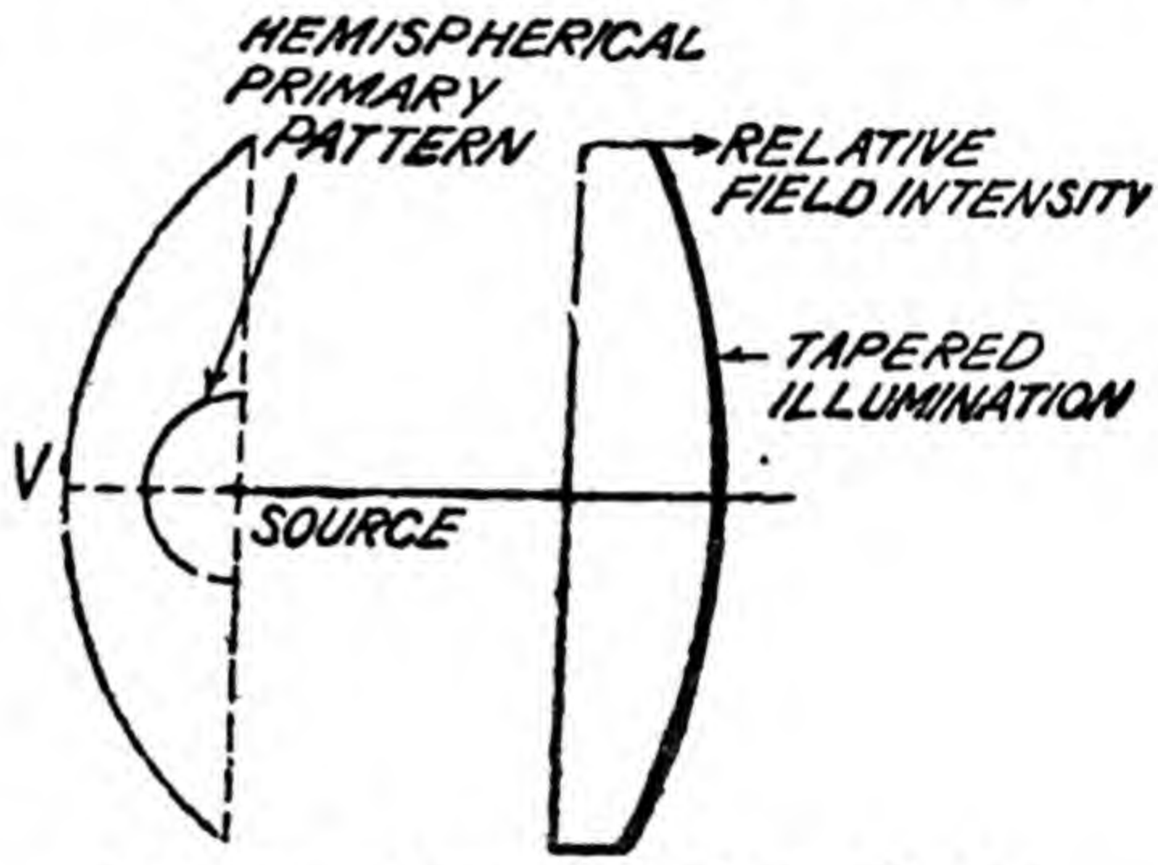


Fig. 47. Hemispherical primary pattern.

With hemispherical primary pattern, as shown in fig. 47, the amplitude of relative field intensity is tapered. An excessive taper causes the loss of gain and sharpness of the pattern is reduced but, on the other hand, too little taper results in unnecessary appearance of the side lobes and back lobes.

28·19. RECTANGULAR HORNS :

A broad class of radiating antennas, particularly useful in the microwave region, are horn type radiators, paraboloidal reflectors, and lenses. In each case electromagnetic energy must pass through a physically distinct aperture. It is possible to show that the field in the aperture behaves like an equivalent source. This fact is related to the Huygen's principle which states that *each element of a wavefront may be considered as a secondary source.* By 'equivalent theorem' we can replace an element of area of an advancing wavefront (say $dx\ dy$ if z is the direction of propagation) by electric and magnetic current elements. Therefore, in order to

calculate the field at a point due to aperture radiation, we shall simply determine the radiation fields of these current elements. In Chapter 27 we have pictured the field configuration inside the guide operating in $TE_{1,0}$ mode. Obviously, there exists field distribution across the mouth of the guide. By the concept of 'equivalent source', one can explain the action of a horn as though the mouth of the horn is replaced by a distribution of elementary dipoles in which each dipole radiates a field, proportional to $E(1+\cos\theta)$, where θ is the angle measured with respect to the direction in which the horn points, and E is the intensity of the electric field where the dipole is assumed to exist. Dipoles are polarised in the same direction as the electric field at the mouth of the horn. The actual field pattern of the radiation from the horn is the summation of the radiated fields produced by all these elementary dipoles as obtained on integration.

If a rectangular aperture lies in the yz-plane of a cartesian coordinate system with its centre at the origin, then the beam axis will coincide with the x-coordinate axis. If the sides of the aperture rectangle are parallel to the y-axis and z-axis, the pattern in the xy-plane is given by the integral :

$$E(\alpha)=\frac{1}{D_y}\int_{-D_y/2}^{D_y/2} \exp.\left[\, j\left(\frac{2\pi y \sin\alpha}{\lambda}\right)\right] dy,$$

where α is the azimuth angle in the xy-plane and D_y is the length of the aperture in the y dimension. The corresponding expression for the xz plane pattern is obtained by putting z for y everywhere in the above relation and substituting $\cos\beta$ for $\sin\alpha$, where β is azimuth angle in xy plane.

The solution of this integral is

$$E(\alpha)=\frac{\sin\left(\frac{\pi D_y \sin\alpha}{\lambda}\right)}{\left(\frac{\pi D_y \sin\alpha}{\lambda}\right)}$$

This pattern is often plotted in terms of a coordinate designated u, defined as

$$u=\frac{\pi D_y \sin\alpha}{\lambda}$$

so that the solution of the integral is written as

$$E(u)=\frac{\sin u}{u}.$$

At $u=0$, this expression is indeterminate so that, by differential calculus, it can be shown that

$$\lim_{u\to 0}\frac{\cos u\, du}{E(u)}=1$$

or at $u=0$, $E(u)=1$ which is maximum value of $E(u)$ so that normalised field is

$$E_n(u)=\frac{E(u)}{E(u)\,|_{u=0}}=\frac{\sin u}{u}.$$

At $u=\pi$, 2π, 3π etc., $E_n(u)=0$. Therefore these values of u give null directions. At $u=3\pi/2$, $5\pi/2$, $7\pi/2$, secondary maxima occur.

In order to secure greater directivity, a waveguide can be flared out to form an electro-magnetic horn. There can be various types of horn antennas *e.g.*, rectangular horns, circular horns. In fig. 48 only rectangular horns are shown. Rectangular horns are fed from rectangular waveguides.

Such radiators concentrate the radiated wave into a sharply defined pattern and are capable of giving highly directional patterns when the mouth of the guide has a dimension large compared with wavelength. At microwave frequencies ($\lambda=1$ cm.), horns become practical because since then a dimension large compared with wavelength does not involve excessive physical size.

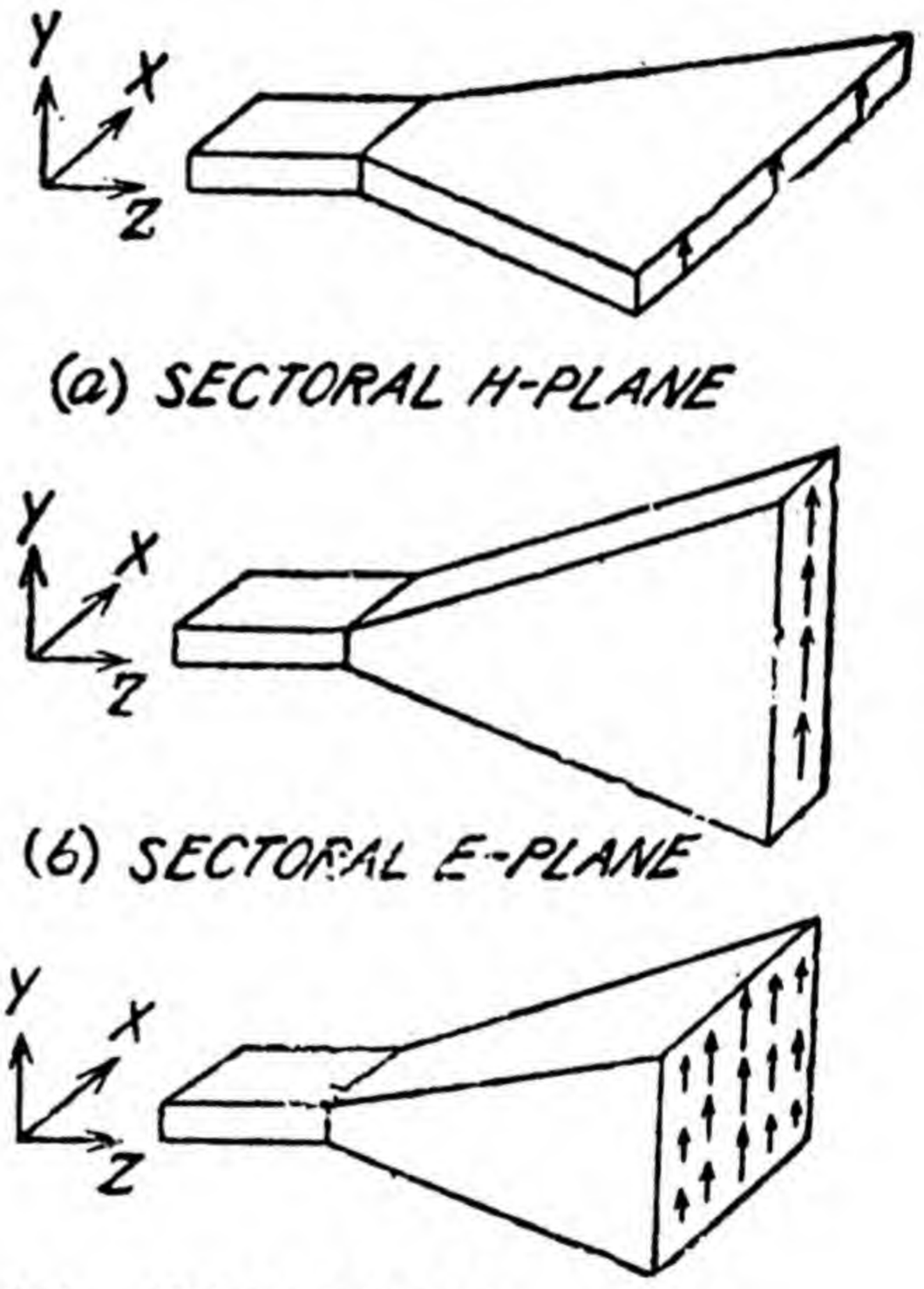


Fig. 48. Types of rectangular horn antenna. Arrows on the aperture indicate the electric field E.

A rectangular guide flared out in one plane only constitutes a sectoral horn, whereas a guide flared out in both the planes forms a pyramidal horn as shown in figs. 48 (*a*), (*b*), and (*c*).

Assume that rectangular wave-guide is operating in $TE_{1,0}$ mode (E in *y*-direction or vertical and H in *x*-direction or horizontal). Then horn shown in fig. 48 (*a*) is flared out in *H* plane and is therefore called a sectoral horn flared in *H*-plane or simply *H*-plane sectoral horn. Similarly, horn, shown in fig. 48 (*b*) is flared out in *E*-plane and is named as *E* plane sectoral horn. When horn is flared in both planes, it is called pyramidal horn.

The field pattern obtained by exciting a horn radiator is determined by the *shape* and *dimensions* of the mouth measured in wavelengths, and the *variations in the magnitude and phase of the field*

distribution produced across the mouth of the horn. The nature of the field distribution across the mouth opening is determined by the horn's shape and by the mode of excitation. The relative phases of the field at different parts of the mouth are determined by the relative distances of different parts of the opening from the throat to the horn and are also influenced by the mode of excitation where higher order modes are involved.

Optimum Horn Dimensions : It has been observed experimentally that the directivity of horn, for given aperture dimensions, is maximum if the field distribution over the aperture is uniform. Variations in magnitude or phase of field across the aperture decreases the directivity (or increases the beamwidth and side lobes).

To obtain uniform field distribution across the aperture, a very long horn with small flare angle is required, but a long horn is not convenient from practical stand point. Therefore, an approximate horn length that compromises well with directivity requirement and practical convenience is chosen. Such a horn is called an *optimum horn*. To calculate the dimensions of optimum horn, we proceed as follows :

Let us consider the longitudinal section of sectoral *H*-plane horn as shown in fig. 49. Let

L be the axial length of horn (from throat to the aperture),

ϕ_0 be the total flare angle, (it is in horizontal plane),

A be the aperture,

δ be the difference in path length of the two waves-one reaching aperture at the axial point and the other reaching aperture at the side of horn.

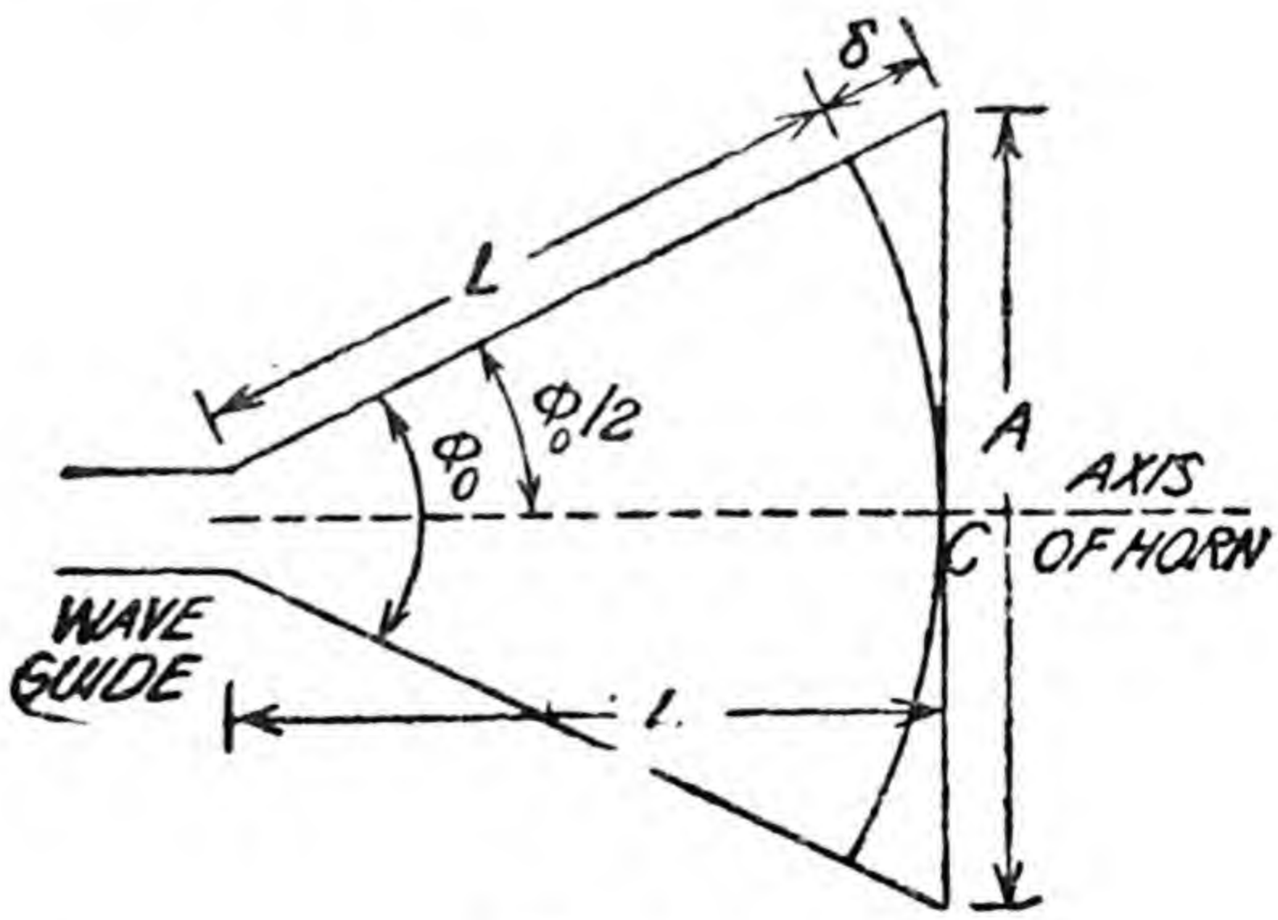


Fig. 49.

Field distribution over the entire aperture is uniform provided δ is only a very small fraction of a wavelength. For a constant length L, the directivity of horn increases as the aperture A and

flare angle ϕ_0 are increased. If the flare angle is not very large, then the ratio

$$\frac{L}{L+\delta} \approx 1,$$

so that the wave that reaches the axial point and the side point of the aperture do not have any significant phase difference and the effect of δ on the distribution of field magnitude over the aperture can be neglected. However, if A and ϕ_0 are so large as to make δ equivalent to 180° electrical degrees, the phase of the wave reaching the side point will be reversed compared to the phase of the wave reaching the axial point of the aperture. This results in the decrease of directivity, increasing thereby the beamwidth and side lobes. The maximum directivity occurs at the largest flare angle for which δ does not exceed a certain value (δ_0). Thus the optimum horn dimensions can be related by

$$\frac{L}{L+\delta_0} = \cos \phi_0/2$$

or
$$\varepsilon_0 = \frac{L}{\cos \phi_0/2} - L \qquad \ldots(1)$$

or
$$L = \frac{\delta_0 \cos \phi_0/2}{1-\cos \phi_0/2} \qquad \ldots(2)$$

or
$$\phi_0 = 2 \cos^{-1} \frac{L}{L+\delta_0}. \qquad \ldots(3)$$

These relations determine the optimum horn dimensions. The value of δ_0 usually preferred lies in the range of 0·1 to 0·4 free space wavelength.* Suppose we want to calculate the flare angle for the maximum directivity of a 10 wavelength optimum horn ($L=10\lambda$) with $\delta_0=0{\cdot}25\lambda$ then from equation (3), we have

$$\phi_0 = 26°.$$

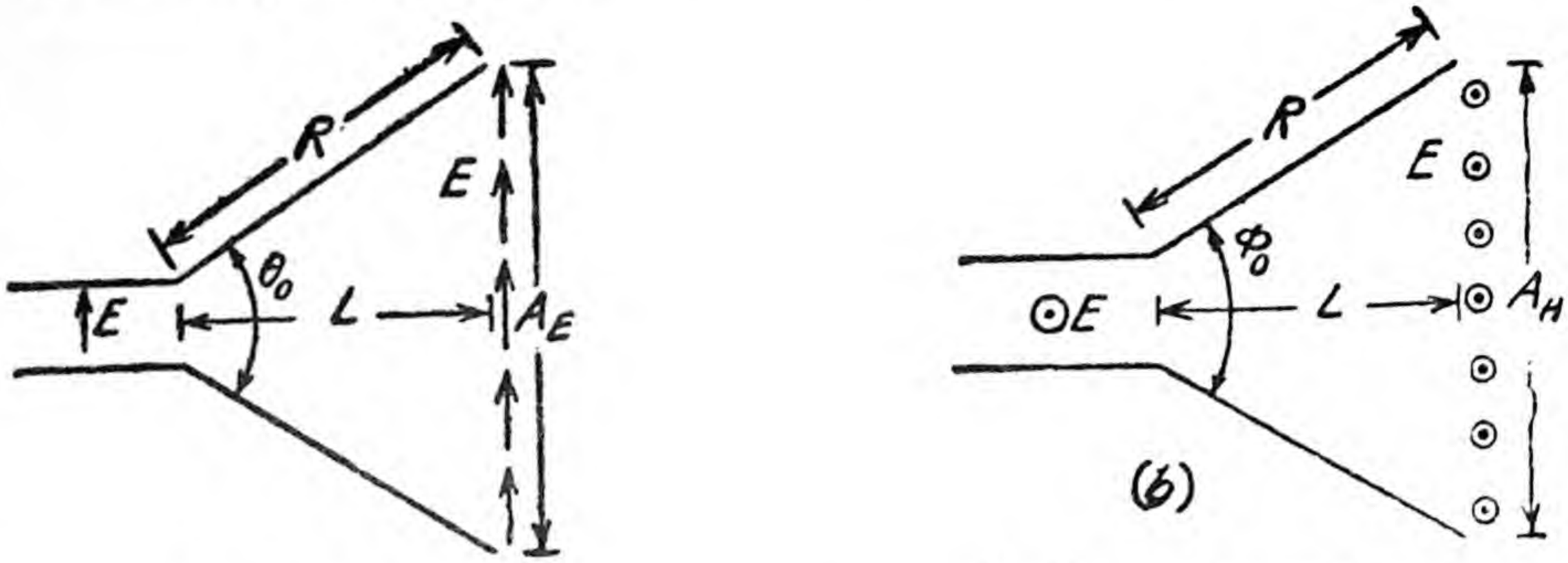
We can estimate length L of the horn. Guides are preferred at and above 30,000 mc/s. So that λ is 1 0 cm. This means flared length will be only 10 cm. in the case of $L=10\lambda$.

Field Patterns : In fig. 50 (*a*) and 50 (*b*) ; *E* plane cross section and *H* plane cross section are shown respectively. θ_0 is the total flare angle in *E* plane (*vertical plane*) and ϕ_0 is the tolal flare angle in *H* plane (*horizontal plane*). Let A_E and A_H denote the apertures in *E* plane and *H* plane respectively and guide is excited in $TE_{1,0}$ mode.

When the aperture in both planes of a rectangular horn exceeds one wavelength, the pattern in one plane is substantially *independent* of the aperture in other plane. This means that *H* plane patterns of a *H* plane sectoral horn fig 48 (*a*) and a pyramidal horn with the

*Since wavelength in horn λ_h depends upon the horn dimensions, it is more convenient to express δ_0 in terms of free space wavelength λ.

same H plane cross section, fig. 48 (c), will be the same. Same is true for E plane patterns.



E-plane cross-section *H*-plane cross-section

Fig. 50.

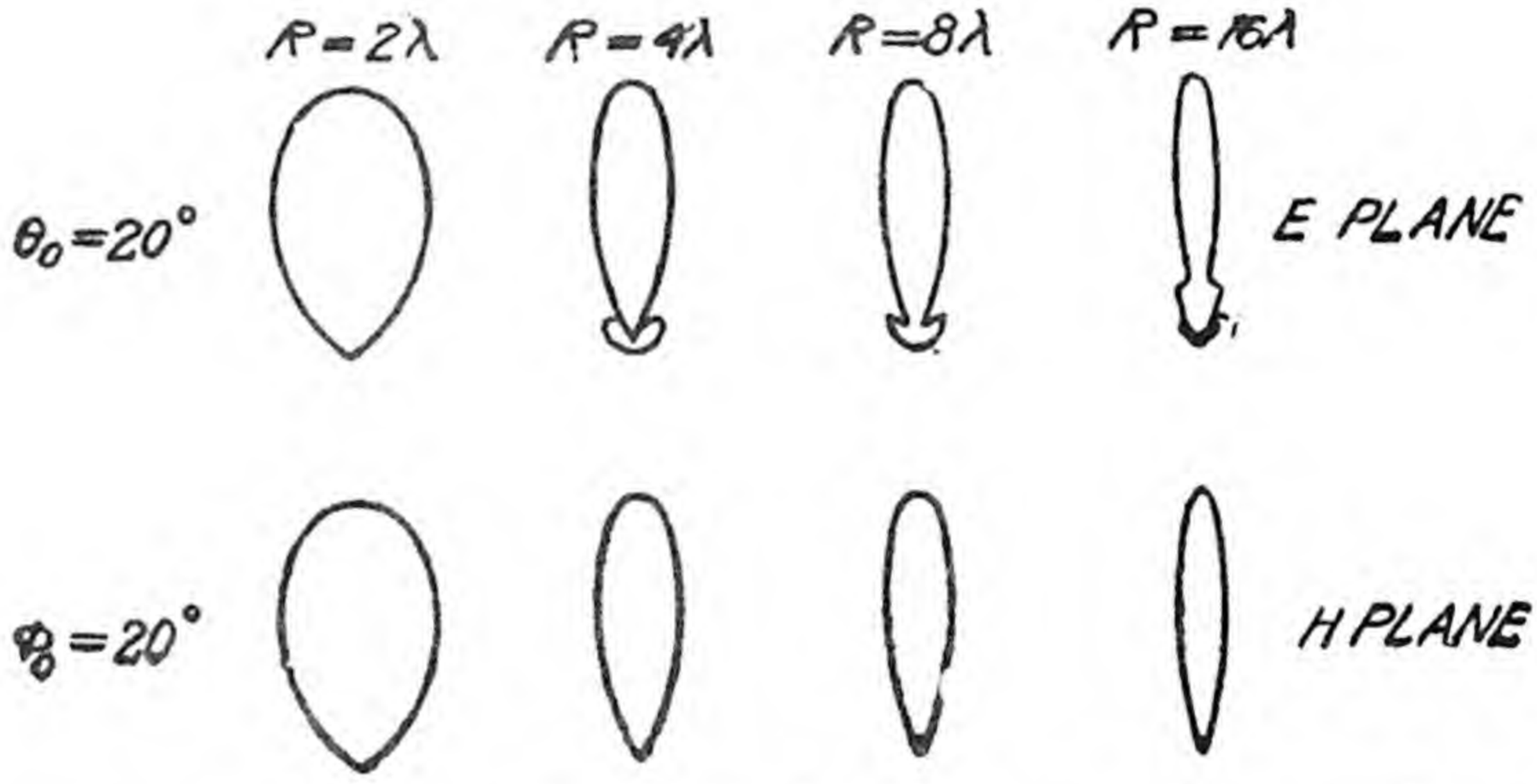


Fig. 51. (*a*). Measured E and H plane patterns of rectangul r horn operating in $TE_{1,0}$ mode, as a function of horn length

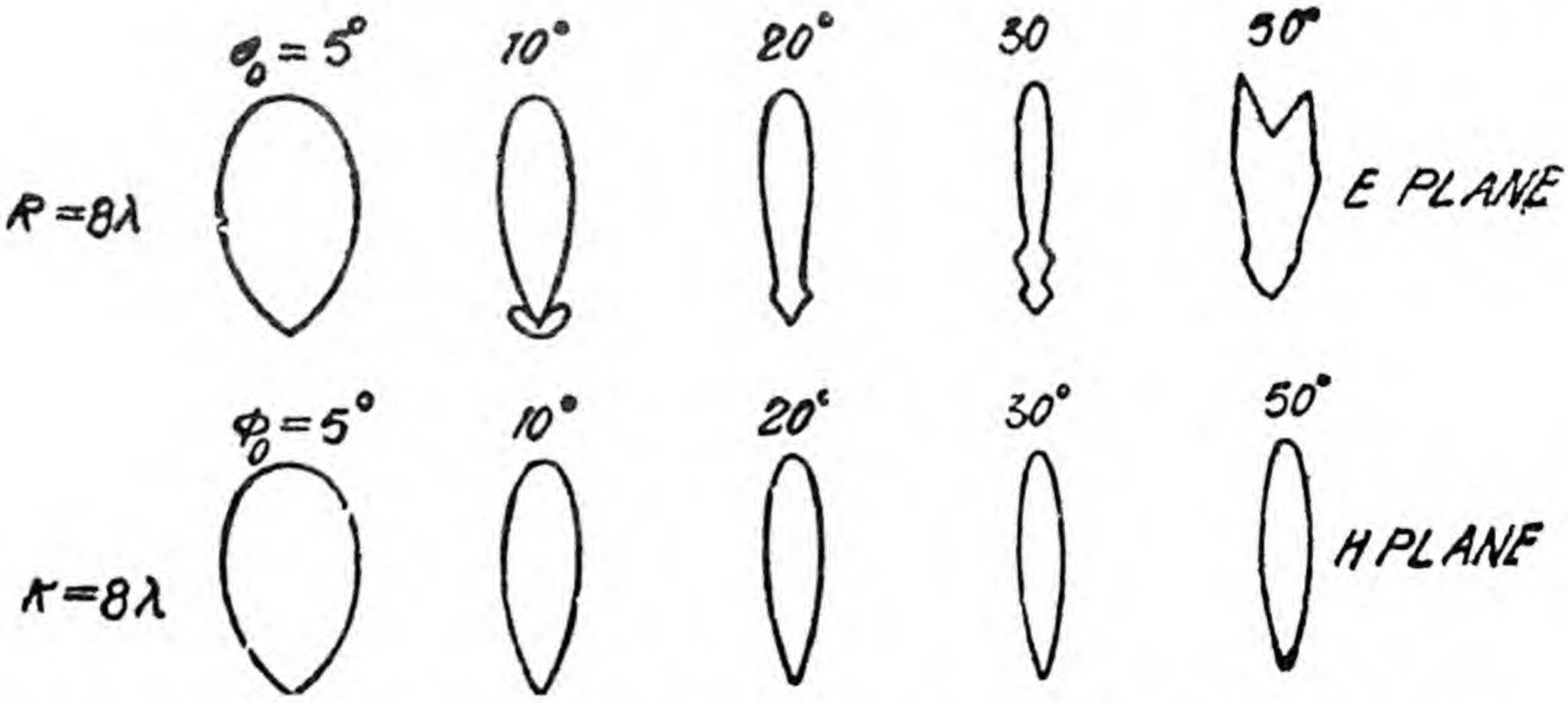


Fig. 51. (*b*). Measured E and H plane patterns of rectangular horns. operating in $TE_{1,0}$ mode, as a function of flare angle.

Measured patterns in E plane and H plane are compared as a function of R in fig. 51 (*a*) and as a function of corresponding

flare angles θ_0 and θ_0 in fig. 51 (*b*). In fig. 51 (*a*) it is to be noted that *E* plane patterns have minor lobes whereas *H* plane patterns have practically none. Same is true about patterns in fig. 51 (*b*).

Field Calculation: W. L. Barrow and L.J. Chu in 1936 have shown by developing from Maxwell's equations a wave equation in cylindrical coordinates subject to the boundary conditions at the walls of the horn, that the fields inside the *E* plane sectoral horn are given, in terms of cylindrical coordinates (r, θ, z), by

$$E_\theta = K_1 \cos\left(\frac{\pi z}{w}\right)[H_1^{(2)}(kr) + K_2 H_1^{(1)}(kr)]$$

$$H_r = \frac{j\pi K_1}{\mu\omega w} \sin\left(\frac{\pi z}{w}\right)[H_1^{(2)}(kr) + K_2 H_1^{(1)}(kr)]$$

$$H_z = \frac{j\beta K_1}{\mu\omega} \cos\left(\frac{\pi z}{w}\right)[H_0^{(2)}(kr) + K_2 H_0^{(1)}(kr)]$$

where K_1 = complex constant,

K_2 = ratio of reflected to incident wave amplitudes at a point in the horn.

$k = \sqrt{\left[\beta^2 - \left(\frac{\pi}{w}\right)^2\right]}$,

w = height of horn,

$\omega = 2\pi f$,

$H_0^{(1)}$ = Hankel function of first kind and zere order,

$H_0^{(2)}$ = Hankel function of second kind and zero order,

$H_0^{(1)}$ = Hankel function of first kind and first order,

$H_2^{(2)}$ = Hankel function of second kind and first order.

28·20. SELF AND MUTUAL IMPEDANCES OF ANTENNAS:

The impedance presented by an antenna to a transmission line is termed as terminal or *driving point impedance*. If the antenna is remote from ground or other objects and is lossless then this impedance is the same as the *self impedance* of the antenna. This impedance has both reactive and resistive components, respectively termed as *self reactance* and *self resistance* (radiation resistance). In case the antenna is not isolated but is placed in the vicinity of another antenna then the two act as if they are coupled *Mutual impedance* then comes into play. The magnitude and phase angle of mutual impedance depend upon spacing and geometrical configuration of the two antenna systems.

We shall use induced e.m.f. method to determine the self impedance of an antenna. In this method the electric field E_θ produced parallel to the antenna by a known (or assumed) current distribution is calculated and used to compute the voltage $E_z\,dz$ induced in each element dz of the antenna. Then the reciprocity theorem is invoked to determine the voltage at the antennas terminals. The reciprocity theorem may be stated as follows: *If an e. m. f. is applied to the terminals of an antenna A and the current is*

measured at the terminals of another antenna B, then an equal current (in both amplitude and phase) *will be obtained at terminals of antenna A if the same e. m. f. is applied to the terminals of antenna B.* Therefore

$$\frac{V_A}{I_B}=Z_{AB}=Z_{BA}=\frac{V_B}{I_A},$$

where Z_{AB} or Z_{BA} is termed as transfer impedance.

It is, however, assumed that e.m.fs. are of the same frequency and that the medium is linear, passive, and also isotropic. The reciprocity theorem holds not only for two separate antennas but also for two points on the same antenna.

SELF-IMPEDANCE OF A THIN LINEAR ANTENNA :

A centre fed linear $\lambda/2$ antenna with sinusoidal current distribution is shown in the figure 52.

The current distribution along the length of the antenna is given by

$$I_z=I_1 \sin \beta z. \qquad ...(1)$$

Obviously we are considering only those lengths for which at centre point $I_z=I_1$ that is odd multiple of $\lambda/2$. The assumed current distribution is then symmetrical with a current maximum at the terminals.

Suppose V_{11} is the e.m.f. applied at the terminals of the antenna, producing a current I_z at a distance z from the lower end. Then transfer impedance Z_{1z} is

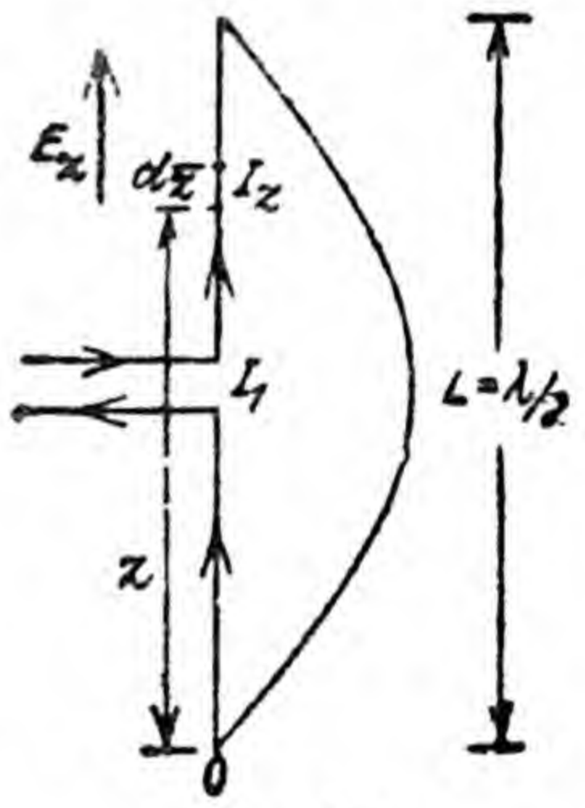


Fig. 52.

$$Z_{1z}=\frac{V_{11}}{I_z}. \qquad ...(2)$$

Let the applied field at the antenna and parallel to it be E_z. This is the field produced by the own current of the antenna. This field induces a field E_{zi} at the conductor such that the boundary conditions are satisfied. For a perfect conductor these are that the total field is zero or that

$$E_z+E_{zi}=0$$

and therefore,

$$E_{zi}=-E_z. \qquad ...(3)$$

The e.m.f. dV_z produced by the induced field over a length dz is

$$-E_z\, dz$$

or

$$dV_z=-E_z\, dz. \qquad ...(4)$$

If the antenna is short-circuited this e.m.f. will produce a current dI_1 at the terminals. Then the transfer impedance Z_{z1} is given by

$$Z_{z1}=\frac{dV_z}{dI_1}. \qquad ...(5)$$

According to reciprocity theorem,

$$\frac{V_{11}}{I_z}=Z_{1z}=Z_{z1}=\frac{dV_z}{dI_1}=\frac{-E_z\,dz}{dI_1} \qquad ...(6)$$

and $$V_{11}\,dI_1=-I_zE_z\,dz.$$

The driving point or terminal impedance is then

$$Z_{11}=\frac{V_{11}}{I_1}. \qquad ...(7)$$

Since it is constant and independent of current amplitude (system being linear , we can write

$$Z_{11}=\frac{dV_{11}}{dI_1} \qquad .. (8)$$

or $$\frac{V_{11}}{I_1}=\frac{dV_{11}}{dI_1}$$

giving $$V_{11}\,dI_1=dV_{11}\,I_1. \qquad ...(9)$$

On applying equation (6), we get

$$dV_{11}=\frac{-I_z}{I_1}\,E_z\,dz. \qquad ...(10)$$

or $$V_{11}=-\frac{1}{I_1}\int_0^L I_zE_z\,dz, \qquad ...(11)$$

where the integration is carried over the length of the antenna. Therefore terminal impedance is

$$Z_{11}=\frac{V_{11}}{I_1}=-\frac{1}{I_1^2}\int_0^L I_z\,E_z\,dz. \qquad ...(12)$$

Since the antenna is isolated Z_{11} is called the self-impedance. Putting the value of I_z from eq. (1), we get

$$Z_{11}=-\frac{1}{I_1}\int_0^L E_z\sin\beta z\,dz. \qquad ...(13)$$

As stated earlier E_z is the field produced at the antenna by its own current, we denote it by E_{11} in place of E_z, so that

$$Z_{11}=-\frac{1}{I_1}\int_0^L E'_{11}\sin\beta z\,dz, \qquad ...(14)$$

which is the final expression for the calculation of self impedance provided we put in it the expression for E_{11} which is yet to be calculated.

Calculation of E_{11} : If V be retarded scalar potential due to the charges on the antenna and A the retarded vector potential, then electric field is

$$\mathbf{E}=-\nabla V-j\omega\mathbf{A}$$

or $$E_z=-\frac{\partial V}{\partial z}-j\omega A_z. \qquad ...(15)$$

We have shown in the article of current element that

$$\mathbf{A}=\mu\int\frac{i\,d\mathbf{l}}{4\pi r},$$

which for the present case becomes

$$A_z=\frac{\mu}{4\pi}\int_0^L I_{z_1}\,dz_1.$$

From equation (1), including retarded time factor, we can write

$$I_{z_1}=I_1 \sin \beta z_1\, e^{j\omega(t-r/v)},$$

so that

$$A_z=\frac{\mu I_1\, e^{j\omega t}}{4\pi}\int_0^L \frac{\sin \beta z_1 . e^{-j\beta r}}{r}\,dz_1$$

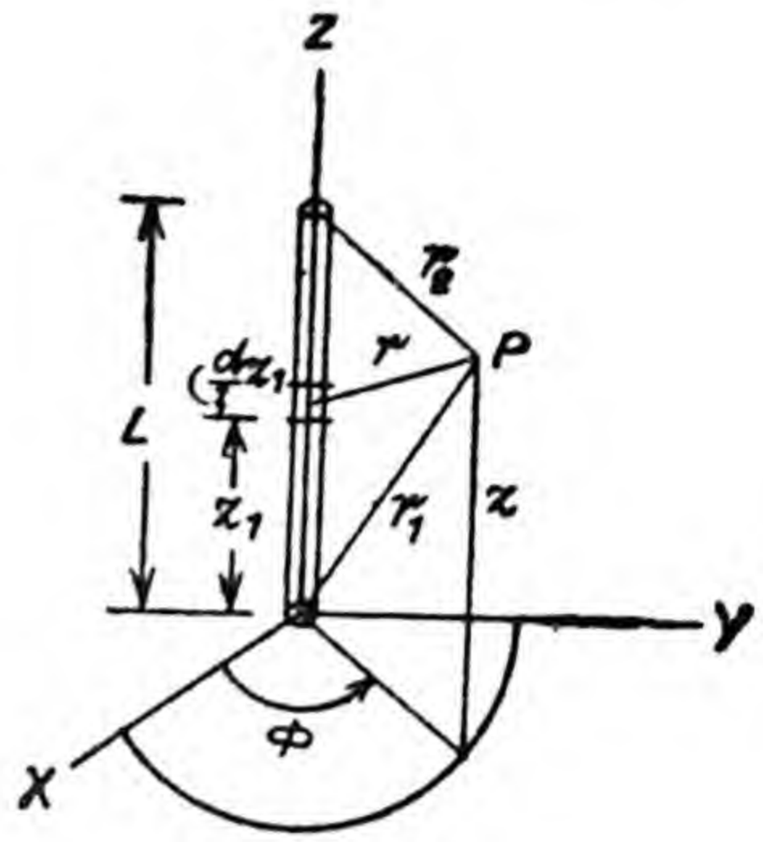


Fig. 53.

$$=\frac{j\mu I_1\, e^{j\omega t}}{8\pi}\int_0^L \frac{e^{-j\beta(z_1+r)}-e^{j\beta(z_1-r)}}{r}\,dz_1 \qquad \ldots(1)$$

The scalar potential V at any point is given by

$$V=\frac{1}{4\pi\epsilon}\iiint\frac{\rho}{r}\,dv$$

where ρ is the volume charge density. In the case of wire it can be reduced to

$$V=\frac{1}{4\pi\epsilon}\int_0^L \frac{\rho_L}{r}\,dz_1,$$

where ρ_L is the linear charge density on the wire. From continuity relation

$$\rho_L=-\int\frac{dI_{z_1}}{dz_1}\,dt$$

$$=\frac{jI_1\, e^{j\omega t}}{4\pi\epsilon c}\int_0^L \frac{\cos \beta z_1\, e^{-j\beta r}}{r}\,dz$$

in which we have put $\frac{\omega}{\beta}=c$. Therefore

$$\rho_L=\frac{jI_1\, e^{j\omega t}}{8\pi\epsilon c}\int_0^L \frac{e^{-j\beta(z_1+r)}-e^{j\beta(z_1-r)}}{r}\,dz_1. \qquad \ldots(17)$$

Putting equations (16) and (17) in equation (15), we get

$$E_z=-\frac{jI_1\, e^{j\omega t}}{8\pi\epsilon c}\int_0^L \frac{\partial}{\partial z}\left[\frac{e^{-j\beta(z_1+r)}+e^{j\beta(z_1-r)}}{r}\right]dz_1$$

$$+\frac{\omega I_1 e^{j\omega t}}{8\pi}\int_0^L\left[\frac{e^{-j\beta(z_1+r)}-e^{j\beta(z_1-r)}}{r}\right]dz_1$$

$$=-\frac{jI_1 e^{j\omega t}}{4\pi\epsilon c}\left(\frac{e^{-j\beta r_1}}{r_1}+\frac{e^{\pm j\beta r_2}}{r_2}\right) \qquad \ldots(18)$$

where $\left.\begin{aligned} r_1&=z\\ r_2&=L-z\end{aligned}\right\}$ at the antenna.

Putting these values of r_1 and r_2, we write finally for E_{11}, th field at the antenna due to its own current, as

$$E_{11}=-j30I_1\left[\frac{e^{-j\beta z}}{z}+\frac{e^{-j\beta(L-z)}}{L-z}\right], \qquad ...(19$$

in which we have put

$$\frac{1}{\pi\epsilon c}=\frac{120\pi}{4\pi}=30$$

and $e^{j\omega t}=1$, the absolute value.

Calculation of self impedance : We now put equation (19) int equation (14) to find the value of self impedance. Therefore,

$$Z_{11}=j30\int_0^L\left[\frac{e^{-j\beta z}}{z}+\frac{e^{-j\beta(L-z)}}{L-z}\right]\sin\beta z\,dz. \qquad ...(20$$

As already written, we have taken odd multiple of $\lambda/2$ as th length of the antenna. Therefore,

$$L=n\lambda/2, \qquad n=1, 3, 5...$$

and then $e^{-j\beta L}=e^{-j\pi n}=-1$,

so that equation (20) becomes

$$Z_{11}=j30\int_0^L\left[\frac{e^{-j\beta z}}{z}-\frac{e^{j\beta z}}{L-z}\right]\sin\beta z\,dz$$

$$=5\int_0^L\left[\frac{e^{-j\beta z}}{z}-\frac{e^{j\beta z}}{L-z}\right](e^{j\beta z}-e^{-j\beta z})\,dz$$

$$=-15\int_0^L\left(\frac{e^{-2j\beta z}-1}{z}+\frac{e^{2j\beta z}-1}{L-z}\right)dz$$

$$=15\int_0^L\frac{1-e^{-2j\beta z}}{z}\,dz+15\int_0^L\frac{1-e^{2j\beta z}}{L-z}\,dz. \qquad ...(21)$$

In the first integral let

$$u=2\beta z \quad \text{or} \quad du=2\beta\,dz$$

so that

$$=15\int_0^L\frac{1-e^{-2j\beta z}}{z}\,dz=15\int_0^{2\pi n}\frac{1-e^{-ju}}{u}\,du. \qquad ...(22)$$

In the second integral let

$$v=2\beta\,(L-z) \text{ or } dv=-2\beta\,dz$$

so that

$$15\int_0^L\frac{1-e^{2j\beta z}}{L-z}\,dz=-15\int_{2\pi n}^0\frac{1-e^{j(2\pi n-v)}}{v}\,dv$$

$$=15\int_0^{2\pi n}\frac{1-e^{-jv}}{v}\,dv. \qquad ...(23)$$

Since equations (22) and (23) are definite integrals of identical form and as their limits are the same, they are equal. Hence equation (21) becomes

$$Z_{11}=30\int_0^{2\pi n}\frac{1-e^{-ju}}{u}\,du.$$

Putting $\omega=ju$, we get

$$Z_{11}=30\int_0^{j2\pi n}\frac{1-e^{-\omega}}{\omega}\,d\omega, \qquad ..(24)$$

which is an exponential integral with imaginary argument. It is designated by $Ein(jy)$, where $y=2\pi n$. Thus

$$Ein(jy)=\int_0^{iy}\frac{1-e^{-\omega}}{\omega}\,d\omega$$
$$=Cin\,(y)+jSi\,(y),$$

where $Cin\,(y)$ and $Si\,(y)$ denote the cosine and sine integrals repectively

$$Ein\,(jy)=0\,577+\log_e y-Ci\,(y)+jSi\,(y).$$

or

Putting this value in equation (24), we get

$$Z_{11}=R_{11}+jX_{11}=30\,[Cin\,(2\pi n)+jSi\,(2\pi n)]$$
$$=30\,[0{\cdot}577+\log_e y-Ci\,(y)+jSi\,(2\pi n)]\text{ ohms.} \qquad ...(25)$$

Therefore the self resistance is

$$R_{11}=30\,Cin\,(2\pi n)$$
$$=30\,[0{\cdot}577+\log_e(2\pi n)-Ci\,(2\pi n)]\text{ ohms.} \qquad ...(26)$$

and the self reactance is

$$X_{11}=30\,Si\,(2\pi n)\text{ ohms.} \qquad ..(27)$$

Equations (26) and (27) are the final results desired. By applying these to any odd value of n, we can calculate self resistance and self reactance of that antenna length. For example, for one half wavelength antenna $(n=1)$,

$$R_{11}=30\,Cin\,(2\pi)$$

and

$$X_{11}=30\,Si\,(2\pi)$$

Evaluating these two integrals we can show that

$$Z_{11}=R_{11}+jX_{11}=73+j42{\cdot}5\text{ ohms.}$$

For the antenna to be resonant one X_{11} should be zero which can be achieved by shortening the antenna a few percent

For $n=3$, the $\frac{3}{2}$ wavelength antenna, the self impedance can be shown to be

$$Z_{11}=30\,[Cin\,(6\pi)+jSi\,(6\pi)]$$
$$=105{\cdot}5+j45{\cdot}5\text{ ohms.}$$

For large value of n, the self resistance expression (26) approaches the value

$$R_{11}=30\,[0{\cdot}577+\log_e(2\pi n)]$$

because $Ci\,(2\pi n)$ approaches zero. Thus, with increasing n, self resistance continues to increase at a logarithmic rate.

MUTUAL IMPEDANCE :

In coupled circuits mutual imdedance is defined as the negative ratio of the e.m.f. V_{21} induced in circuit 2 to the current I_1 flowing in circuit 1 with circuit 2 open. That is

$$Z_{21}=-\frac{V_{21}}{I_1}$$

We consider two parallel coupled antennas shown in fig. 54(*a*):

Suppose a current I_1 in antenna 1 induces an e.m.f. V_{21} at the open terminals of antenna 2. Then the ratio of $-V_{21}$ to I_1 is the mutual impedance Z_{21}. Thus

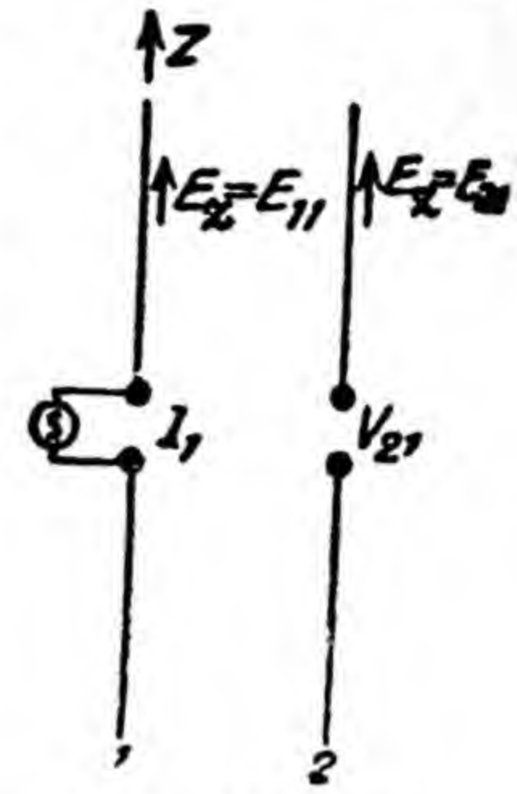


Fig. 54 (*a*)

$$Z_{21} = \frac{-V_{21}}{I_1} \qquad \ldots(28)$$

If the generator is connected to the antenna 2, then by reciprocity theorem

$$Z_{12} = \frac{-V_{12}}{I_2} \qquad \ldots(29)$$

Thus $$\frac{-V_{21}}{I_1} = Z_{21} = Z_{12} = \frac{-V_{12}}{I_2} \qquad \ldots(30)$$

For the calculation of Z_{21}, we require V_{21} and I_1. Refer to equation (11) in which $-V_{11}$ is the e.m.f. induced in an antenna by its own current. To find V_{21} the e.m.f. induced at the open terminals of antenna 2 by the current in antenna 1, we get

$$E_z = E_{21}$$

$$V_{11} = -V_{21}$$

and $$I_1 = I_2$$

in equation (11), so that

$$V_{21} = \frac{1}{I_2}\int_0^L I_z\, E_{21}\, dz$$

where I_2 is the maximum current and I_z the value at a distance z from the lower end of antenna 2 with its terminals closed, and where E_{21} is the electric field along antenna 2 produced by the current in antenna 1. Assuming that this current distribution is sinusoidal, *i.e.*

$$I_z = I_2 \sin \beta z,$$

we can write

$$V_{21} = \int_0^L E_{21} \sin \beta z\, dz$$

then

$$Z_{21} = \frac{-V_{21}}{I_1} = -\frac{1}{I_1}\int_0^L E_1 \sin \beta z\, dz$$

We shall use this expression to calculate the mutual impedance of parallel antenna placed side by side.

We have from figure 54 (*b*)

$$r_1 = \sqrt{(d^2 + z^2)}$$

and $$r_2 = \sqrt{\{d^2 + (L-z)^2\}}$$

so that mutual impedauce is

$$Z_{21} = j30 \int_0^L \left[\frac{e^{-j\beta\sqrt{[d^2+z^2]}}}{\sqrt{(d^2+z^2)}} + \frac{e^{-j\beta\sqrt{[d^2+(L-z)^2]}}}{\sqrt{[d^2+(L-z)^2]}} \right] \sin \beta z\, dz$$

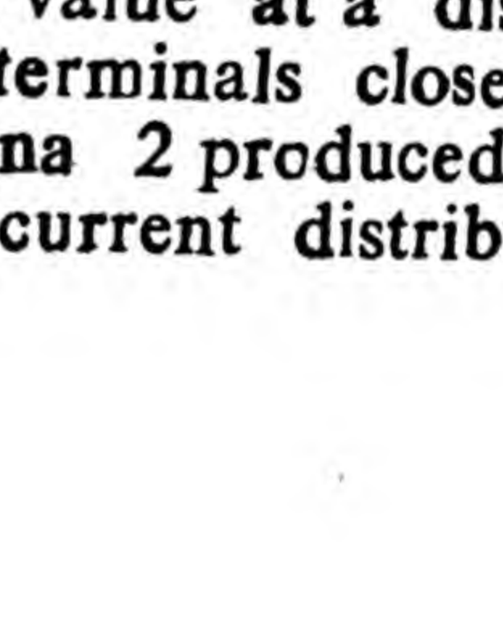


Fig. 54. (b)

Antenna

$$=30\,[2Ei\,(-j\beta d)-Ei\,\{-j\beta\,(\sqrt{(d^2+L^2)}+L)\}\cdot$$
$$-Ei\,\{-j\beta\,(\sqrt{(d^2+L^2)}-L)\}]$$

where the exponential integral

$$Ei\,(\pm jy)=Ci\,(y)\pm jSi\,(y)$$

Thus, the mutual resistance is

$$R_{21}=30\,[2Ci\,(\beta d)-Ci\,\{\beta\,(\sqrt{(d^2+L^2)}+L)\}$$
$$-Ci\,\{\beta\,(\sqrt{(d^2+L^2)}-L)\}]$$

and the mutual reactance is

$$X_{21}=-30\,[2Si\,(\beta d)-Si\,\{\beta(\sqrt{(d^2+L^2)}+L)\}$$
$$-Si\,\{\beta\,(\sqrt{(d^2+L^2)}-L)\}],$$

where

$$R_{21}+jX_{21}=Z_{21}=R_{12}+jX_{12}$$

Example. *A vertical half wave parasitic antenna is placed at a fixed distance in the vicinity of another similar antenna fed with 20 volts. The radiation resistance of half wave resonant antenna is 73 ohms. Mutual impedance for this spacing of the system is (50−j18) ohms. Calculate the radiation intensity in the horizontal plane for the pair along the direction of maximum radiation.*

If Z_{11} denotes the impedance of primary antenna, then the power radiated by first (primary) antenna

$$P_1=I_1^2Z_{11}+I_{11}\,(I_2Z_{12}),$$

where I_2 is the current flowing in the second antenna.

If Z_{22} denotes the impedance of the second antenna, then the power radiated by antenna 2 will be

$$P_2=I_1^2Z_{22}+I_2\,(I_1Z_{12})$$

So the total power radiated by the system is

$$P=P_1+P_2$$
$$=I_1^2Z_{11}+I_2^2Z_{22}+2I_1I_2Z_{12}. \qquad \ldots(1)$$

In the question, the matual impedance is given by

$$Z_{12}=50-j18$$

Say $Z_{12}=r\,e^{-i\theta}=r\cos\theta-ir\sin\theta$; then comparing with its given value we should have

$$r=\sqrt{(50^2+18^2)}=53\cdot1$$

and
$$\tan\theta=\frac{18}{50}.$$
$$\therefore\quad\theta=19\cdot8.$$

Thus
$$Z_{12}=53\cdot1\ \angle\ (-19\cdot8).$$

Since the phase difference in I_1 and I_2 is 90°, they can be represented as

$$I_1=I_2\ \angle\ 90$$

or
$$I_2=I_1\ \angle\ (-90).$$

Now we calculate the values of P_1 and P_2 from the relations deduced previously, *i.e.*

$$P_1 = Z_{11} I_1^2 + I_1 (I_2 Z_{12})$$
$$= 73 I_1^2 + I_1 . I_1 \angle (-90)\ 53{\cdot}1 \angle (-19{\cdot}8)$$
$$= I_1^2\ [73 + 53{\cdot}1 \cos 109\ 8].$$

We have considered only the real part because we are concerned with resistive impedance of the antenna system. Thus

$$P_1 = I_1^2\ [\ 3 + 53{\cdot}1 \times (-{\cdot}33874)]$$
$$= 54{\cdot}95 I_1^2$$

Similarly, $P_2 = I_2^2\ [73 + 53{\cdot}1 \angle (90 - 19{\cdot}8)]$
$$= I_2^2\ [73 + 53{\cdot}1 \cos 70{\cdot}2]$$
$$= 91{\cdot}05 I_2^2.$$

Thus total power radiated by the coupled system is

$$P = P_1 + P_2$$
$$= (54{\cdot}95 + 91{\cdot}05)\ I_1^2 \text{ if } |\ I_1\ | = |\ I_2\ |$$
$$= 146 I_1^2$$
$$= 146 \left(\frac{E_1}{R_1}\right)^2$$
$$= 146 \times \frac{20 \times 20}{73 \times 73} \text{ watts}$$
$$= \mathbf{10{\cdot}95\ watts.}$$

28·21. ANTENNA MEASUREMENTS :

In this article we shall discuss the measurements of
(*i*) total resistance of aerial,
(*ii*) radiation resistance and
(*iii*) effective height.

(*i*) **Total aerial resistance measurements :** The methods, generally adopted for the measurement of total aerial resistance, are given below :

(*a*) resistance variation method,
(*b*) substitution method, and
(*c*) ammeter and voltmeter method.

We shall describe the first two methods :

(*a*) *Resistance Variation Method* : An equivalent $\lambda/4$ aerial is shown in figure 55 (*a*). A thermo-ammeter *A* and a known resistance R_1 which must be non-inductive and non-capacitive are connected to earth lead of the aerial.

By means of switch *S*, R_1 can be introduced into the circuit or shunted off. Carrier wave oscillator is loosely coupled to a small coil *L*. We proceed as follows :

(*i*) First short-circuit resistance R_1 and tune the aeria l

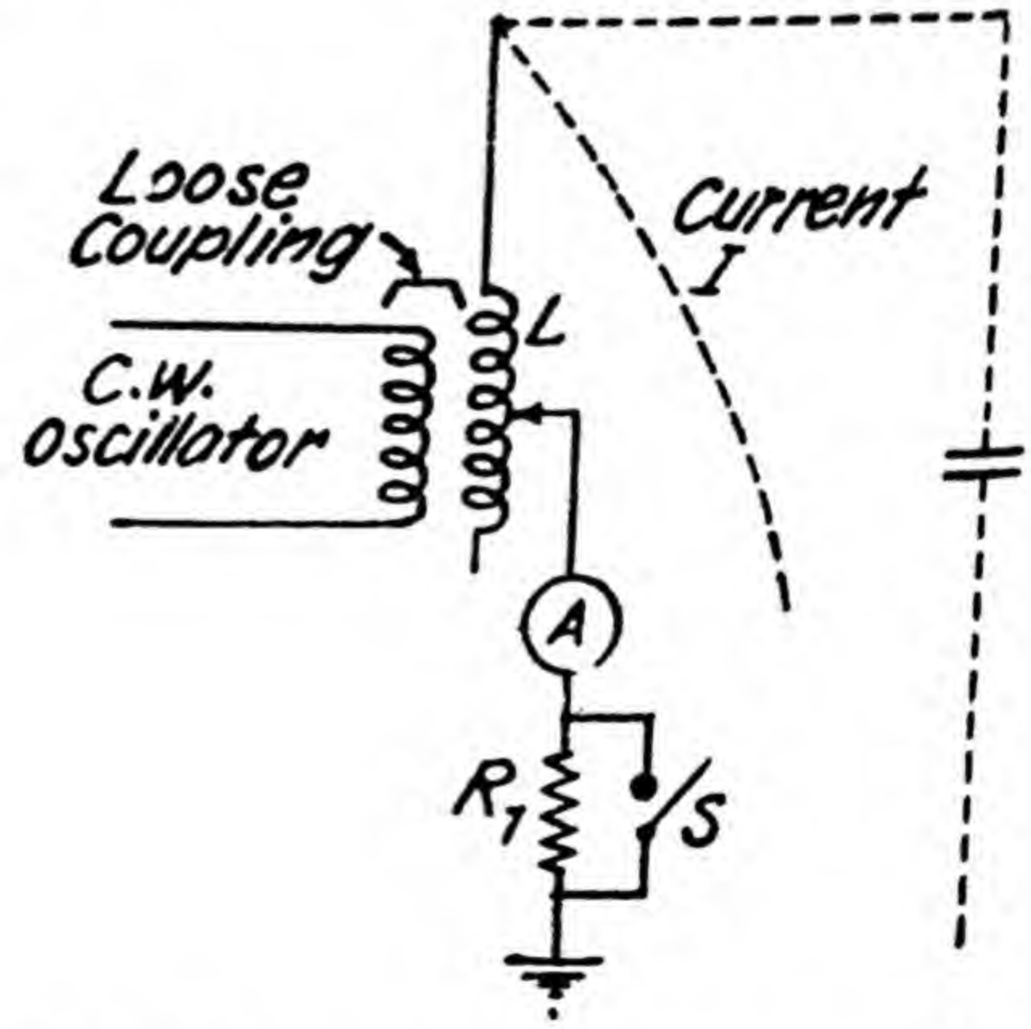


Fig. 55 (*a*). Measurement of aerial resistance.

to resonance. Let the aerial current in this case be I_1 and total effective resistance of aerial be R_e, then

$$I_1 = \frac{e}{R_e},$$

where e is the voltage induced in the coil.

(ii) Now switch R_1 into the circuit, so that aerial current falls to some other value, say, I_2. Since the aerial is resonant, we may write

$$I_2 = \frac{e}{R_1 + R_e}.$$

Therefore from first and second relations, we have

$$\frac{I_1}{I_2} = \frac{R_1 + R_e}{R_e},$$

R_1 being known resistance, R_e can be calculated.

(b) Substitution method : Fig. 55 *(b)* represents an equivalent $\lambda/4$ aerial.

In this method, an artificial aerial is substituted in place of an actual aerial with the help of a change over switch. When switch is changed in actual aerial direction, then the latter has its circuit completed to earth through ammeter A and coupling coil. Secondly, if the switch is changed over in other direction, then substitute aerial comes into the circuit.

The procedure is to obtain the same current value in ammeter 'A' with actual and artificial aerials separately with a given input from carrier wave oscillator. First with a given input, the current is noted for the actual aerial being in the circuit. Now switch is changed over so as to introduce the substitution aerial. The constants of this circuit are adjusted to give the same value of current. If artificial aerial components are calibrated, aerial resistance and aerial capacity can be reod directly. It is necessary that the resistance of artificial circuit should be non-inductive and non-capacitive and capacity itself should be pure.

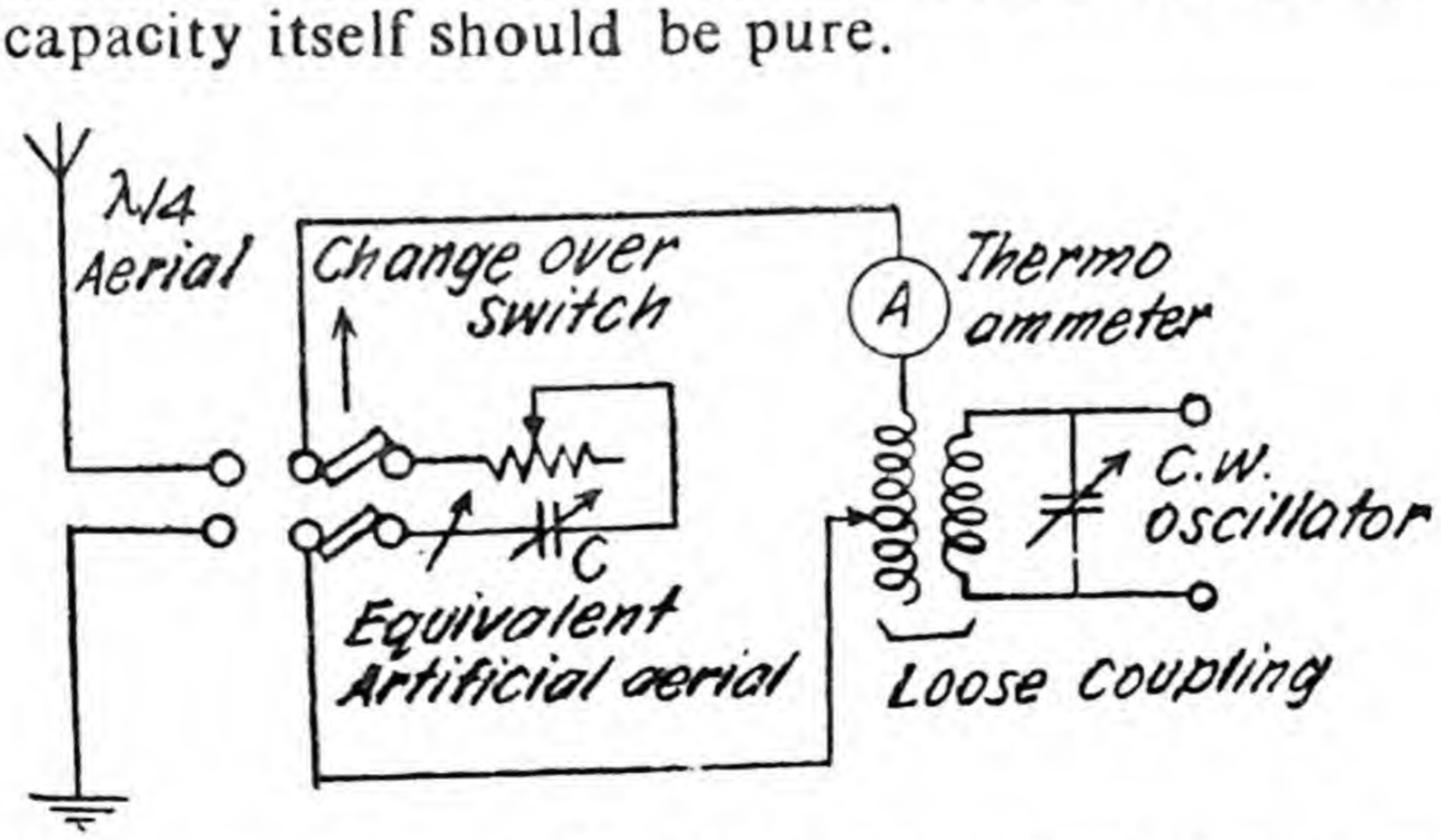


Fig. 55 (*b*) Measurement of aerial resistance (substitutio method).

(ii) Measurement of Radiation Resistance : The radiation resistance of test antenna (resonant) is measured in terms of some standard antenna constants by employing a reflection method.

Consider that a half-wavelength resonant antenna 1 under test is placed in the field of transmitting antenna. The reflected power, received by another antenna is measured. The power radiated by the antenna under test is

$$P_1 = K.\frac{(E_t.l_{e1})^2}{R_1}, \qquad ...(1)$$

where E_t = field of transmitting antenna at the antenna 1 under test,

l_{e1} = effective length of resonant test antenna 1,

R_1 = radiation resistance of resonant test antenna 1,

K = constant involving the distance between antennas.

Now we take some standard half-wavelength resonant antenna 2 (*i.e.*, whose radiation resistance R_2 is known) of different thick-

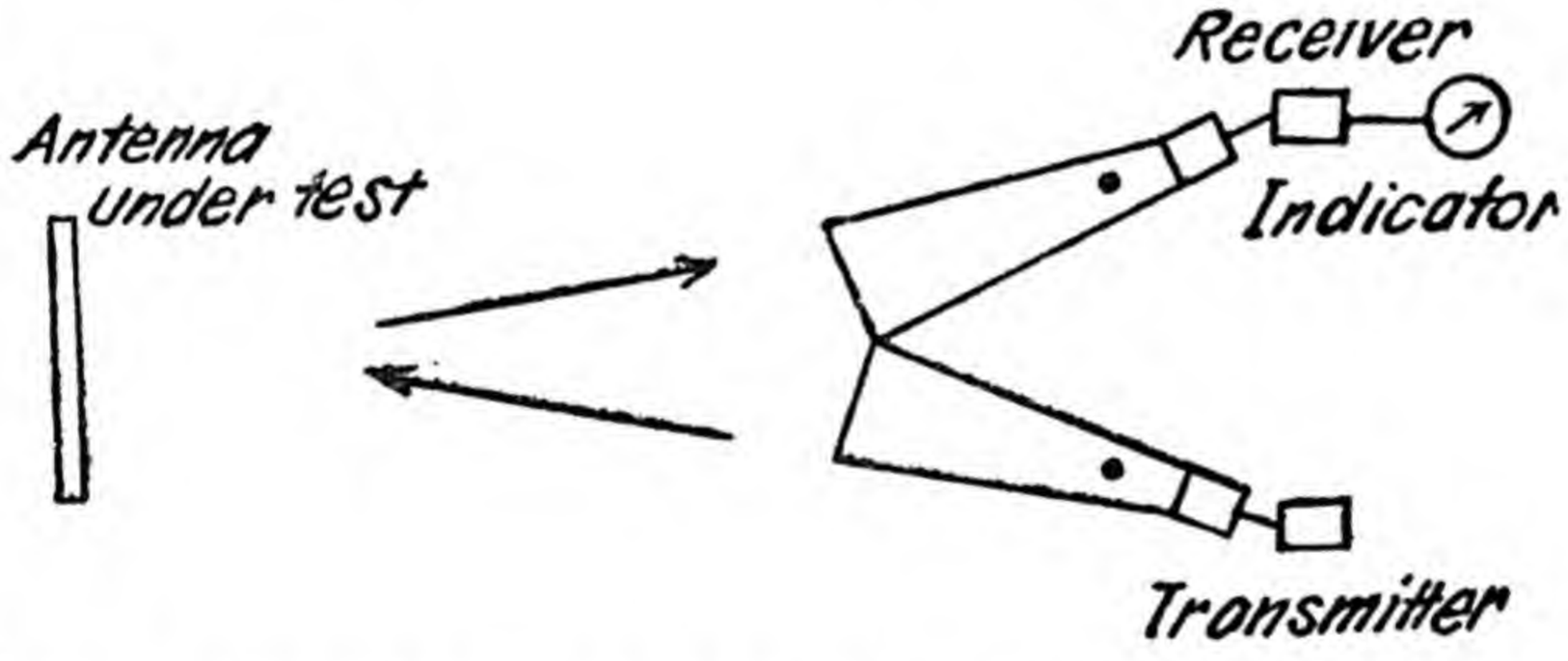


Fig. 55 (c) Measurement of radiation resistance by reflection method.

ness and hence different effective length l_{e2}, but substantially of the same pattern. The reflected power by this antenna 2 will be

$$P_2 = K.\frac{(E_t.l_{e2})^2}{R_2}. \qquad ..(2)$$

where l_{e2} = effective length of antenna 2,

R_2 = radiation resistance of antenna 2,

K = same constant as in previous case with test antenna.

Thus from equations (1) and (2), we get

$$\frac{R_2}{R_1} = \frac{P_1}{P_2}\left(\frac{l_{e2}}{l_{e1}}\right)^2 \qquad ..(3)$$

Assuming the small changes in length, the effective length is proportional to the physical length l_1 of antenna 1 and l_2 of antenna 2, we write

$$\frac{R_2}{R_1} = \frac{P_1}{P_2}\left(\frac{l_2}{l_1}\right)^2. \qquad ...(4)$$

The ratio $\frac{P_1}{P_2}$ can be measured in terms of the ratio $\left(\frac{V_1}{V_2}\right)^2$,

where V_1=voltage received when test antenna 1 is observed.
V_2=voltage received when standard antenna 2 is observed.

Thus $$R_1 = R_2 \cdot \frac{P_2}{P_1}\left(\frac{l_1}{l_2}\right)^2$$

The right-hand side being known, the radiation resistance R_1 of antenna 1 can be calcula ed.

(iii) Effective Height Determination : In equation (3) of art. 28 9 it has been shown that

$$\mathcal{E} = \frac{377 I_{r.m.s.}\, l_e}{\lambda r} \text{ volt/metre.}$$

From this relation effective height can be calculated since

$$l_e = \frac{\mathcal{E}\lambda r}{377 I_{r.m.s.}} \qquad \ldots(5)$$

In order to calculate effective height, we measure the field strength $\mathcal{E}$ at a point distant 5λ from the transmitter. The accuracy of effective height determination will thus depend upon the accuracy with which field-strength is measured. Field strength measurement is usually accomplished by means of loop aerial. The signals are received on the loop aerial tuned by a condenser and having a thermo milliammeter as shown in fig. 55 (*d*).

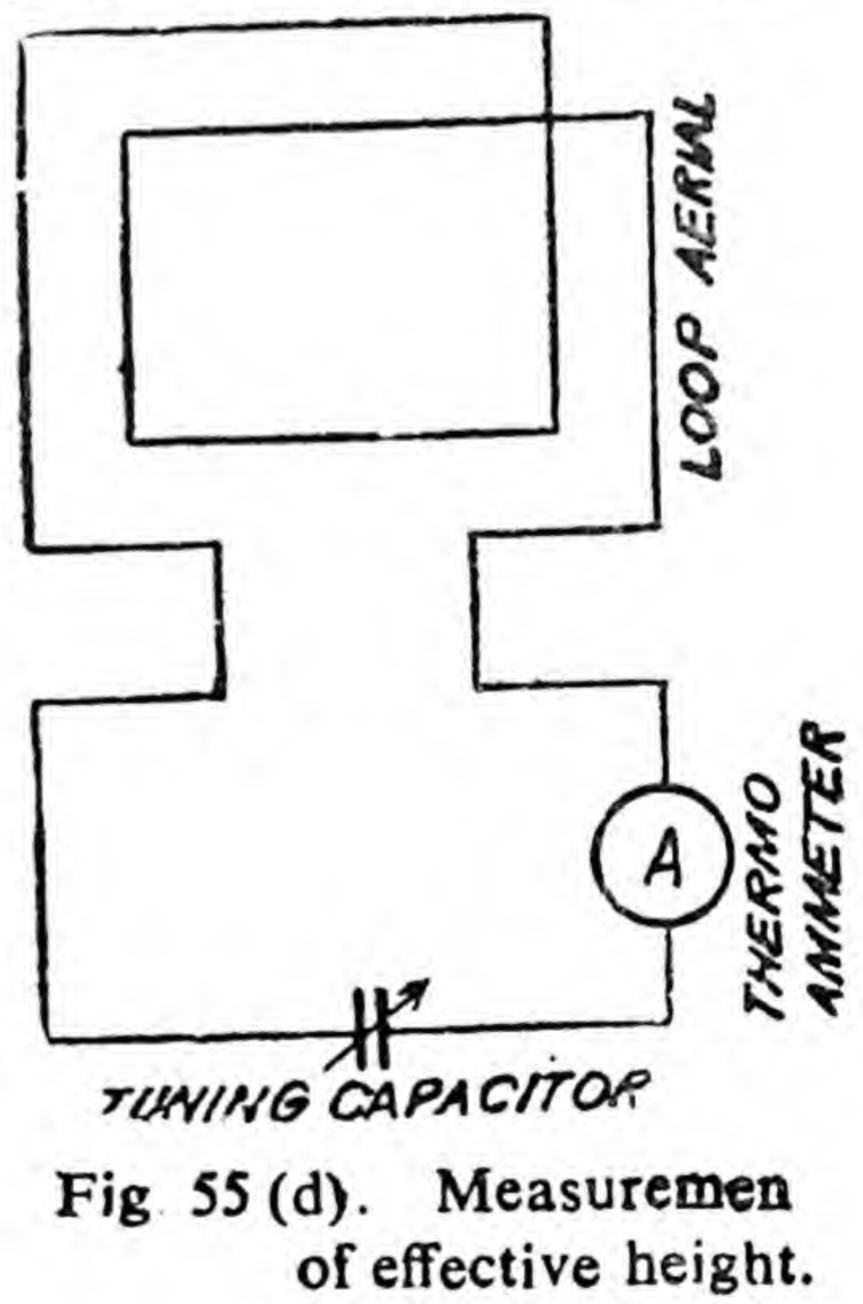


Fig. 55 (d). Measuremen of effective height.

It can be shown that e.m.f induced in the loop aerial is given by

$$\frac{2\pi A\mathcal{E}}{\lambda},$$

wnere A is area of the loop.

This loop e. m. f. will produce a resonant current I determined by the resistance of the circuit and is given by

$$e = IR.$$

Therefore, $$e = IR = \frac{2\pi\mathcal{E}A}{\lambda}. \qquad \ldots(6)$$

Thus by measuring either I if $\mathcal{E}$ is large or induced voltage e if $\mathcal{E}$ is small, value of field strength can be determined. Then by using relation (5) effective height is calculated

28 22. BROAD-BAND ANTENNA PRINCIPLES :

First we shall explain how the input impedance and pattern properties of an antenna change with the change in frequency. Conventional antennas are essentially resonant structures Their properties exhibit maxima and minima at particular frequencies.

For example, we consider centre fed half wave antenna. At the frequency at which its length is approximately (accounting for end effects) a half wavelength, the reactive component of input impedance is minimum (zero) [that is, input impedance is purely resistive]. This frequency is called rosonant frequency. Below this frequency antenna impedance becomes capacitive and above it, impedance becomes inductive [Refer to impedance curves of open ended transmission line remembering that a half wavelength antenna corresponds to a quarter wave open circuited transmission line] At even multiple of this frequency input reactance is maximum, and minimum at an odd multiple of this frequency. Thus with frequency variation, input impedance goes through maximum and minimum values. Similarly, the radiation which is maximum in the direction perpendicular to the dipole when length is $\lambda/2$ also goes through successive maxima and minima when the frequency is increased Since this pattern property does not exhibit same nature of variation instep with the variation in input impedance with the increase in frequency, we shall consider pattern and impedance bandwidths separately.

At the lower frequencies *ELF* (below 3 kc/s), *VLF* (3-30 kc/s) and *LF* (30–300 kc/s), efficiency is the main requirement. Although antennas in these frequency ranges are quite large physically, they fall in the class of electrically small antennas because of the large dimensions of a wavelength. At *MF* (300 Kc/s – 3 Mc/s) which include 'broadcast band', tower antennas are frequently employed in case where efficiency is the prime factor, and arrays of towers are employed if directivity is desired. For reception in this band efficiency is unimportant and relatively short wire antennas suffice. At *HF* (3 – 30 Mc/s), *VHF* (30 – 300 Mc/s) and *UHF* (300 Mc/s–3Gc/s) elevated wires and rods are used and arrays of such antennas are designed to provide more directivity. *At such frequencies, the operation over a large bandwidth is often required* (say from 54 – 890 Mc/s), *the due consideration is to be given while designing the antenna or array, to maintain desired impedance and pattern characteristics over a wide band of frequencies.* At *SHF* (3–30 Gc/s), aperature antennas are employed. At such frequencies the antenna design considerations include the production of a narrow beamwidth with low side-lobe level, and maintenance of the desired radiation pattern characteristics over a wide range of frequencies.

In television transmitting systems it is necessary that the impedance match between the antenna and the transmission line leading from the transmitter be good enough to keep the reflection coefficient to within a very small value throughout the frequency band covered by the television signal. This bandwith is of the order of 5 mc/s in typical systems, and with the assigned television channels, is of the order of 10 to 15 percent of the carrier frequency. The maximum allowed reflection coefficient that can be tolerated is of the order of 5 percent, which means that the vector impedance must not vary

more than ± 10 percent from the mean value. Hence when the magnitude of the impedance is substantially constant over the band, the reactive component of the impedance must never exceed 10% of the resistive componnet.

Band width of an antenna : It is the range of frequencies over which the antenna maintains certain required impe dance, pattern characteristics. If an antenna were capable of operating satisfactorily from a minimum frequency of 195 mc/s to a maximum frequency of 205 mc/s, its bandwidth would be 10 mc/s. It would also be said to have 5 percent bandwidth defined as

$$\frac{\text{actual bandwidth in mc/s}}{\text{centre frequency of the band}} \times 100.$$

5 percent bandwidth may be considered a moderate bandwid h

The design of an antenna which is required to operate only at a fixed frequency with a signal that is narrow in its bandwidth does not pose any bandwidth problem But some applications require much greater bandwidths. In the case of frequency modulated waves, width of frequency spectrum that contains all components having an amplitude equal to or greater than one percent of the amplitude of unmodulated wave is defined as bandwidth Suppose a given wave has n significant side frequency pairs then bandwidth is equal to $BW = 2n\, f_n$.

Television signals have a bandwidth equal to 4 mc/s and therefore a television intermediate frequency amplifier must have a band width of approximately 4 mc/s inorder to pass all frequency components of a television signal. By the same reason. a television transmitting antenna must also have atleast this bandwidth. A television receiving antenna, on the otherhand, must have sufficient bandwidth to receive all the channels to which the receiving set can be tuned.

The bandwidth is classified into two categories according to the parameters involved. They are

(*a*) pattern bandwidth

(*b*) impedance bandwidth.

In pattern bandwidth, antenna would maintain satisfactorily the parameters-beamwidth, gain, side-lobe level, beam direction, and polorization over that specified range of frequencies (called band width); whereas in impedance bandwidth, associated parameters are input impedance, radiation resistance and efficiency that are satisfactorily maintained by the antenna while operating in that specified range of frequencies.

The definition of the antenna bandwidth is less precise than the definitions of other parameters due to the fact that there is no established criterion of 'satisfactory' operation. In some applications, for example, an impedance variation of a factor of 2 over

the operating frequency band may be acceptable. whereas in other applications only a 10 percent variation may be tolerated.

Basic Broad-Band Principles : If a resonant circuit is 'loaded' by connecting or coupling a dissipative element (*e.g.*, a resistor) to it, then

$$Q = \frac{1}{R}\sqrt{\left(\frac{L}{C}\right)}$$

would decrease. Decrease in Q implies increase in bandwidth *i.e.* sharpness of resonance, amplitude of variation of impedance and other properties will be reduced. This method of loading, however, decreases antenna efficiency. It is used in terminated rhombic and *V* antennas. However, a resistive loading is unnecessary when ground losses and other conductor losses reduce the Q sufficiently.

At higher frequencies, conductor losses being low, simple antennas do not have adequate bandwidth. But at these frequencies, deliberate loading with resistance is seldom used. The methods employed to improve bandwidth are divided into two classes : those that apply to basic radiators, and those that apply to complex antennas such as arrays, reflector-feed combinations and lens antennas. The methods that apply to basic radiators, such as dipoles, are mostly geometric in nature. That is, bandwidth is increased by using certain shapes and certain size ratios. This method has helped in the development of the log-periodic antennas, or frequency independent antennas. Other methods apply to complex antennas. Sometimes it is possible to combine elements whose variations of impedance or other properties are complementary in such a way that they tend to offset each other, as the frequency is increased. We shall, for example, discuss here that how increased bandwidth is achieved by a folded dipole.

Folded Dipole : A folded dipole is preferred over a simple dipole due to the following two advantages :

(*i*) it provides high input impedance for best matching with the feed which may be a two wire transmission line, and

(*ii*) the bandwidth is increased.

We shall now explain how the above characteristics are achieved :

(*i*) *High Input impedance* : A simple dipole has an input impedance that is too low for direct connection to an ordinary open wire transmission line. For example, a simple half wave dipole will provide 73 ohms as the input resistance and the transmission lines have characteristic impedance of the order of not less than 300 ohms. Therefore, for best matching, the line should be terminated by an impedance equal to 300 ohms otherwise reflections on the line would occur. Therefore, with simple dipole, some

sort of impedance matching arrangement is required if the desirable condition of no standing wave on the line is required. However, with folded dipole no such arrangement is required because it offers a high input impedance.

A tolded dipole is shown in the fig. 56. It consists of two half wave radiators very close to each other and connected together at top and bottom. From the point of view of antenna currents or radiating currents, the two elements are in parallel and if their dia neters are the same, the currents in the elements will be equal and in the same direction. If I amp. current flows in each element (at the centre) the total effective current is $2I$ amp , so that the power radiated will be

$$(2I)^2 R_{rad} \approx 4 \times 73 I^2.$$

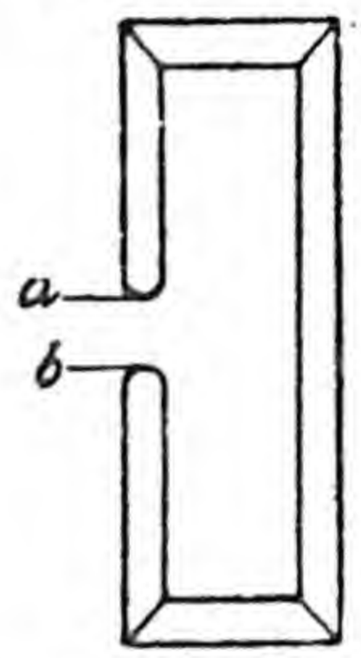


Fig. 56. A folded dipole.

Therefore input resi tance will be 4×73 ohms. That is 4 times that of a simple dipole.

(*ii*) *Increased Bandwidth* : In fig. 57 (*a*) a simple half wave dipole connected at its input terminals to a shorted quarter wave transmission line is shown. We shall explain how this arrangement maintains the input resistance at its value when the frequency is changed above or below the resonance. That is how such an arrange ment reduces the input impedance variations with the variation in frequency or in other words impedance bandwidth is increased.

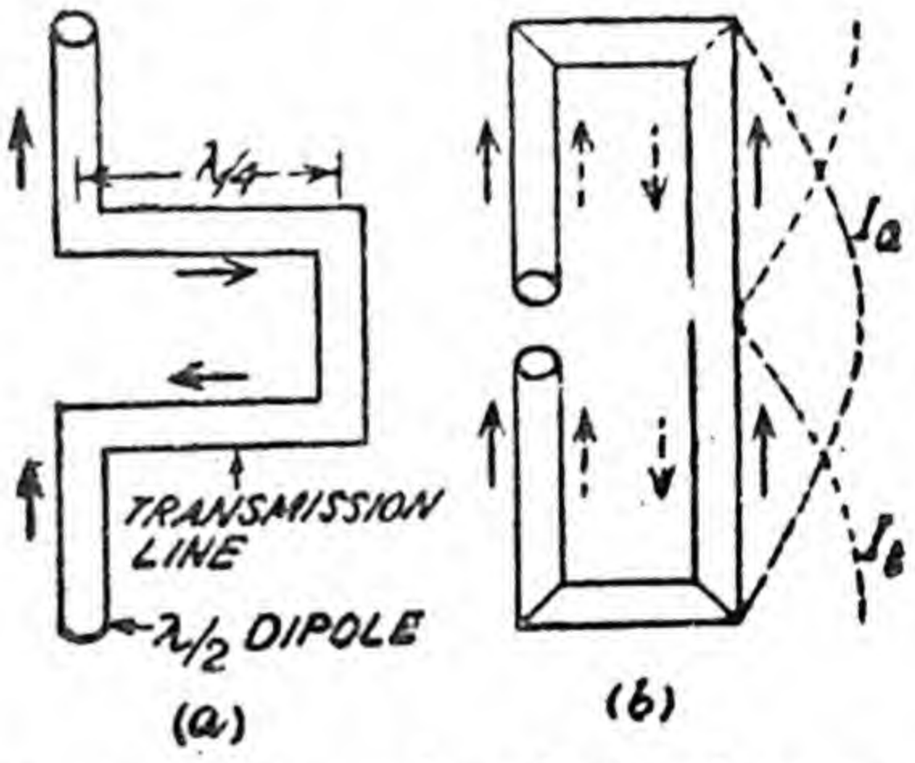


Fig. 57 (*a*) Simple λ/2 dip le connected at this input to a shorted quarter wave transmission line.

(*b*) a folded dipole
I_a—antenna current
I_t—transmission line current.

Let us first remember that for a short circuited transmission λ/4 line, below resonance, input impedance becomes inductive and above resonance, input impedance becomes capacitive. But for λ/2 dipole which is equivalent to a open circuited λ/4 line, this impedance variation is in reverse order that is, below resonance its input impedance will become capacitive whereas above resonance it will be inductive.

Now we discuss the arrangement shown in fig. 57 (*a*) that how the antenna impedance and transmission line impedance compensate the variations of each other frequency in order to leave input impedance of the system at resonant frequency unaffected. At the resonant frequency the dipole resistance is in parallel with the input impedance of the transmission line, which is a resistance of very

high value.* Below resonance, the antenna impedance becomes capacitive, but the transmission line impedance becomes inductive. and the parallel combination tends to remain nearer unity power factor than does the antenna alone. Conversely, above resonance the antenna impedance becomes inductive and the transmission line impedance becomes capacitive so that compensation is again obtained. Although compensation is far from perfect, because the susceptances are not equal and opposite, yet the frequency is shifted far enough in either direction from the resonant frequency of the dipole, a point of perfect susceptance compensation (where the input impedance is a pure resistance) is again obtained.

The above considerations apply directly to the folded dipole of fig. 57 (*b*) which has the stub line as a built in feature. The elements of the folded dipole then carry both, the antenna currents, which are in the same direction in the two elements, and the transmission line currents, which are in opposite directions in the two elements of dipole. The antenna currents

$$I_a = \frac{V}{R_{in}} = \frac{V}{4R_{rad}}$$

in each element at the centre are relatively large at the resonant frequency whereas transmission line currents are zero at the centre, but have the value $I_t = \frac{V}{2Z_{0t}}$ at the ends, where Z_{0t} is the characteristic impedance of each of the two shorted $\lambda/4$ transmission line sections.

Thus with folded dipole, input impedance is maintained substaintially constant over a large band of frequencies. This leads to the better matching between the transmission line and antenna over the wide range of frequencies.

Multi-element Folded Dipole :

The additional step up of the simple dipole impedance may be obtain d by adding more "folded" conductors. A three wire folded dipole is shown in fig. 58. Its input impedance is approximately 9 times that of a single dipole which is 73 ohms. The step up ratio may be altered by making the folded conductors of different size from the basic conductor.

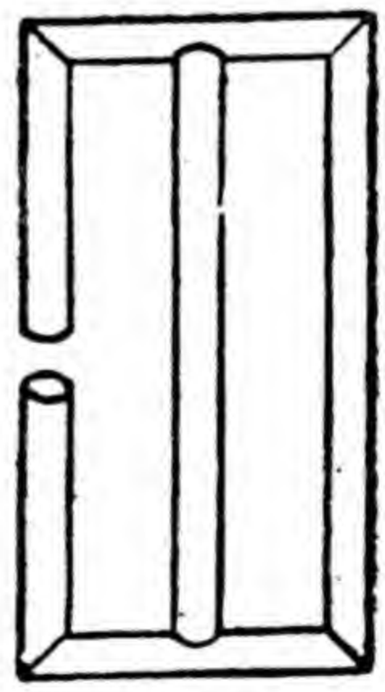


Fig. 58. A three wire folded dipole

A typical folded dipole for *TV* reception is made of aluminium or plated steel

*The input impedance of a short circuited line is given by

$$Z_S = jZ_0 \tan \beta s.$$

For $s = \lambda/4$,

$$Z_S = \infty,$$

that is very high.

tubing $\frac{1}{4}$ to $\frac{1}{2}$ inch in diameter. Since the dipole length is perhaps 2 or 4 feet, tubing of this size provides necessary stiffness and strength to make it self supporting.

Even the folded dipoles of the type discussed above are inadequate to meet the rigid requirements of a high grade television transmitter with a 50 mc/s carrier. By combining two folded dipoles in a turnstile array, still further improvement is possible.

Superturnstile Antenna : A broad band dipole frequently used as the element of a *TV* transmitting array, known as the superturnstile, is shown in the fig. 59. This element is partly a dipole and partly a slotted sheet antenna. The slot is vertical and the dipole horizontal ; hence the polarization is horizontal. It is fed by a balanced line at the centre of the slot. The input impedance is about 75 ohms. This radiator has a 30 to 35% impedance bandwidth for a voltage standing wave ratio not exceeding 1·1 : 1. Its broad band properties stem partly from large dipole width and partly from the fact that the dipole and the slot have complementary impedance variation characteristics that tend to offset each other as the frequency is varied.

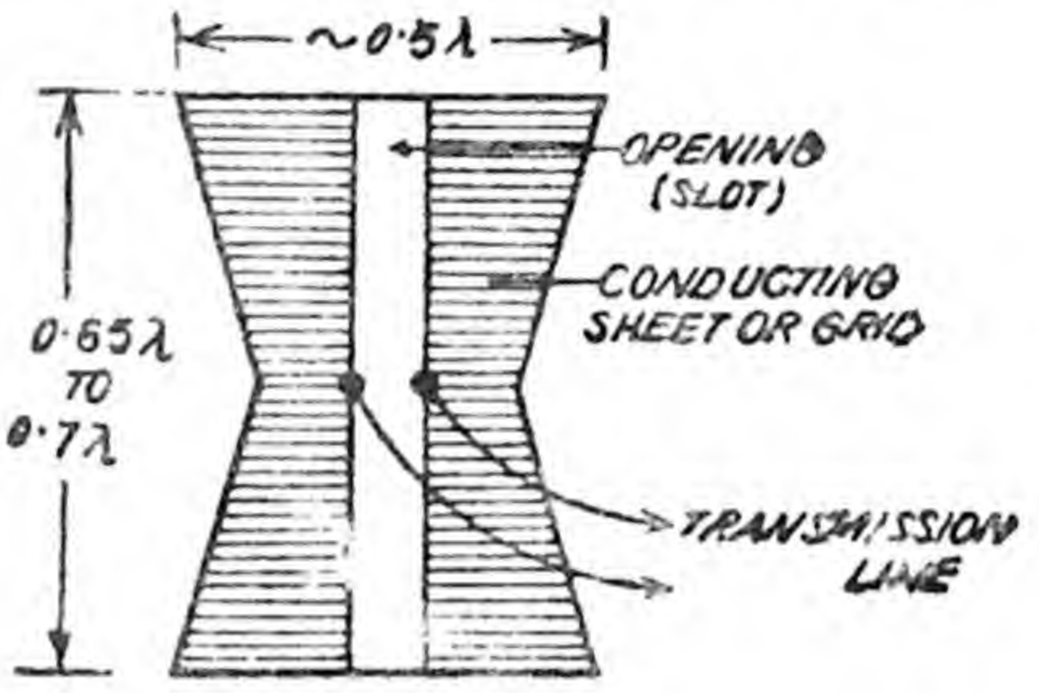


Fig 59.

28·23. PRACTICAL TRANSMITTING ANTENNAS :

(1) **Transmitting antennas for frequencies below 500 kc.** (*Low frequency range*) : For the transmission of low radio frequencies grounded vertical wire type antennas are used. Such an antenna consists of flat top, usually arranged in the form of an inverted *L* or a *T*, and sufficiently large so that its capacity to ground is much greater as compared to the capacity between vertical wire and ground. Due to the latter cause, the effective height and the actual height of the transmitting antenna are nearly equal. This, on the other hand, increases the radiation efficiency.

At low frequencies radiation resistance is usually very low because the antenna height is only a small fraction of wavelength. Due to this reason power radiated by low radio frequency antenna is also low and therefore to enable the antenna to radiate even small fraction of the energy supplied, losses should be reduced to minimum. The impedance losses occur in the ground, in the antenna tuning inductance, in the insulation of the conductors.

For the elimination of night effect, vertical tower radiators are used at such frequency ranges.

The Adcock type uses base insulated and self-supporting towers consisting of four radiators with height 40 metres and spacing 180 metres on a diagonal.

(2) Transmitting antennas for frequencies 550-1500 kc. (*Mid frequency range*): The antenna used for such frequency ranges consists of a tower that functions as a vertical radiator. In some cases it is not feasible to construct a simple radiator of the required height because of some commercial air line passing nearby. For such cases, tower is top loaded which then gives the same field distribution in the vertical plane as is obtained with a simple vertical radiator of somewhat greater vertical height. Top-loading is usually in the form of a ring, spider or other equivalent arrangement providing a lumped capacity at the top of the tower. Since the amount of top loading can be modified, the equivalent electrical height of a top-loaded antenna is subject to change accordingly. This factor shows that flexibility also demands the top-loading.

The height of this vertical radiator is made either $\lambda/4$ or greater than $\lambda/2$ For the antenna of height $\lambda/4$, radiated field is nearly proportional to the cosine of the angle of elevation. With antenna heights of the order of $\lambda/2$ or slightly greater, energy is much concentrated along the horizontal plane and consequently at high vertical angles, radiation is reduced. A height of $0{\cdot}25\lambda$ gives the maximum concentration of radiation along the horizontal but at such a height undesirable large lobe radiation appears at high angles and hence does not suit much. The optimum value of height depends upon the

(*i*) transmitter power.

(*ii*) transmitter frequency and

(*iii*) earth connections.

Radiation efficiency of a tower antenna is 80% or 90% when its length is atleast $\lambda/4$ and a ground system is employed.

There are, however, two ways of exciting the tower antenna :

(*a*) *Series excitation* : In such a type of excitation, radio frequency voltage is applied between the lower end of the tower and ground.

(*b*) *Shunt excitation* : In this type of excitation, the tower

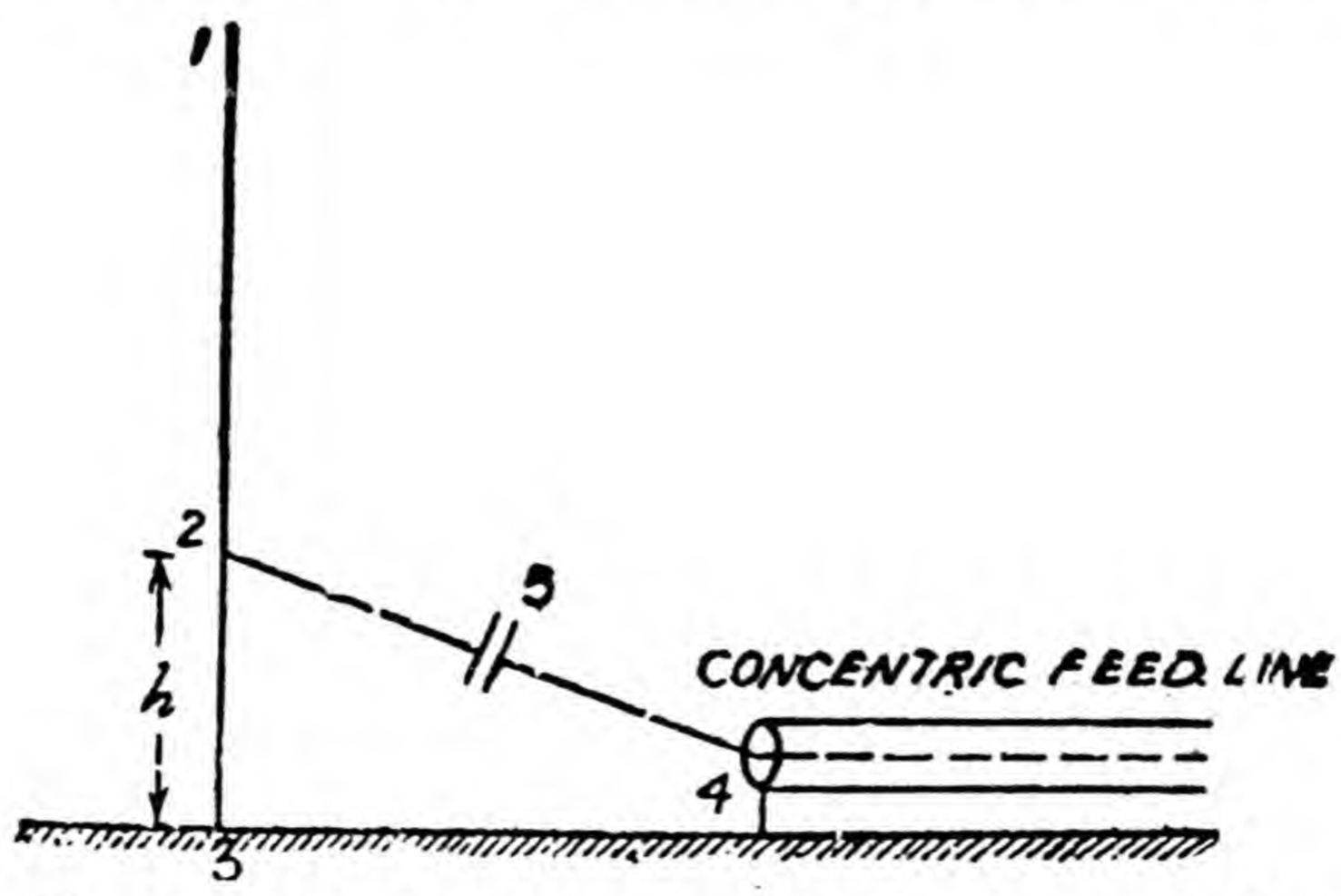


Fig. 60. Shunt-excited grounded vertical radiator.

may be grounded and the exciting voltage is applied across a section *h* of the tower. In fig. 60 a shunt-excited grounded vertical radiator in shown

In such a case, loop 23452 carries a current which develops a voltage across section *h*. This voltage is then applied to the rest of the tower. The reactance of the loop is tuned by means of condenser 5.

Losses : When the ground system is employed, the important losses occur due to ground. Currents flow from tower to the ground through the capacity and then return through the earth to the base of the tower. Because of skin effect, current is largely confined to the region near the earth's surface and since the earth is a bad conductor, appreciable energy losses occur. To minimise losses, a low resistance path should be provided to the current to return to the tower. For this purpose, a system of buried wires radiating out from the base of the tower is employed.

Antenna Arrays : For broadcast purposes, simple antenna arrays consisting of three to four spaced tower radiators are used extensively. Such arrays serve two principal purpose in broadcast work :

(*a*) because of its directional property, it helps to concentrate the radiated energy in a desired direction and thus gives better signals, and

(*b*) since a directional array can reduce the field strength radiated in a particular direction to minimum, interference between the signals of two transmitters that are separated by a small distance and operating on the same channel can be avoided.

The directivity of the array depends upon the spacing and the relative magnitude and phase of currents in the various towers.

(3) Transmitting antennas for frequencies 15-30 mc/s. (*High frequency range*) :

Generally two arrangements, depending upon whether or not the directivity of radiation is required, are used :

(*a*) *Half wave antenna arrangement* : Ungrounded half-wavelength resonant wire is usually employed as an antenna if directivity is not the chief consideration. Such a half wave antenna may be placed vertically or horizontally. When it is placed with its axis vertical, maximum field is radiated along horizon but when placed with its axis parallel to ground, field radiated in the horizontal direction is zero. The latter placement is most common.

Some ways in which a halfwave resonant antenna may be excited delivering it the power from a balanced two wire transmission line are shown in fig. 61

In method (*a*) it is not essential that antenna length should approximate a half-wavelength. A change in antenna length simply affects the resonance transmission feed

line. This arrangement fails to provide characteristic impedance termination.

In method (*b*) the capacity shown in figure completes the termination of transmission line. This method, however, fails to provide impeda nce matching.

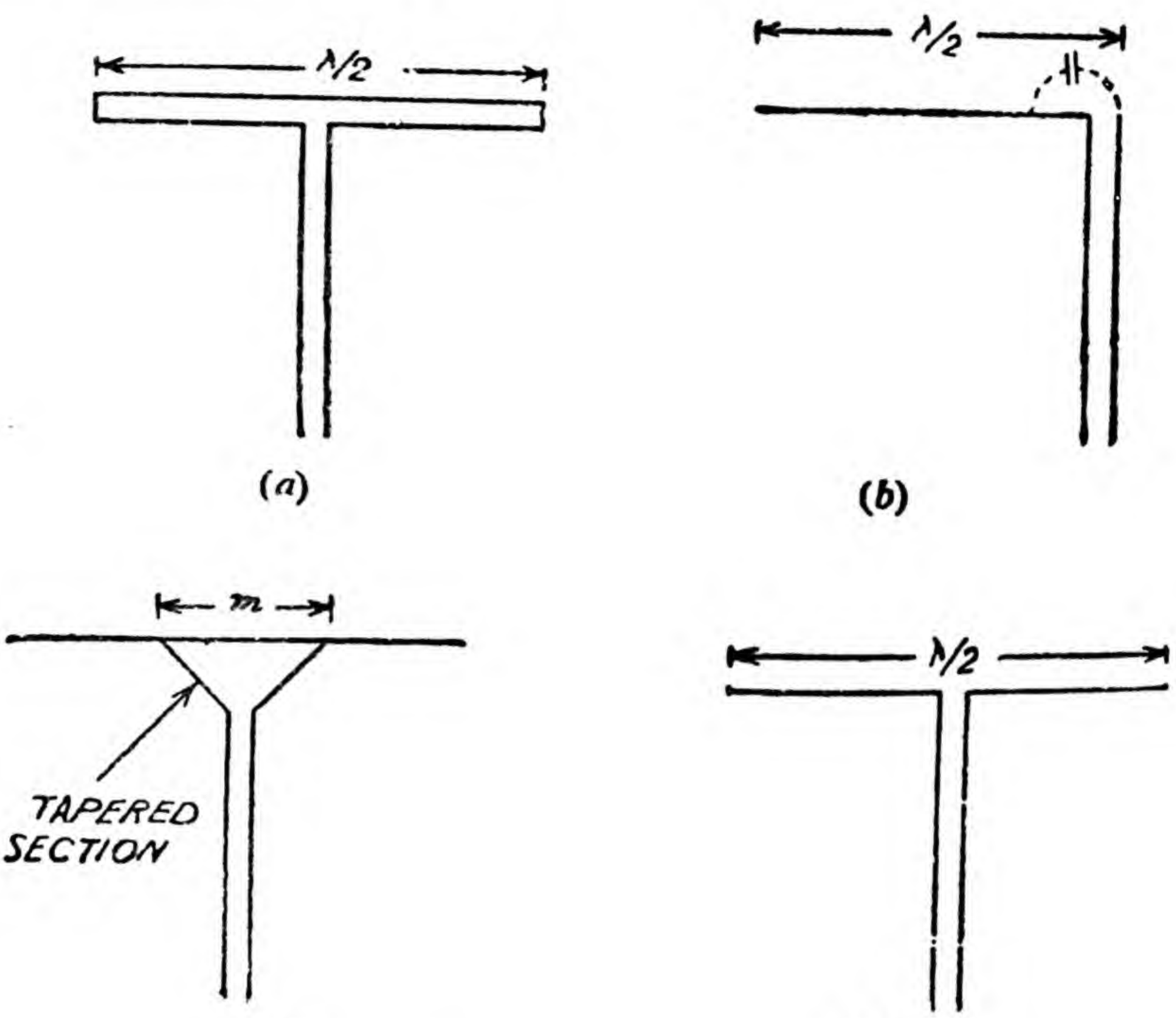


Fig. 61. Half-wave antenna arrangements.

Methods (*c*) and (*d*) provide impedance matching that permits non-resonant operation. In fig. (*c*), the spacing *m* is adjusted to provide an impedance match. This spacing in the case of a horizontal antenna above the perfect earth, depends upon the height. Arrangement at fig. (*d*) is basically similar to that at fig. (*a*), except that two closely spaced halfwave antennas are joined together at their ends and one of the dipoles is excited at the centre. This arrangement transforms the impedance of the antenna system. The two antennas function in *parallel* and each carriers *half* the total current. The result is that the impedance offered to the transmission line is 4 times the loop radiation resistance of a halfwave antenna, approximately 290 ohms.

(*b*) *Directional shortwave trasmitting antenna* : For moderate directivity, antenna arrays consisting of several radiators are sufficent. For l arge directivit either a rhombic antenna or an array of *V* antennas in the horizontal position is best suited. Readers should consult *An*

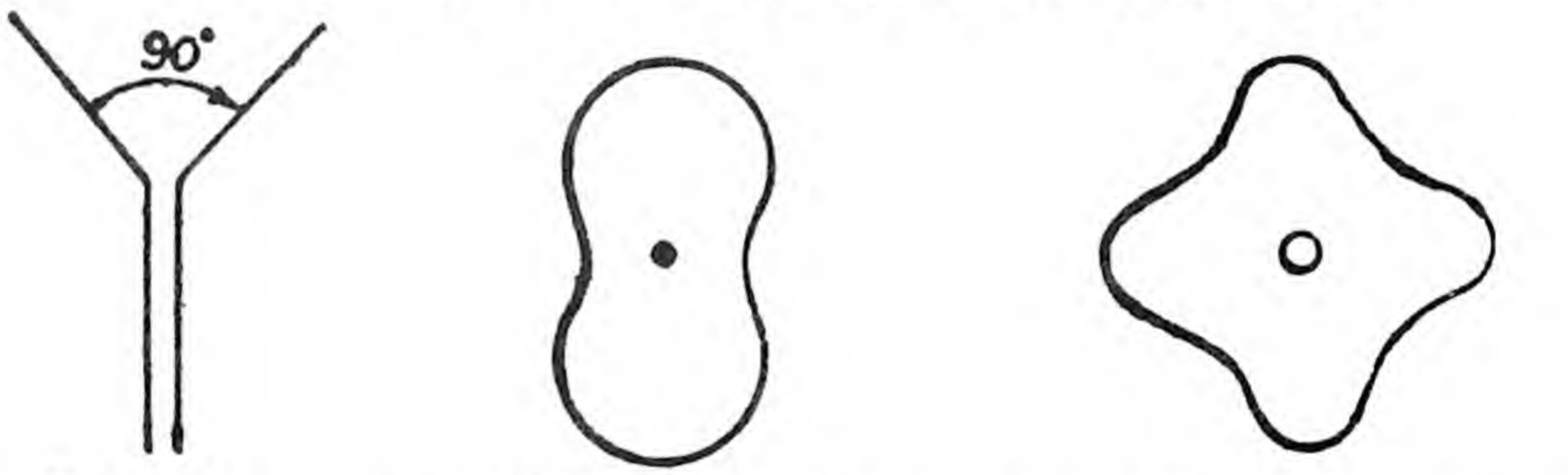
unt-excited grounded

Air plane antennas : Installation of transmitting antennas at high speed aircraft is rather difficult. At frequencies below 2 mc/s., or even at still higher frequencies trailing wire serves as a suitable transmitting antenna This trailing wire antenna consists of a wire, taken underside the plane, of such a length as to give quarter wave resonance when air-plane is used as a substitute for ground. Trailing wire is terminated with a weight.

(4) Transmitting antennas for frequencies 30-600 mc/s. (*Ultra high frequency range*) : The antennas suited for short wave operation are also suitable for ultra high frequency operation. To name simply, the different antennas commonly employed in this region are resonant half wave radiators, cone antennas, corner reflectors, Yagi antennas, *V* and rhombic arrays, high frequency loops.

Corner reflector antenna is especially suitable for ultra high frequencies because it is compact, portable and gives good gain.

For broadcrast purpose, at such frequency range, it is desired to radiate horizontally polarised signals equally strong in all directions in the horizontal plane. A *V* antenna with legs $\lambda/2$ or $\lambda/4$ long and an apex angle of 90° (fig. 62) serves the purpose.



(*a*) *V* antenna (*b*) Directional pattern for $L=\lambda/4$. (*c*) Directional pattern for $L=\lambda/2$.

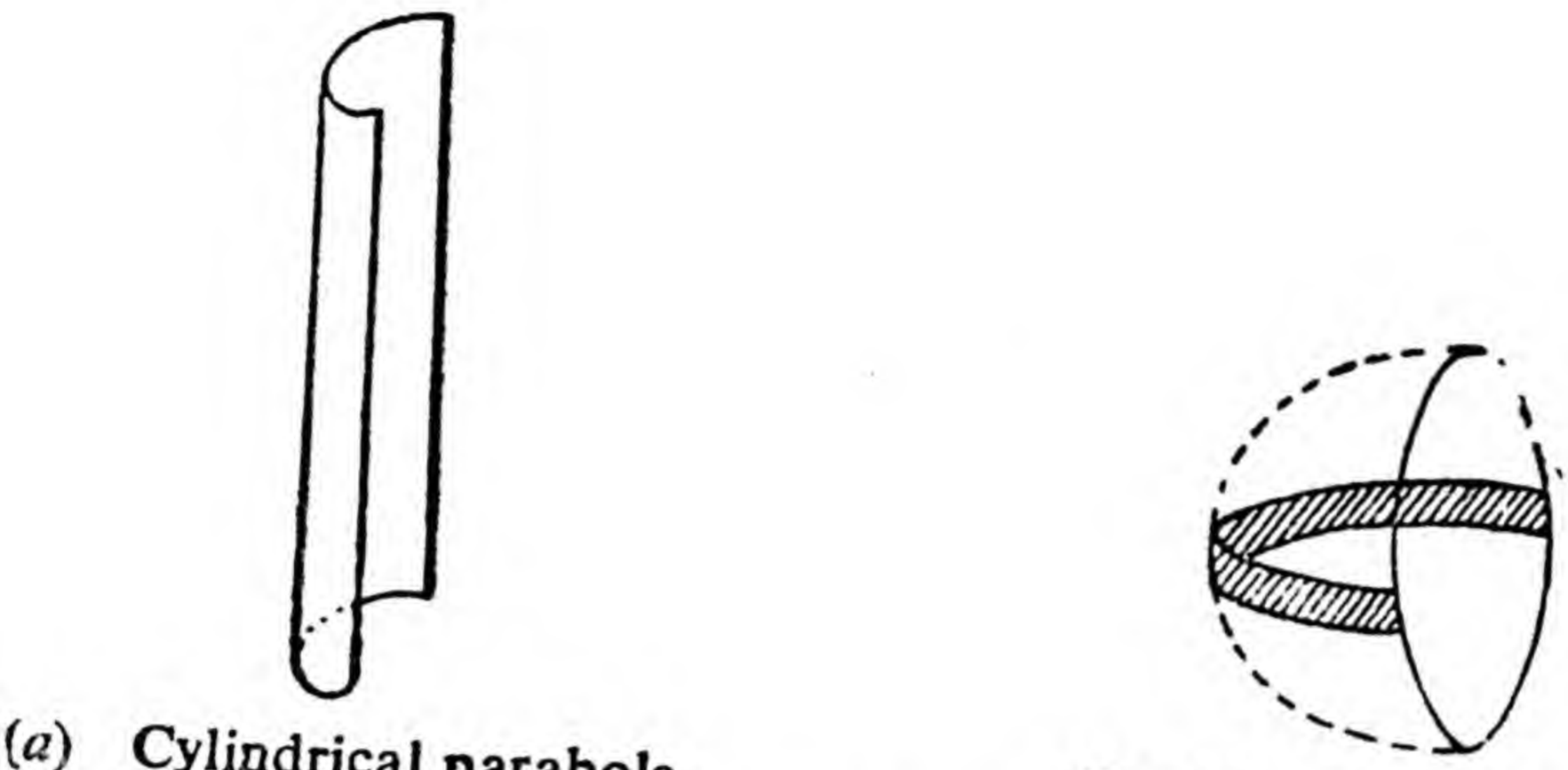
Fig. 62. *V* **antenna with short legs and directional characteristics.**

To obtain the directional pattern with vertical polarised waves, a broadside array, consisting of a number of half wave radiators staked vertically and excited in the same phase, can be employed.

(5) Transmitting antennas of frequencies higher than 600 mc/s. (*Microwave range*) : Directional antennas are most commonly employed for the radiation in such frequency range.

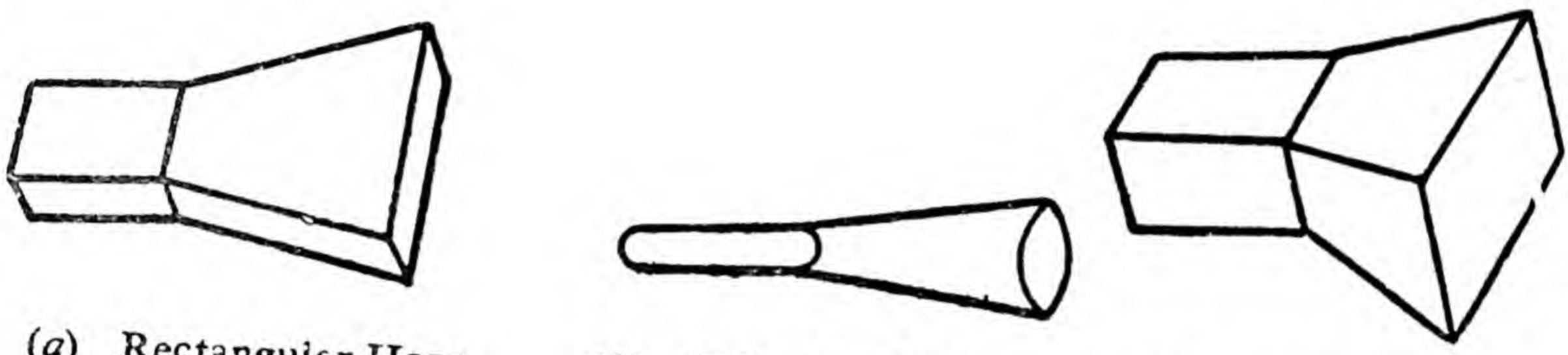
High Directivity : When high directivity is chief consideration, parabolic and horn antennas are generally employed : parabolic reflectors may assume a number of forms. In fig. 63 two forms are shown. The narrowness of the radiated beam depends upon the shape of the reflecting surface and the area of its aperture in wavelengths. Directivity is high when the aperature area in wavelengths is large.

Horn antennas are extensively used at such high frequencies. The directivity of the radiation is determined by the angle of flare and length of the horn. Few types of horn radiators are shown in fig. 64. Greater is the flare angle and horn length, narrower will be the radiated beam.



(*a*) Cylindrical parabola (*b*) Truncated paraboloid.

Fig. 63. Types of parabolic reflectors.



(*a*) Rectangular Horn (*b*) Cylindrical Horn (*c*) Pyramidal Horn

Fig. 64. Types of Horn antenna.

Moderate Directivity : When moderate amount of directivity is required, corner reflector is best suited. A corner reflector is shown in fig. 65. The reflector is usually illuminated by a half wave radiator located on a line bisecting the corner angle.

When low directivity is required, half wave radiators, excited from a concentric line through some form of transformer, are used.

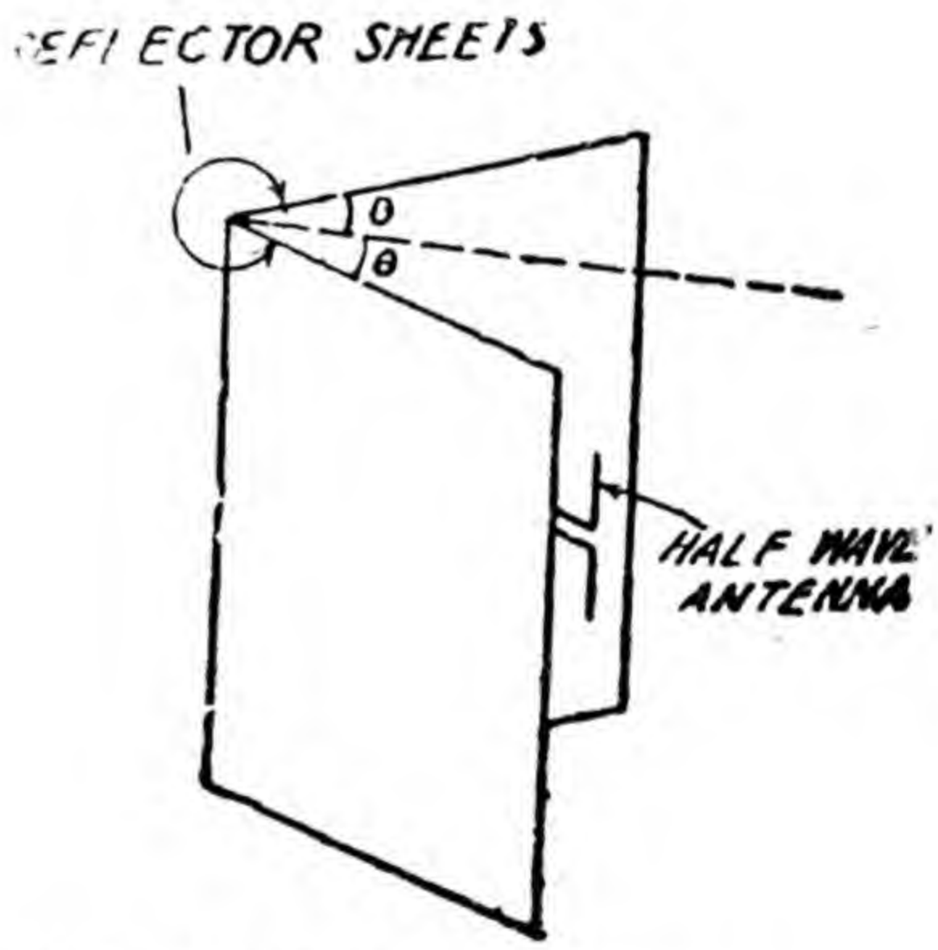


Fig. 65 A corner reflector.

28·24. RECEIVING ANTENNAS :

When electromagnetic waves are cut by an antenna, a voltage is induced in the latte by the magnetic flux of the wave. This induced voltage is distributed over the entire length of the antenna. Because of this dirtributed voltage, current flows in the receiving antenna. If a load is connected in series with this antenna, the energy abstracted from the passing electromagnetic wave will be delivered to it.

The important consideration in the choice of a receiving antenna is its sensitivity to locally generated noise fields. In manmade noise electric induction field predominates and if antenna responds much to such fields, large disturbances would occur in the receiver. In this respect, loop antenna is best suited as compared to simple wire antenna since the latter responds much to the induction field.

The ideal property of an antenna is its ability to pick up minimum possible noise energy due to natural disturbances which, on the other hand, is comparable to the noise generated within the input portion of the receiving system. Higher is the noise level, lower is the frequency. At microwaves, noise level is low and it becomes necessary to employ greater directivity since it helps :

(*a*) to increase the intercept area,

(*b*) to reduce the noise energy picked by the receiving antenna from the noise field, provided all noise sources differ in direction from that of desired signal,

(*c*) to discriminate undesired signals originating in directions other than that of desired signal.

Half-wave radiators serve the purpose of normal reception for frequencies upto 30 mc/s. For the reception of frequencies lying in the range 20 to 600 mcs, *V*-antennas rhombic antennas, arrays of half wave radiators, or arrays of high frequency loops are best suited.

28 25 DIFFERENCE IN RECEIVING AND TRANSMITTING ANETENAS :

As regards the important characteristics of an antenna, *viz.* impedance and field pattern, both types of antenna bear the point of similarity. But there are differences in other properties that are peculiar to their uses. We shall list them as follows :

(1) **Current distribution :** Current distribution on receiving antenna is different from that on transmitting antenna. In latter antenna when it is thin, the current distribution is almost sinusoidal, while in the case of receiving antenna, except for resonant length antenna ($h=\lambda/4$), current distribution is not approximately sinusoidal even for thin antennas.

(2) **Optimum field pattern :** Although field patterns for both the antennas are the same, it is not necessary that optimum pattern for transmitting antennas should be optimum for receiving one. The meaning of 'optimum pattern' for the two antennas is, however, different. For transmiting antenna, the pattern which puts most of the signals in desired direction is optimum, while for receiving antenna optimum condition is not maximum received power, but rather maximum signal to noise ratio. It is not necessary that in every case, transmitting optimum condition will lead to receiving optimum condition. For example, if a minor lobe in the receiving

antenna pattern points towards the noise sources, it will bring a large amount of noise and hence gives rise to low signal to noise ratio. If this lobe happens to be the pattern for transmitting antenna, it will have no ill effect except low loss in the power radiated.

(3) Directivity consideration as regards signal to noise ratio : If the receiver is in correct direction, an increase in the directivity of the transmitting antenna will definitely improve signal to noise ratio at the receiver, but such is not the case with receiving antenna. The improvement in the latter case, assuming that set noise is negligible part of the total noise, depends upon the direction of noise source as stated below :

(*a*) if noise is coming from the direction of desired signal, there will be no improvement in signal to noise ratio,

(*b*) if less noise is coming from the direction of desired signal, improvement will be greater as compared to transmitting case,

(*e*) if noise is coming equally from all directions. the improvement in signal to noise will be exactly the same as that in transmitting case.

(4) Antenna coupling : The antenna coupling in the transmitting case is adjusted to feed the antenna with a specified amount of power, but not for maximum power transfer, while receiving antenna coupling is adjusted so as to yield the maximum signal to noise ratio. Maximum signal to noise ratio is not always accomplished with maximum power transference. For example, in the medium frequency reception, since the gain is ample to deliver sufficient audio power and set noise is negligible as compared with atmospheric noise, there is no use to increase the amount of power delivered by the antenna because the noise power increases in the same ratio as the signal power.

(5) Operation of an antenna of given electrical length as a function of frequency : A half wave transmitting dipole neglecting losses, fed with definite watts of power, will produce the same field intensity in a given direction whatever be the frequency, while a half wave dipole, used for reception, delivers to a matched load an amount of power that is proportional to the square of wavelength.

(6) Economic factor : Since a single transmitter feeds several receivers. it will not be partial if more money is spent to obtain the given improvement of the transmitting rather than on receiving antenna.

EXERCISES AND PROBLEMS

1. (a) Derive an expression for the electric field strength due to a short grounded transmitting antenna neglecting ground absorption How would you measure the field strength ?

(b) Derive an expression for the radiation resistance of a short antenna remote from ground and sketch the field patterns in horizontal and vertical planes.

2. Assuming earth to be a perfect conductor, deduce an expression for the electric intensity of the radiation field at a point r cm. away from a short vertical aerial.

Show that power radiated by it is given by

$$P = 1578 \left(\frac{h}{\lambda}\right)^2 I^2 \text{ watts.}$$

where h is the height of the aerial, λ is the wavelength and I is the current at the base.

3. Define electric field strength due to a radio transmitter. Neglecting ground absorption, deduce an expression for the electric field strength at a point on the earth's surface (assumed as flat) in terms of power radiated by the transmitting aerial distance.

Describe how would you find the value of the electric field strength at a point experimentally.

4. What do you understand by the effective height of an aerial ? Describe how is it determined experimentally ?

Obtain an expression for the radiation resistance of a short vertical grounded aerial.

5. What is meant by the gain of an antenna array ? Derive an expression for the polar diagram in horizontal plane for antenna array, consisting of N elements, spaced half a wavelength apart, in a straight line and energized in phase.

6. Two similar aerials A and B are situated at a distance d from each other and they are fed by high frequency currents of equal amplitude and frequency which differ in phase by θ. Deduce an expression for computing the polar diagram of such an aerial system and show it possesses directive property.

7. A short vertical grounded aerial is designed to radiate on 1 mc/s. If the effective height of the aerial is 30 meters, what is its radiation resistance ?

8. Explain what is meant by the radiation resistance of an antenna. Derive expression for the electric field and the magnetic field due to a symmetrical thin, linear, centre fed antenna Hence show that the radiation resistance of a half wave dipole is 73 ohms.

9. Give the theory of dierctional transmission by means of an antenna array.

A broad side array consists of six vertical radiators space $\lambda/2$ apart and fed in phase. Draw the horizontal polar diagram of radiation.

Hint : In horizontal plane, they radiate uniformly in all directions and consequently can be treated like an array of six isotropic point sources. Use the formula with $n=6$, $\delta=0$]

10. Write in detail the need of directional transmissior and how is it achieved at different frequency ranges ?

Calculate the expression for the secondary field patterns of parabolic reflectors.

11. At what frequencies horn and lense antennas are used and why ?

What do you understand by an optimum horn ? Find the formula for the design of such a horn and sketch the field patterns for *E* plane and *H* plane sectoral horns at various flare angles.

12. Write an essay on broad band antennas, giving the description of each with special reference to their use in radar and television.

13. The electric field strength of a plane wave is 0·1 *V/m*. What is magnetic field strength in micro-webers/meter^2 ?

14. Show that the radiation field electric intensity of a quarter wave monopole is

$$E_\theta = \frac{60 I_m}{r_0} \left[\frac{\cos (\pi/2 \cos \theta)}{\sin \theta} \right] V/m,$$

What is the power radiated for peak antenna current equal to 10 amperes ?

15. Deduce an expression for the induced electromotive force in a loop antenna. Discuss its application as directional finder.

16. Discuss the effect of ground on the radiation by an antenna which is very close to earth. Explain the concept of image antenna.

Find the expression for field distribution around a vertical grounded antenna of length $\lambda/2$.

17. Derive an expression for the radiation field for a continuous array consisting of point sources of uniform amplitude and of the same phase.

18. Define radiation intensity. Derive an expression for the radiation intensity of a current element, starting from the expression for the radiation field. Derive an expression for the radiation intensity or an end fire couplet.

19. Explain what is meant by retarded magnetic vector potential. Derive an expression for the electric field intensity at a distant point due to a half wave dipole and hence discuss its radiation resistance.

20. Explain the basis from which the current distribution on a symmetrically fed, thin, linear antenna may be estimated. Draw a nature of voltage and current distribution along a linear $3\lambda/2$ antenna, fed at the centre, and calculate its distribution in horizontal and vertical planes.

21. Comment briefly but critically on (i) end array, and (ii) aperture radiator.

29 PROPAGATION OF ELECTROMAGNETIC WAVES

In chapter 28 it has been established that high frequency currents in aerials produce a detachment of electromagnetic energy in the form of waves of the same nature as light—a process usually called *radiation* and one which may be directional in nature.

In this chapter, we shall study the phenomena which occur in the media between the transmitting and receiving antennas. There are several mechanisms by which radio waves may travel from a transmitting to a receiving antenna, the important of these being designated by the terms *ground waves, sky waves, and space or tropospheric waves.*

***Ground wave or Surface wave :** The ground wave is of practical importance at broadcast and lower frequencies. This wave can exist when the transmitting and receiving antennas are close to the surface of earth, and is supported at its lower edge by the presence of ground. The wave is vertically polarised because horizontal electric field components, being in contact with earth, suffer energy losses and thus at any considerably distance from transmitting antenna, only the vertical component of the field is left. At day time, broadcast signals are usually carried by ground waves.

Sky wave : The sky waves are of practical importance for very long distance radio communication. The energy by these waves reaches the receiving antenna as a result of bending of their path due to ionisation in the upper atmosphere. This ionised region, which begins about 80 kilometers above the surface of earth, is termed as ionosphere.

Space wave : Space wave represents the energy that travels from transmitting to the receiving antenna in the earth's troposphere. Troposphere is a portion of earth's atmosphere extending upto first 10 miles from the earth's surface. Radio transmission at frequencies higher than 30 mc/s. is usually space wave propagation. The space wave commonly consists of two components, one of which travels directly from transmitting to the receiving

---

*Sometimes the term 'ground wave' is referred to mean both 'surface wave' and 'space wave' together.

antenna and the other reaches as a result of reflection from the surface of the earth.

29·1. SURFACE WAVE PROPAGATION :

A ground wave can be divided into two components—a surface wave and a space wave. The magnitude of these individual components and also of the wave is affected by (*i*) the resistivity and dielectric constant of the earth, (*ii*) the earth's curvature, (*iii*) the height of transmitting and receiving antennas, (*iv*) the distance to transmitter, (*v*) the frequency of the signal, (*iv*) variation of refractive index of earth's atmosphere with height.

Surface wave is a wave that is guided along the surface of earth and around its bends and curvature, much as electromagnetic wave is guided by a transmission line. To make up for the ground losses, energy is supplied from this wave and hence the attenuation of this wave must be affected directly by the constants of the earth.

When the transmitting and receiving antennas both are located at the earth's surface, the two parts of a space wave—the direct and ground reflected—cancel each other (earth's reflection coefficient is —1 for such a case) and the surface wave is left purely to carry the intelligence. Due to the bending round curvatures, this wave has an advantage of propagating the signal through long distances.

Surface wave field intensity due to a vertical dipole antenna :

The field intensity at a point, considering all types of propagation, will be the sum of the intensities due to ground wave (both components, *i.e.*, surface wave and space wave) and sky wave. In the present article, we are simply concerned with ground wave. Now out of its two components which would predominate depends upon all the six factors stated earlier.

In cylindrical co-ordinates, the axial component of electric intensity E_z for a dipole antenna of length dl and carrying a current I has been calculated by K. A. Norton and is given by

$$E_z = j.30\beta I\, dl \left[\cos^2\psi \left(\frac{e^{-j\beta R_1}}{R_1} + R_v \frac{e^{-j\beta R_2}}{R_2} \right) + (1-R_v)(1-u^2+u^4\cos^2\psi)\, F.\left(\frac{e^{-j\beta R_2}}{R_2} \right) \right], \quad \ldots(1)$$

and the radial components,

$$E_\rho = -j30.\beta I\, dl \left[\sin\psi\cos\psi \left(\frac{e^{-j\beta R_1}}{R_1} + R_v \frac{e^{-j\beta R_2}}{R_2} \right) - \cos\psi\,(1-R_v)\, u\sqrt{(1-u^2\cos^2\psi)}\, F.\frac{e^{-j\beta R_2}}{R_2}\left(1+\sin^2\frac{\phi}{2} \right) \right] \quad \ldots(2)$$

where $\beta=\frac{2\pi}{\lambda}$, λ being the wavelength of transmitted signal,

R_v=reflection coefficient of the ground for a plane wave,

R_1=distance between transmitting and receiving points by a direct wave path,

R_2=distance between transmitting and receiving points by a ground reflected wave path,

ψ=angle made with the horizon by the wave leaving the transmitting antenna (if the height of receiving antenna is very small compared to its distance from the transmitting antenna, then $\psi_1=\psi_2=\psi$, fig. 1).

$u^2=\frac{1}{k+jx}$, k being the relative dielectric constant of earth,

F=attenuation factor that depends upon earth's constant,

and $x=\frac{\sigma}{\omega k}$, σ being the conductivity of earth, and ω the angular frequency of radiation

$=\frac{1\cdot8\times10^4\,\sigma \text{ ohm/metre}}{f_{mc}}$, where f_{mc} is frequency in mc/s.

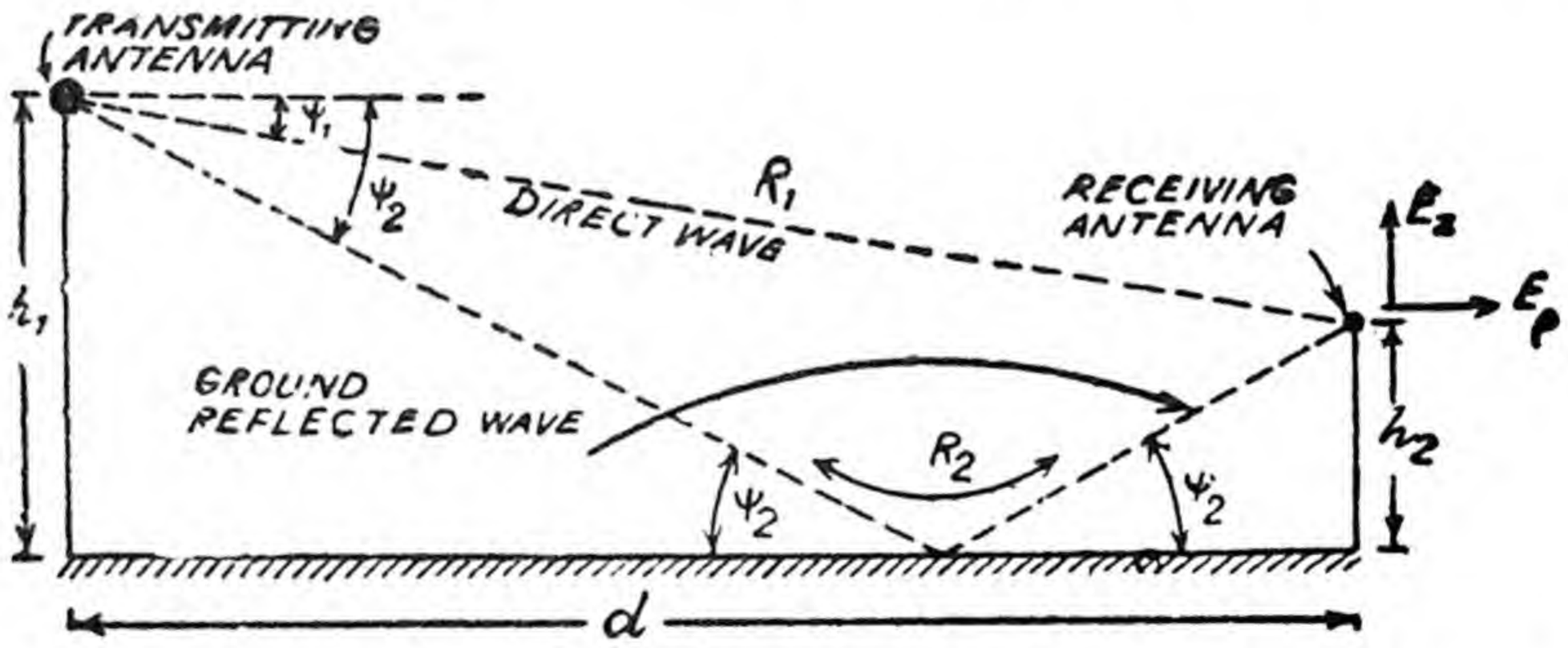


Fig. 1. Geometry for direct and ground reflected waves illustrating different parameters n ground wave propagation.

Equations (1) and (2) show that first bracketed term in each of the equations is independent of ground attenuation factor F, *i.e.*, unaffected by earth factors and these both should represent space wave. Hence

$$E_{space}=\sqrt{\{E_z^2\,(\text{space})+E_\rho^2\,(\text{space})\}}$$

$$=j.30.\beta I\,dl\cos\psi\left(\frac{e^{-j\beta R_1}}{R_1}+R_v\,\frac{e^{-j\beta R_2}}{R_2}\right)$$

The second term, depending upon F, is a surface wave. The total surface wave, neglecting term containing u^4 and adding vectorially, will be

$$E_{surface}=j30\;\beta I\,dl\,(1-R_v)\,F.\frac{e^{-j\beta R_2}}{R_2}$$

$$\times\sqrt{\{1-2u^2+u^2\cos^2\psi\,(1+\sin^2\psi/2)\}}$$

We shall here discuss the propagation due to surface wave. For long distance (much greater than λ) transmission $R_1=R_2=R$ and when the dipole is at the surface of earth, expression for the field due to surface wave can be written as :

$$E_{surface}=j.30.\beta I\ dl.\ (1-R_v)\ F.\left(\frac{e^{-j\beta R}}{R}\right)$$

$$\times\left[\mathbf{k}(1-u^2)+\mathbf{r}.\cos\psi\left(1+\sin^2\frac{\psi}{2}\right)u\sqrt{(1-u^2\cos^2\psi)}\right], \quad ...(3)$$

where **k** and **r** are the unit vectors respectively parallel and perpendicular to the vertical dipole.

Function F depends upon distance, frequency and on earth's constants. As the attenuation of surface wave increases with frequency, this mode of propagation is important only for lower frequencies (long and medium waves) and for the points near earth's face. Value of F approaches unity as distance R tends to zero, *i.e.*, for distance within a few wavelengths of dipole, $F=1$. Value of F is given by

$$F=[1+j\sqrt{(\pi q)}\ e^{-q}\ \text{erf}\ (-j\sqrt{q})] \quad ...(4)$$

where $$q=\frac{j\beta Ru^2\ (1-u^2\cos^2\psi)}{2}\left[1+\frac{\sin\psi}{u.\sqrt{(1-u^2\cos^2\psi)}}\right]$$

and $$\text{erf}\ (-j\sqrt{q})=\frac{2}{\sqrt{\pi}}\int_{-\sqrt{q}}^{\infty}e^{-v^2}\,dv$$

Putting $F=1$ in equatlon (3), field pattern for the wave called 'unattenuated surface wave' can be sketched. In fig. 2, plots of pattern for two values of earth conductivity $n=\left(\frac{x}{k}\right)$ are shown. For low frequencies and good conductivity ($n=100$), unattenuated surface wave field intensity is small except for angles near grazing $\psi=0$, the value being 2 at this value of ψ. For higher frequencies and poor conductivity ($n=1$), field intensity is appreciably high at high angles, *i.e.*, surface wave is stronger, though its value at $\psi=0$ is still 2. Attenuation of this wave is quite rapid as the distance increases because of the factor F. Thus for $n=1$ and at the points near earth's surface, surface wave is much stronger than space wave at long distances.

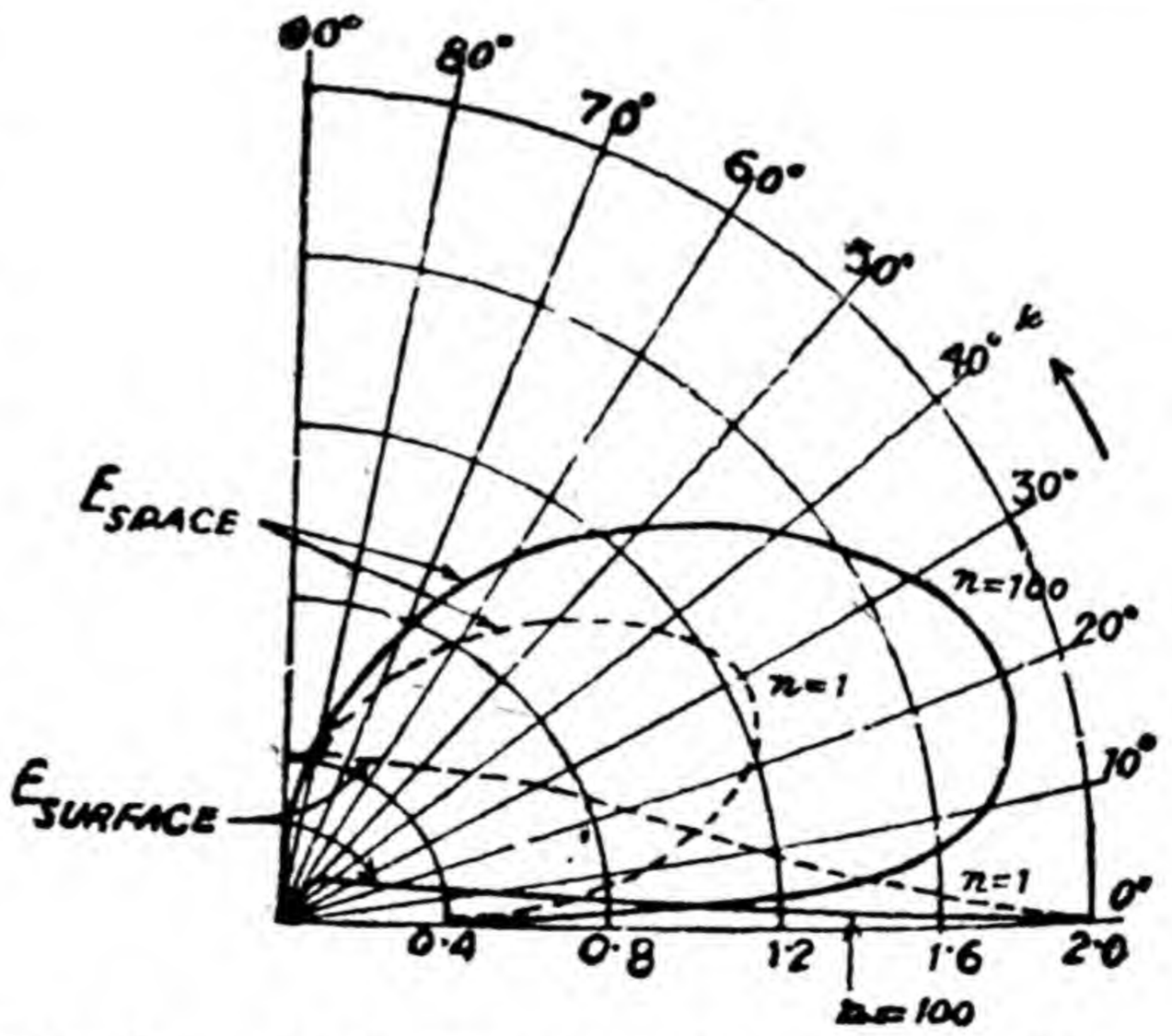


Fig. 2. **Vertical field pattern of a vertical dipole at the surface of the earth for n=1, and n=100.**

The magnitude of F for $\psi=0$, *i.e.*, at the surface of earth is called 'ground wave attenuation factor' and is denoted by A, *i.e.*,

$$A=|F|=|1+j\sqrt{(\pi q)}\, e^{-q} \operatorname{erf}(-j\sqrt{q})|_{\psi=0}$$
$$=|1+j\sqrt{(\pi p_1)}\, e^{-p_1} \operatorname{erf}(-j\sqrt{p_1})|, \quad \ldots(5)$$

where p_1 is the value of q for $\psi=0$.

p_1 is complex in nature and can be expressed as

$$p_1=pe^{jb},$$

where p is called the 'numerical distance' and b the phase constant.

Evaluating q, at $\psi=0$, shows that

$$p=\frac{\pi R \cos^2 b''}{\lambda x \cos b'} \approx \frac{\pi R}{\lambda x} \cos b$$

and
$$b=(2b''-b') \approx \tan^{-1} \frac{k+1}{x}$$

where
$$b''=\tan^{-1} \frac{k}{x}$$

and
$$b'=\tan^{-1} \frac{k-\cos^2 \psi}{x}=\tan^{-1} \frac{k-1}{x}.$$

In figure 3, a plot of A against the numerical distance p for various values of b is shown. The numerical distance p is directly proportional to the actual distance to the transmitter and the square of frequency, and varies almost inversely, as the ground conductivity. b'' is called 'power factor angle' and hence b is a measure of power factor angle of earth.

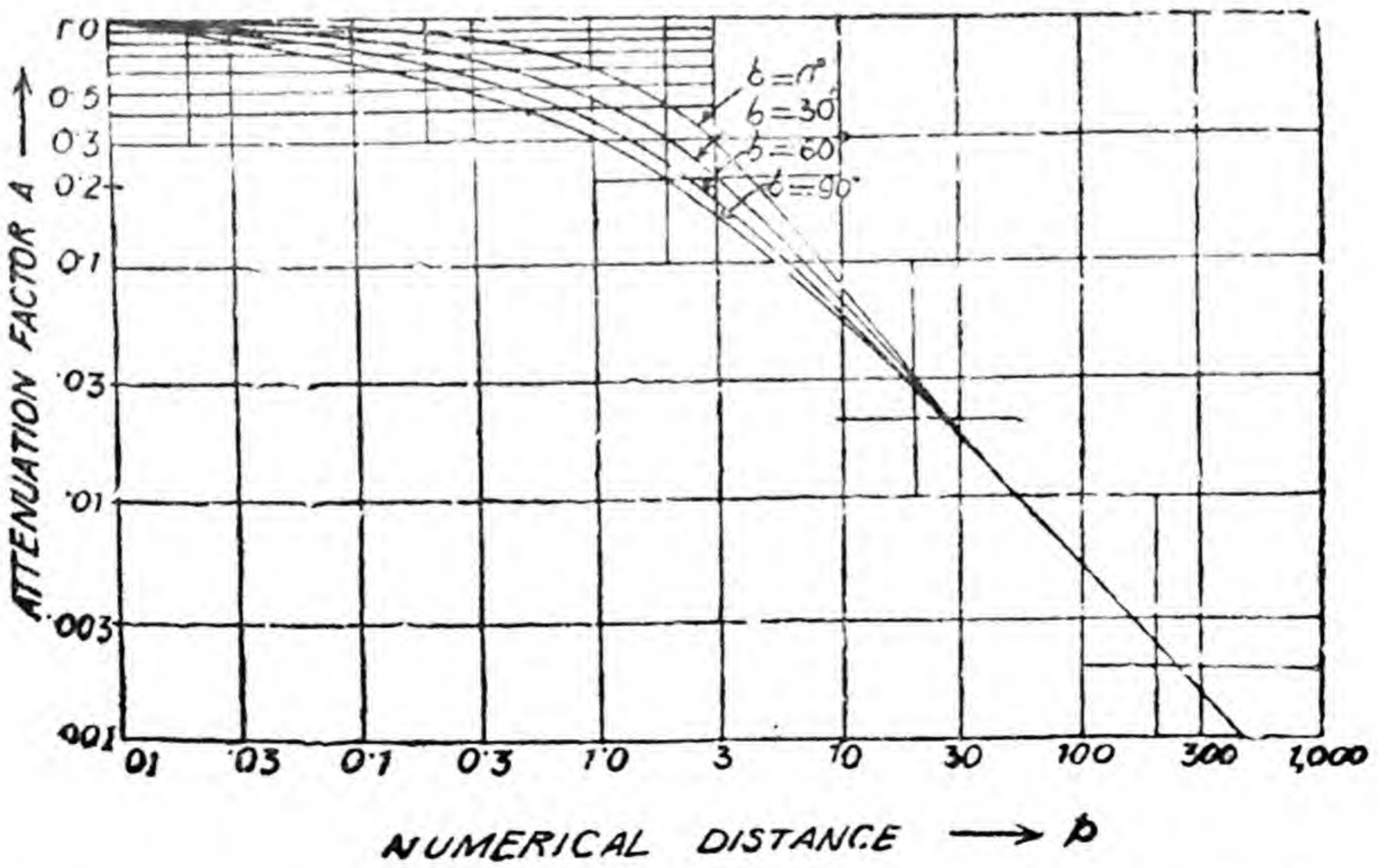


Fig. 3. Variation of attenuation factor with numerical distance for different values of b.

Earth impedance can be resistive or reactive. For better earth at broadcast frequencies, $x \gg k$ and impedance of earth is mainly resistive, the power factor angle being nearly zero. At high frequencies and over poor earth $x \ll k$ and impedance of earth is reactive.

For all values of b, attenuation factor A can be expressed as

$$A = A_1 - \sin b . \sqrt{(p/2)} . e^{-5p/8}. \qquad ...(6)$$

For $b < 5°$ $A_1 = \dfrac{2 + 0{\cdot}3\,p}{2 + p + 0{\cdot}6p^2}$

Now for $b < 5°$ and $p < 4{\cdot}5$, *i.e.*, for short numerical distance

$$A = e^{-0{\cdot}43p + 0{\cdot}01p^2},$$

and for $b < 5°$ and $p \geqslant 4{\cdot}5$, *i.e.*, for large numerical distance,

$$A = \frac{1}{2p - 5{\cdot}7}$$

When numerical distance p is very small ($p < 1$), ground attenuation factor approaches unity, but for large distances ($p > 10$), magnitude of A varies inversely with p. This causes the field strength to vary inversely as the square of the distance and hence rapidly decreases with p.

As stated earlier, when $\psi = 0$, *i.e.*, when transmitting and receiving antennas both are located at earth's surface, earth's reflection coefficient is nearly one and both components of space wave cancel each other. Propagation is then purely due to surface wave. But for antennas elevated above the ground as in the case of high frequency transmission, resultant field strength is a combination due to surface and space waves.

Curves of fig. 3 are valid upto a distance $\dfrac{80}{f_{mo}^{1/3}}$ km. When the distance is large compared to wavelength, curvature of earth affects the propagation and the above expressions deduced by Nortons for plane earth are not accurate. Curvature of earth also affects the space wave and the effective antenna height.

29·2. SPACE WAVE PROPAGATION :

Space wave propagation finds its wide scope at higher radio frequencies like V. H. F., U. H. F. and microwaves since at these frequencies, propagation due to surface wave and sky wave almost fail. Although at lower frequencies, space wave propagation may also be adopted, but sky waves and surface waves are important to give a major contributions to the propagation. Space wave propagation is practically limited to the line of sight† distance and is also limited by earth's curvature.

---

†If a direct ray passes from transmitter to the receiving point without being intercepted by the buldge in earth's surface, considering variation of refractive index of earth's atmosphere with height then the distance between the two is said to be a line of sight distance.

As pointed out in previous articles, field intensity due to space wave is

$$E_{space}=j30.\beta\ I.l_e \cos\psi \left[\frac{e^{-j\beta R_1}}{R_1}+R\,\frac{e^{-j\beta R_2}}{R_2}\right], \qquad \ldots (1)$$

where R is the ground reflected coefficient and l_e is the effective antenna length.

Ground Reflection Coefficient : The vector ratio of the reflected to the incident wave is termed as reflection coefficient R and expressed as $R=r\,|\underline{\rho}$, where r represents the magnitude and ρ is the modification of phase produced by the imperfections of the earth.

The exact value of reflection coefficient depends upon (*i*) dielectric constant k, (*it*) conductivity of earth σ, (*iii*) frequency, and (*iv*) angle of incidence with which the wave strikes the surface of earth. For the two polarisations, it is expressed as :

For vertical polarisation,

$$R_v=r|\underline{\rho}=\frac{(k\quad j\lambda)\sin\psi-\sqrt{\{(k-jx)-\cos^2\psi)}}{(k-jx)\ \sin\psi+\sqrt{\{(k-jx)-\cos^2\psi\}}} \qquad \ldots(2)$$

for horizontal polarisation :

$$R_h=r\,|\underline{\rho}=\frac{\sin\psi-\sqrt{\{(k-jx)-\cos^2\psi\}}}{\sin\psi+\sqrt{\{(k-jx)-\cos^2\psi\}}}. \qquad \ldots(3)$$

Plots in fig. 4 and fig. 5 for two different frequencies 4 mc/s. and 100 mc/s., show how the magnitudes and phases of these complex reflection coefficients vary with angle of incidence Full lines correspond to R_v (vertical polarisation) and dotted lines to R_h (horizontal polarisation).

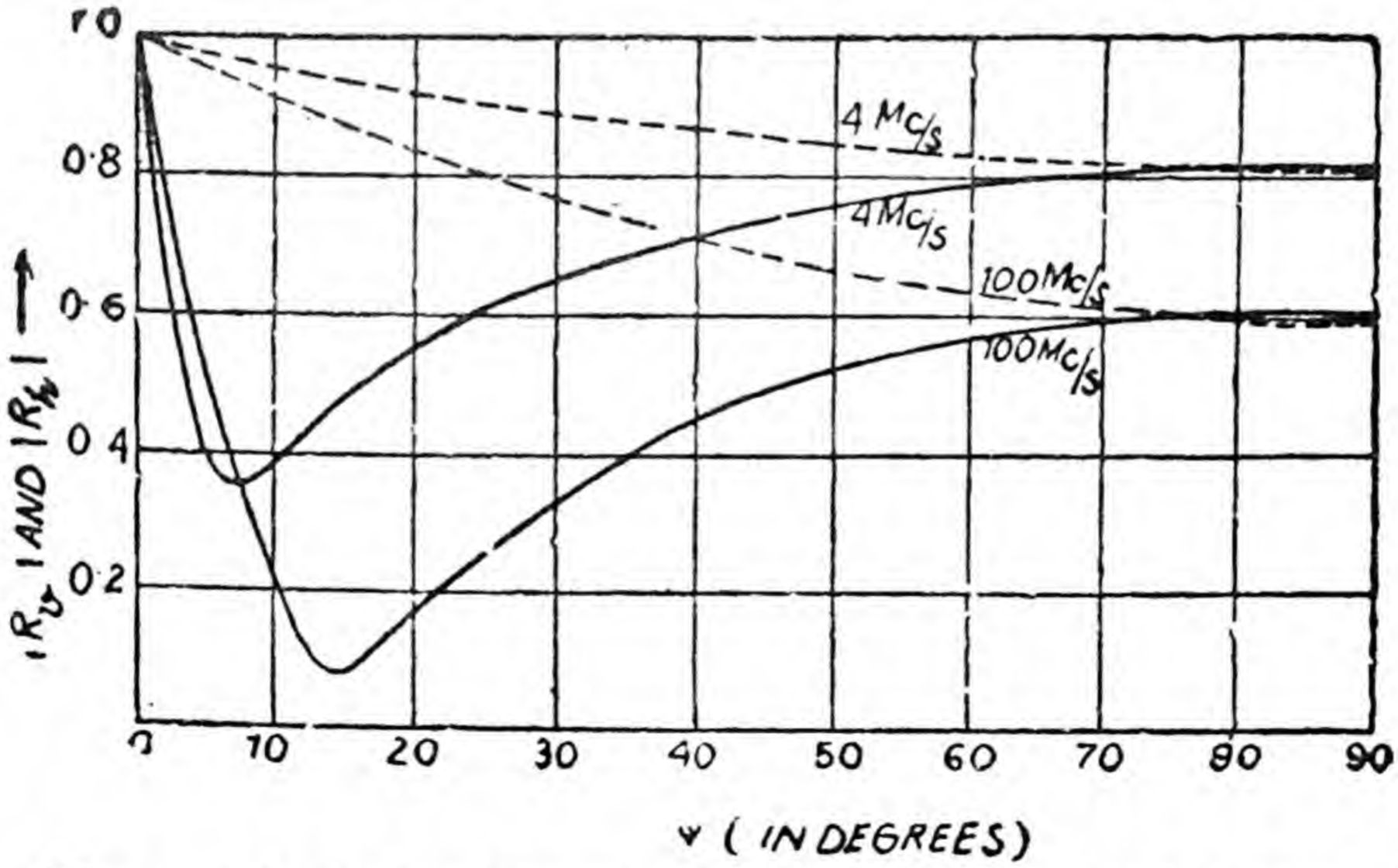


Fig 4. Variation of the magnitude of reflection coefficients with angle of incidence.

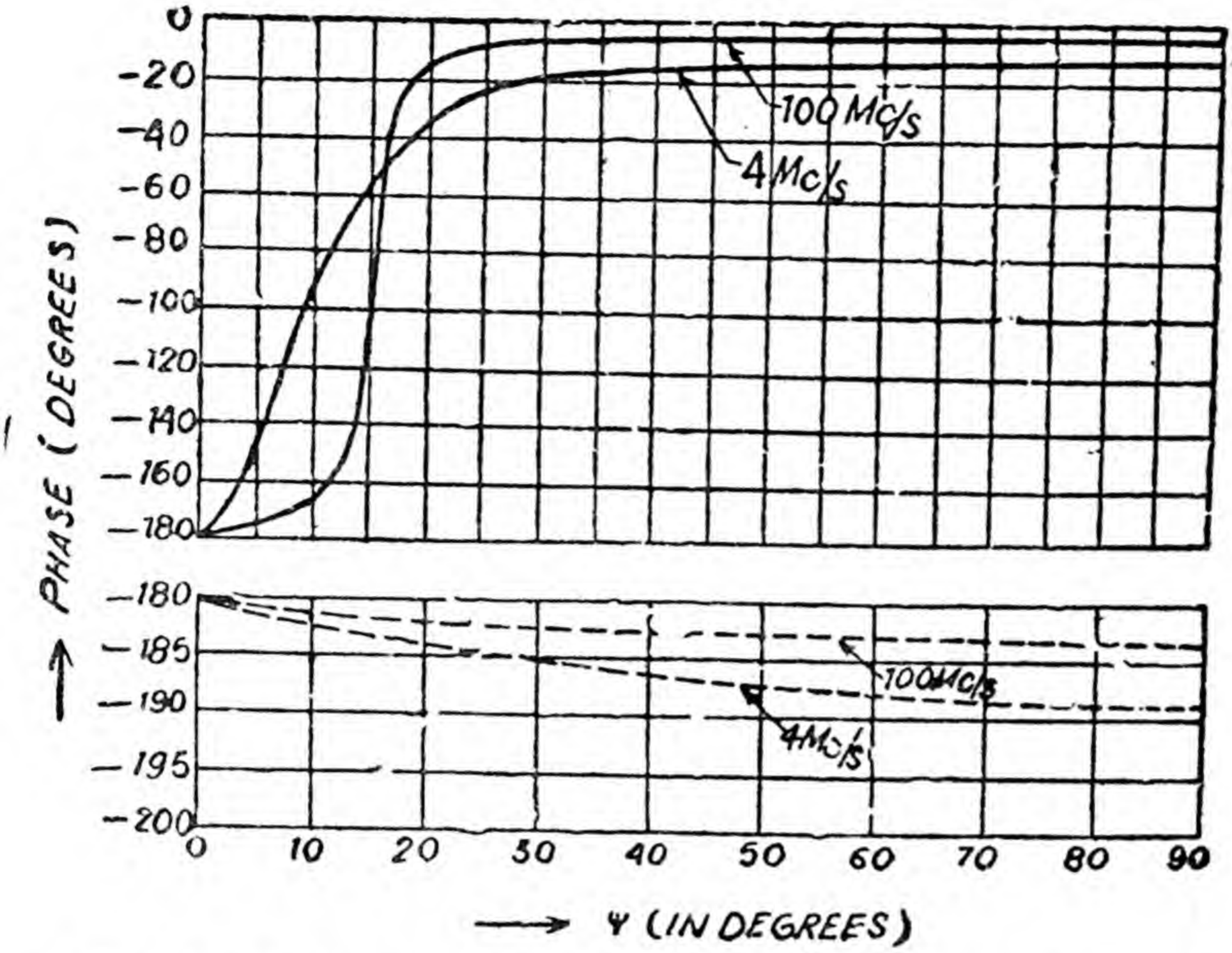


Fig. 5. Variation of the phase of reflection coefficients with angle of incidence.

From the figures we note that

(*a*) for vertical polarisation (full line curves) the magnitude goes through a minimum (and phase through 90°) for some value of angle of incidence ψ for a given earth condition. This particular value of ψ is called pseudo Brewster angle (or just Brewster angle). Its value depends on the wave frequency and earth constant, *i.e.*,

(*i*) for low frequencies and high conductivities ($x \gg k$), the Brewster angle is less and approches zero as x tends to infinity (for perfec t earth).

(*ii*) for high frequencies and low conductivities ($x \ll k$), the Brewster angle is practically of the same value as it has for a per fect dielectric.

(*b*) for horizontal polarisation (dotted line curves), the magnitude of reflection coefficient is unity for grazing incidence ($\psi=0$). It decreases steadily as the value of ψ increases.

(*c*) for normal incidence ($\psi=90°$), electric intensities for both types of polarisation are parallel to the reflecting surface but oppositely directed. Thus magnitudes of R_h and R_v are the same but phase differs by 180°. This apparent difference in phase is purely due to the ways in which one defines the positive directions for the reflected wave in two cases.

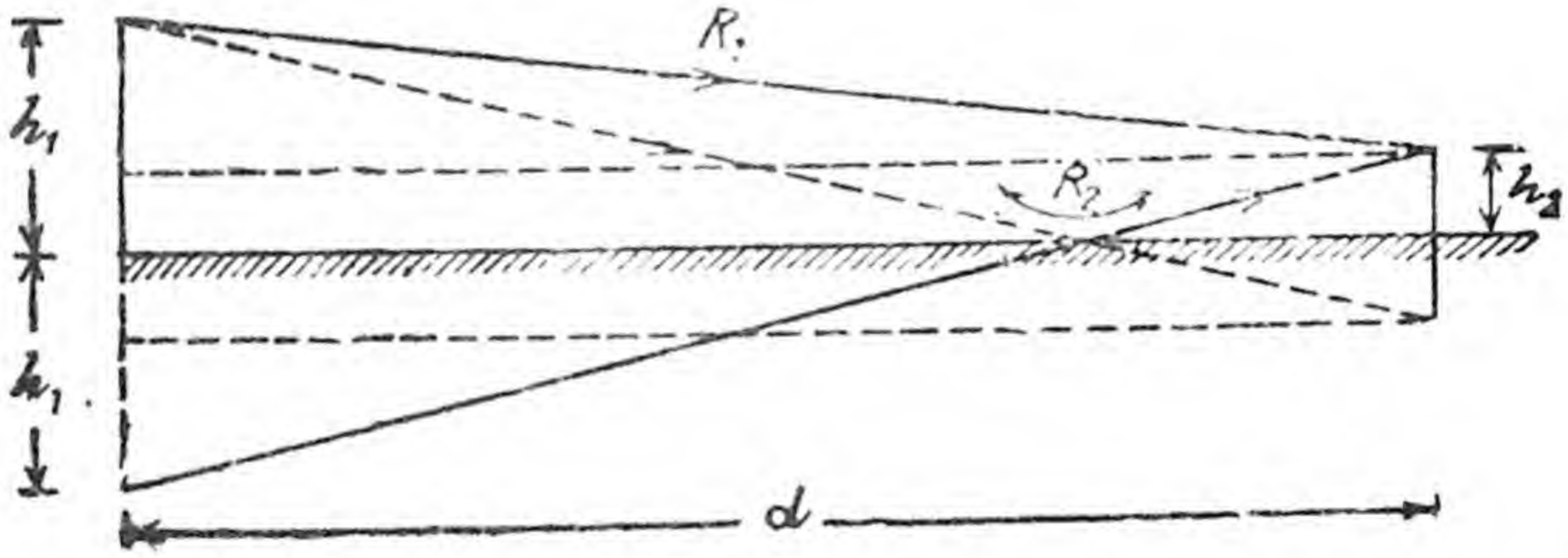
Elevated Dipole Antennas above a Plane Earth : As mentioned earlier, the propagation is entirely due to surface wave when transmitting and receiving atennas are located at earth, reflection coefficient being -1 due to which two components of space wave

cancel each otner. This is the case with day time reception of ordinary broadcast signals.

At higher frequencies, where wavelength becomes too short, it is possible to elevate the antennas a fraction of wavelength above the ground. The space wave will no longer be zero and the intensity at a point will be the resultant of space wave as well as surface wave field intensities.

Since in this article we are concerned with space wave propagation, an expression for the field intensity due to space wave at the receiving point will be calculated. This expression, on the other hand, will also represent the ground wave propagation at very high or ultra high frequencies since at these frequencies surface wave is almost attenuated because of ground factors.

Consider two vertical antennas elevated at heights h_1 and h_2 above the surface of earth (fig. 6). The resultant space wave field intensity at a point depends upon the relative phases of ground reflected and direct components. Since this phase difference varies with path difference signal intensity in space will exhibit maxima and minima.



Fi g 6. Elevated antennas of heights h_1 and h_2 with direct and ground reflected wavepaths.

Field intensity due to space wave from equation (1), is

$$E=j.30\ \beta.I.\ l_e \cos\psi \left(\frac{e^{-j\beta R_1}}{R_1}+R\,\frac{e^{-j\beta R_2}}{R_2}\right)$$

First we assume that R_1 is nearly equal to R_2, so that in magnitude, e may put $R_1\approx R_2=d$; but for phase, the difference (R_2-R_1) always remains finite and affects the phase difference between two signals. Secondly, for simplicity, we assume that $R=-1$. Thus,

$$E=j.3\,0.\beta I l_e \cos\psi\left(\frac{e^{-j\beta R}}{d}-\frac{e^{-j\beta R_2}}{d}\right)$$

$$=j.30\ \beta I l_e \cos\psi\ \frac{e^{-j\beta R_1}}{d}\left[1-e^{-j\beta\,(R_2-R_1)}\right].$$

From figure 6,

$$R_1=\sqrt{\{d^2+(h_1-h_2)^2\}}$$

$$=d\left[1+\left(\frac{h_1-h_2}{d}\right)^2\right]^{1/2}=d+\frac{(h_1-h_2)^2}{2d}$$

Similarly, $R_1=d+\frac{(h_1+h_2)^2}{2d}$.

Now $(R_2-R_1)=\frac{2h_1h_2}{d}$ and $\beta=\frac{2\pi}{\lambda}$.

Putting these values in expression for field intensity, we get

$$E=j.30.\frac{2\pi}{\lambda}.I.l_e\cos\psi.\frac{e^{-j\ (2\pi/\lambda)\ R_1}}{d}\left[1-e^{j\cdot(4\pi h_1h_2/\lambda d)}\right]$$

or $$|E|=\frac{60\pi Il_e\cos\psi}{\lambda d}\left[1-\cos\frac{4\pi h_1h_2}{\lambda d}^{-2}+j\sin\frac{4\pi h_1h_2}{\lambda d}\right]$$

As $d \gg h_1$ or h_2, we may put, $\cos\frac{4\pi h_1h_2}{\lambda d}=1$, so that

$$|E|=\frac{60\pi Il_e\cos\psi}{\lambda d}.\sin\frac{4\pi h_1h_2}{\lambda d} \qquad ...(4)$$

Since sine factor involves h_1, h_2, λ and d, field intensity will vary simultaneously with the variation of any of these quantities. Thus the received field intensity will exhibit sinusoidal variation of maxima and minima if we vary antenna height, distance d from the transmitter or frequency of the signal.

Suppose λ or d is very large, then $\frac{4\pi h_1h_2}{\lambda d}$ will be small for which we can write

$$\sin\frac{4\pi h_1h_2}{\lambda d}^2=\frac{4\pi h_1h_2}{\lambda d},$$

then $$|E|=240\pi^2\ Il_e\cos\psi.\left(\frac{h_1h_2}{\lambda^2d^2}\right). \qquad ...(5)$$

Thus for long waves, *i.e.*, low frequencies, space wave field strength varies inversely as the square of the distance. On the other hand, even if frequency is moderately high, equation (5) holds good provided distance d is sufficiently large.

In fig. 7 variation of field intensity with distance for two frequencies 50 mc/s. and 300 mc/s. is shown. It can be noted from the figure that :

(*a*) For 50 mc/s. curve shows no maxima and minima and field intensity decreases as $1/d^2$.

(*b*) For 300 mc/s., curve shows maxima and minima for smaller values of d. For large values of d, *i.e.*, larger distances, field intensity decreases as $1/d^2$ like case (*a*).

Dotted curve represents the variation of free space field intensity with distance (neglecting ground reflection).

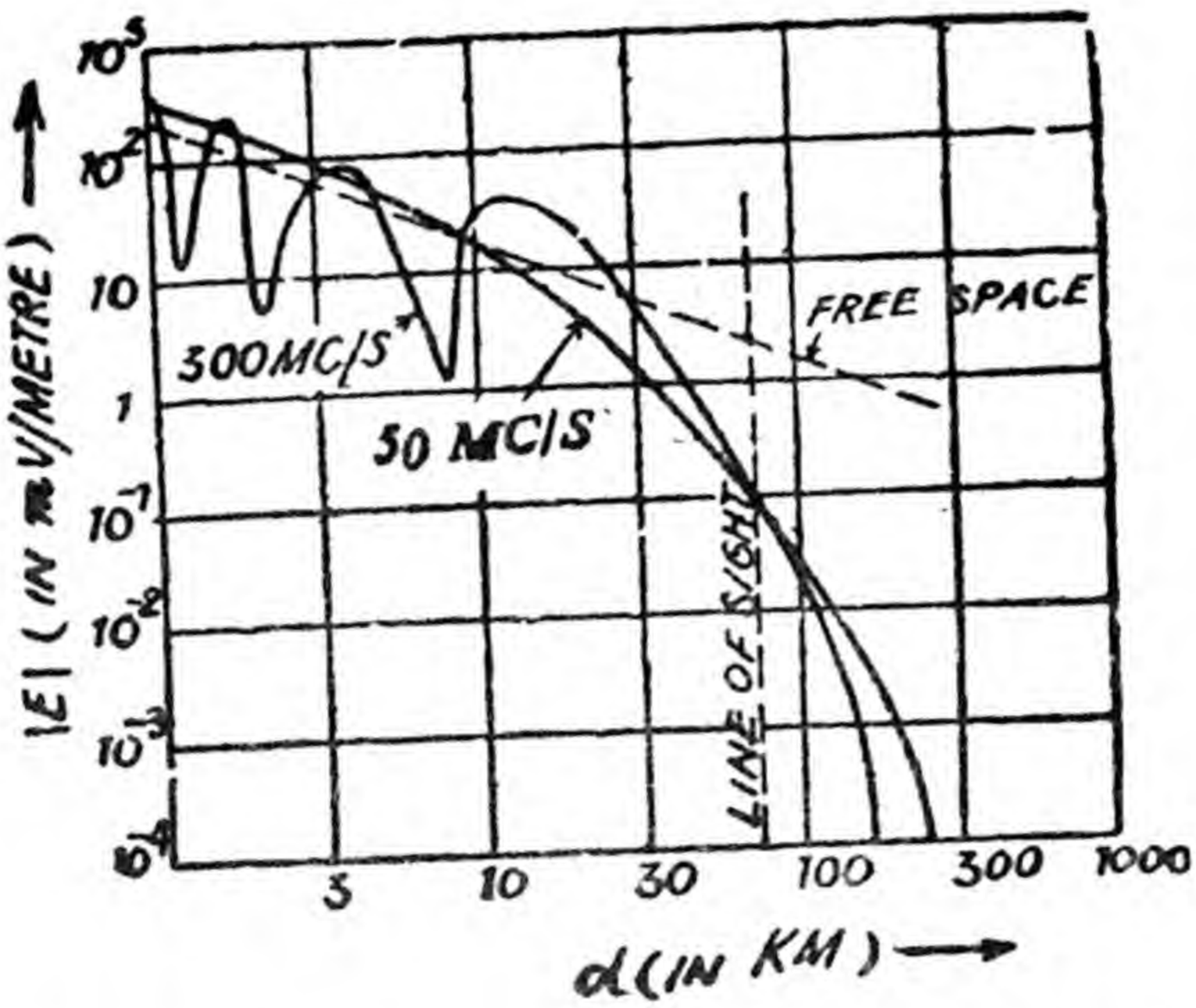


Fig. 7. Variation of field strength with d'stance.

Tropospheric refraction : At frequencies V.H.F., U.H.F. and microwaves, propagation by sky waves is not possible, since the high frequency waves sent upwards are not reflected from the ionosphere Propagation along the earth's surface (surface wave propagation) is also not possible since the surface wave field intensity decreases very rapidly for distances beyond the line of sight. Thus apparently it becomes quite unexpected to think about the transmission much beyond the horizon But this transmission does occur because of the refraction of such frequency waves in troposphere.

Troposphere is a region of earth's atmosphere quite adjacent to earth and extending upwards upto 10 km. In troposphere, temparature falls at the the rate of 6·5°C per kilometre till it reaches −5°C at the upper boundary. Region next to troposphere is stratosphere in which temperature almost remain constant, −50°C.

Tropospheric refraction is the phenomenon in which radio-waves gradually bend downward due decrease in refractive index with height. This refraction increases the range of propagation beyond the 'line of sight' distance. This increase in range is, however, explained by using effective radius of curvature of earth which is 4/3 times the actual radius (*i.e* assuming the space waves propagate along the straight line paths and that the earth has an effective radius slightly larger than its actual radius). One can compute the range of propagation for a normal atmosphere and for a given antenna height using effective radius of earth and is called 'modified line of sight' distance for space wave propagation.

Fading : When space wave signals are received from a distant station, phenomenon of fading often occurs. Fading is simply the variation in field intensity with time. The cause of this fading is

that, because of meterological conditions, the refractive index in troposphere undergoes certain changes in its rate of variation which, in its effect, alters the path difference : and hence phase difference) between the direct and ground reflected components of the space wave at the receiving point. The result is the variation in field intensity at the point of reception.

The second cause of this fading is the variations in tropospheric reflections* which are produced by discontinuities in the troposphere. Since the magnitude, orientation and height of these discontinuities may vary from moment to moment, severe fading is resulted when tropospheric wave component of space wave is dominant at the receiving point.

In high frequency propagation, fading usually occurs since the signal of these frequencies is affected by tropospheric conditions.

Duct propagation : Duct propagation, as a physical picture, is somewhat analogous to guided wave transmission in a leaky waveguide.

Possibility for such a propagation arises when a layer whose modified† refractive index decreases with height (called inversion layer) is present. Because of this negative gradient, a ray coming out of the layer will be bent away from the normal (over a flat earth). When angle of incidence taken with respect to normal, exceeds a critical value i_c, this ray will bend back to the inversion layer and is somewhat trapped between ground and height h_0. Such trapping limits the wave propagation in a duct.

Curve 1 in fig. 8 shows the path of ray that has been trapped between ground and height h_0. Curve 2 shows the ray path of a barely trapped wave and curve 3 is for an untrapped wave. A ray represented by curve 1 will obivously follow the curvature of earth and hence can travel to larger distance far beyond the line of sight

When inversion layer is adjacent to earth surface, duct will have its lower side on the earth surface and such ducts are called surface ducts. But when inversion layer is above the surface of earth, lower side of duct is also elevated and such ducts are called elevated ducts.

---

*Tropospheric reflection arising due to discontinuities is important when the receiving point is well below the l ne of sight of transmitter. For such a location, troposphere waves are much stronger than those arriving at the leaving point by the diffraction of ground wave around the curvature of earth.

†The term modified refers to the fact that the actual index has been modified to account for the curvature of earth.

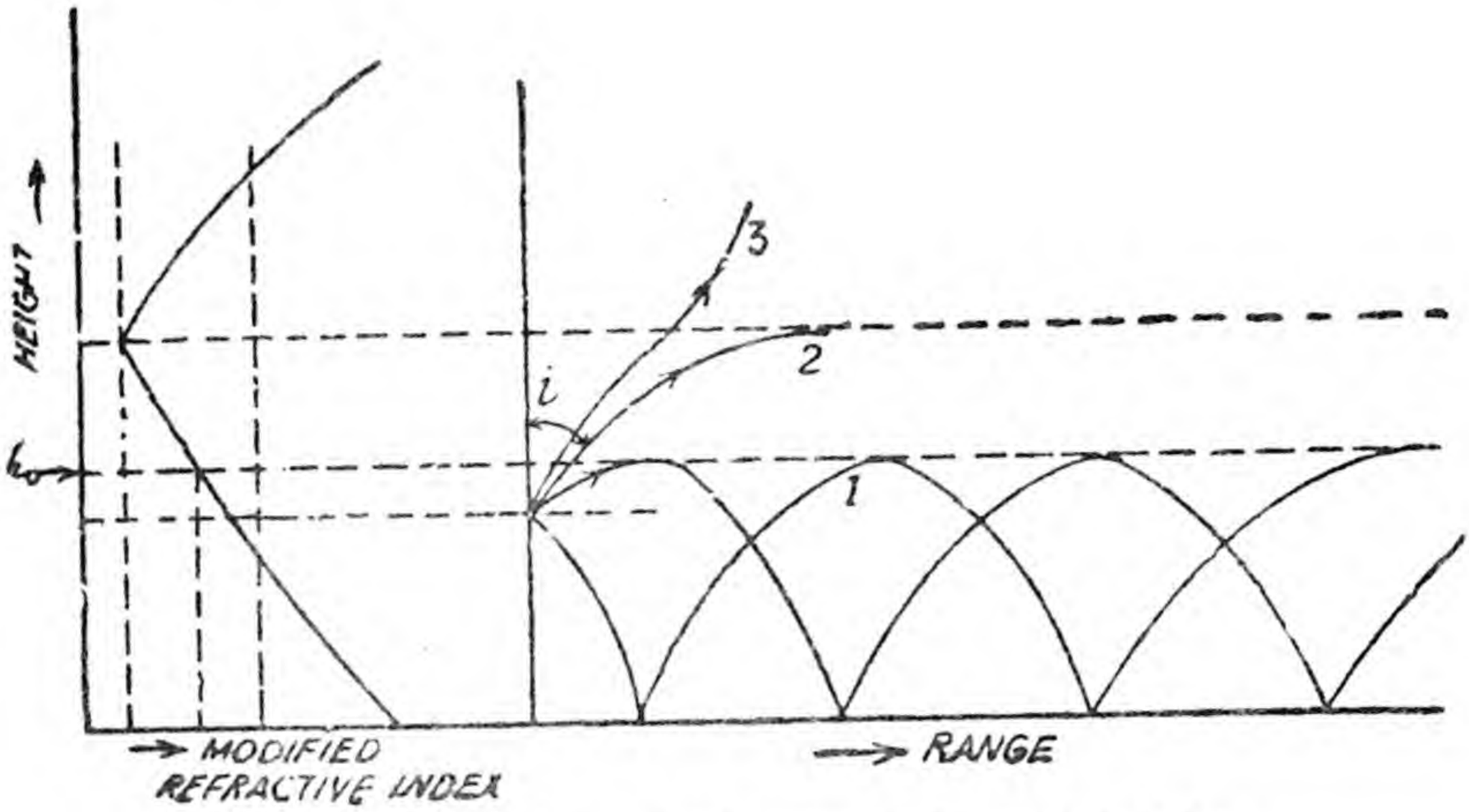


Fig. 8 (*a*) Variation of refractive index with height.
(*b*) Wave trapping.

The physical picture of wave trapping is more analogous to that in a waveguide with the difference that while the modes, in which the field may be resolved in both cases in a waveguide, are confined to the guide, they penetrate the top of the duct to some extent. Like to a waveguide action as a high pass filter (signals having wavelengths above certain maximum value can not propagate through waveguide), signals having wavelengths above certain maximum value, $\lambda_{max.}$, are not allowed by the duct to be trapped. For higher wavelengths, the penetration or leakage from the duct increases rapidly and consequently trapping is very little. The value of $\lambda_{max.}$ is approximately given by

$$\lambda_{max.} = 0{\cdot}084 d^{3/2} \qquad \ldots (6)$$

where $\lambda_{max.}$ is in cm. and duct height d is in metres. It is obvious that for a duct height of 10 metres, $\lambda_{max.} \approx 2{\cdot}5$ cm. This mode of propagation is important at ultra high and microwave frequencies.

Effects of earth's curvature : Curvature of earth affects the space wave propagation in three ways :

(*a*) Plane wave reflection coefficient, which occurs in the expression for field intensities, is different for curved earth. The effect is least important due to circumstances adopted in practice.

(*b*) The energy of a ground reflected wave, being reflected from curved surface, will be diverged more as compared to plane earth case. To account this, ground reflected wave has to be multiplied with divergence factor D, given by

$$D = \left[1 + \frac{2 d_1 d_2}{\alpha a d \cos \psi_2'} \right]^{-1/2}.$$

The distances d_1, d_2, d and angle ψ_2' are shown in fig. 8 (*c*). α gives the fractional increase in the earth's radius of curvature.

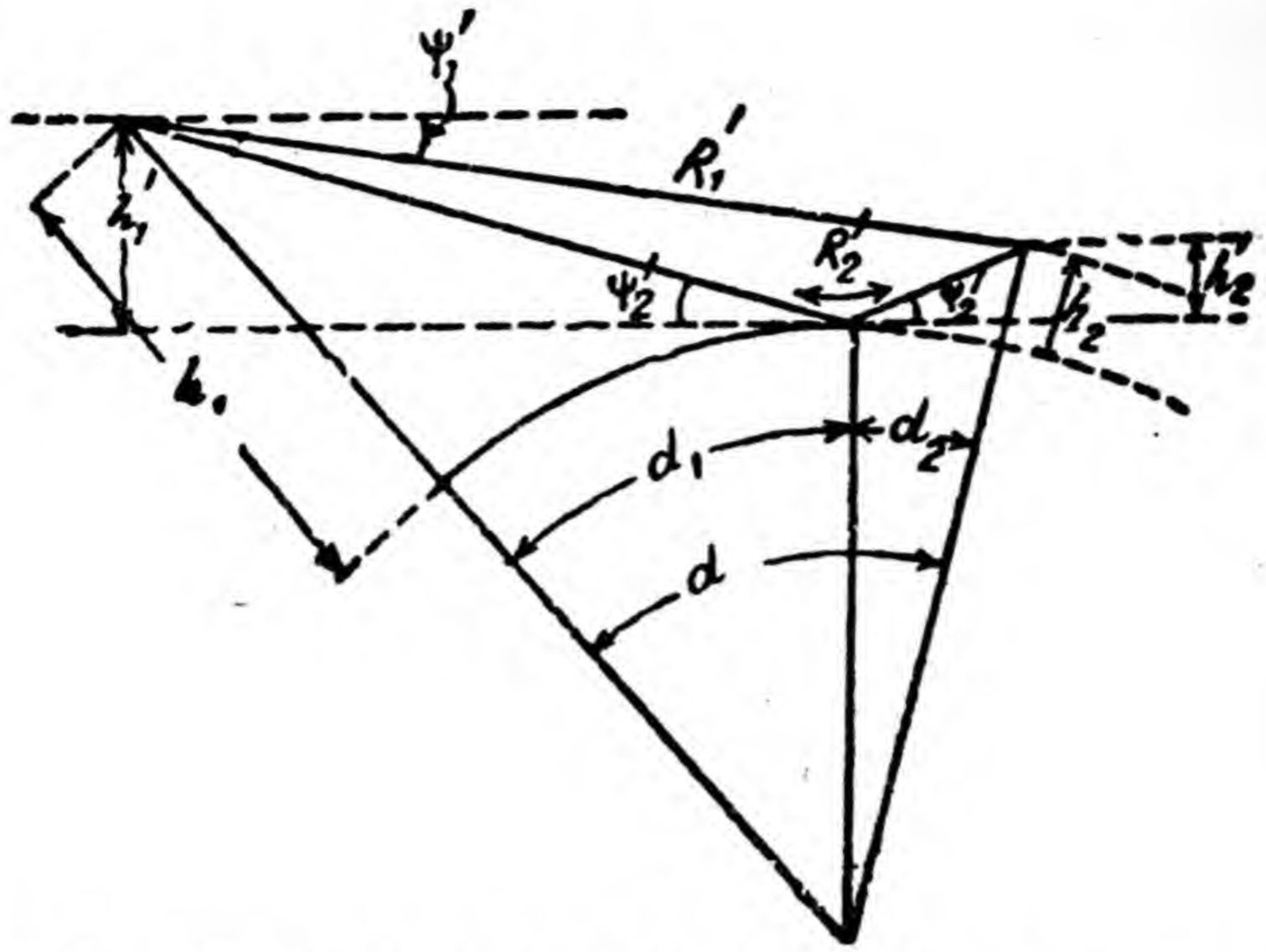


Fig. 8 (*c*) Illustrating the effect of earth's curvature on space wave propagation.

The effective radius of curvature is αa, a being the actual earth's radius. Value of α is usually $\frac{4}{3}$, *i.e.*, (1·33).

(*c*) Effective heights h_1' and h_2' of the transmitting and receiving antennas above the plane that is tangential to earth at the point of reflection are less than the actual heights h_1 and h_2 above the surface of earth. By geometry, we have

$$h_1' = h_1 - \frac{d_1^2}{2\alpha a}$$

$$h_2' = h_2 - \frac{d_2^2}{2\alpha a}$$

$$\tan \psi_2' = \frac{h_1' + h_2'}{2} = \frac{h_1'}{d_1} = \frac{h_2'}{d_2}. \qquad \ldots(7)$$

The maximum distance d_0 for a line of sight propagation in the case of elevated antennas can be shown to be

$$d_0 = \sqrt{(2\alpha a h_1)} + \sqrt{(2\alpha a h_2)}, \qquad \ldots(8)$$

where h_1 and h_2 are the heights of transmitting and receiving antennas respectively; a is the earth's radius.

When the receiving point is at ground, $h_2 = 0$

$$d_0 \simeq \sqrt{(2\alpha a h_1)}. \qquad \ldots(9)$$

Both equations (8) and (9) are valid for h_1 and h_2 less than 20,000 ft.

29·3. PROPAGATION OF SKY WAVES THROUGH IONOSPHERE :

The ionosphere is produced in upper atmosphere chiefly by the ultraviolet radiation from the sun. Since the pressure of air at such heights is very low and ultraviolet radiation is very intense, the formation of ions easily takes place. The air has a variable chemical composition depending on altitude, the lighter gases tending to prevail at greater heights. The ionisation of different gases

will be different because of their different ability to absorb ultra violet radiations of various wavelengths. As a result, *the density of ionisation will vary in an irregular way with altitude*. On the whole, ion density first increases with altitude and then decreases, but it shows 'ledges' where the ion density varies more slowly with altitude. These ledges are named as layers and the process of layer formation is called 'stratification of ionosphere'.

This region, also called plasma, of upper atmosphere ranges in altitude from approximately 50–1000 kilometers, where the ionisation is sufficient to interfere with the propagation of electromagnetic waves. It appears that the conductivity of this ionised region is due almost exclusively to the electron density, except possibly near the lower limit of the ionosphere. The free electron density, N, is typically 10^{11} per metre^3, varying by about 10^{10} to 10^{12} per metre^3 from the lowest to the highest layer. The number of molecules per cubic meter varies from obout 10^{12} to 10^{15}. With altitude percent ionisation increases rapidly but it always remains low. We shall show that for $N=10^{11}$ per metre^3, f_N, the plasma angular frequency, is nearly 3 mc/s. The following analysis will also predict that the electromagnetic waves of frequency lower than 3 mc/s are not transmitted through, whereas waves of frequency higher than 3 mc/s will be sent back to the earth Thus ionosphere helps in wave propagation to different global parts at these frequencies. Frequencies much higher than 3 mc/s penetrate the ionosphere and are thus unaffected by the ionosphere.

PROPAGATION OF ELECTROMAGNETIC WAVES IN THE ABSENCE OF EARTH'S MAGNETIC FIELD :

When an electromagnetic wave passes through ionised medium, current arises from the presence of free electrons and ions in the electric and magnetic fields of the wave. We shall assume that there are no collisions between the electrons and gas atoms or molecules which is, infact, true except in the lowest regions of the ionosphere where the pressure is highest, and at frequencies of the order of 1 mc/s or higher.

Let us consider a plane electromagnetic wave travelling in the positive direction along x-axis with **E** and **H** vectors parallel to y- and z axes respectively An ion of charge q, mass m, and velocity **v** is subjected to a force **F** given by

$$\mathbf{F}=q[\mathbf{E}+(\mathbf{v}\times\mathbf{B})] \text{ in m.k.s. system,}$$

the ion being situated at (x, y, z) with respect to its average position (0, 0, 0). $\mathbf{B}$ $(=\mu\mathbf{H})$ is the magnetic induction. Since $\mathbf{E}=\mathbf{E}_y$ and $\mathbf{B}=\mathbf{B}_z$ we have

$$m\frac{d^2x}{dt^2}=qB_z\ \frac{dy}{dt}=qB\frac{dy}{dt} \qquad \ldots(1)$$

$$m\frac{d^2y}{dt^2}=q\left[E-B\frac{dx}{dt}\right] \qquad \ldots(2)$$

$$m\frac{d^2z}{dt^2}=0. \qquad \ldots(3)$$

From eq. (2),

$$\frac{d^2y}{dt^2}=\frac{qE}{m}\left[1-\frac{1}{c}\frac{dx}{dt}\right] \qquad \ldots(4)$$

in which we have put

$$*E=Bc.$$

We shall show that

$$\frac{1}{c}\frac{dx}{dt}\ll 1,$$

or velocity of the ion along x-axis, due to action of electromagnetic wave, is much smaller than the velocity of light. Then from equation (4), we arrive at

$$\frac{d^2y}{dt^2}=\frac{qE}{m}. \qquad \ldots(5)$$

Let the electric vector of the wave be presented by

$$E=E_0\cos\omega t,$$

then

$$\frac{d^2y}{dt^2}=\frac{qE_0}{m}\cos\omega t$$

or

$$\frac{dy}{dt}=\frac{qE_0}{m\omega}\sin\omega t \qquad \ldots(6)$$

Putting it in equation (1), we get

$$m\frac{d^2x}{dt^2}=\frac{qE_0}{c}\frac{qE_0}{m\omega}\sin\omega t$$

or

$$\frac{d^2x}{dt^2}=\frac{q^2E_0^2}{m^2\omega c}\sin\omega t\cos\omega t$$

giving

$$\frac{dx}{dt}=-\frac{q^2E_0^2}{4\omega^2m^2c}\cos 2\omega t, \qquad \ldots(7)$$

where the constants of integration are set to equal zero because $x=0$, $y=0$, $z=0$ at $t=0$. Now we can show that

$$\left|\frac{1}{c}\frac{dx}{dt}\right|=\frac{q^2E_0^2}{4\omega^2m^2c^2}.$$

For electron, $q=1{\cdot}6.10^{-19}$ coulomb, $m=9{\cdot}11.10^{-31}$ kg., $c=3.10^8$ meters/sec, we have

$$\left|\frac{1}{c}\frac{dx}{dt}\right|=\frac{2{\cdot}17\times10^3}{f^2}E_0^2, \qquad \ldots(8)$$

which is roughly 10^{-8} at 1 km. from an antenna radiating 50kw of power isotropically at a frequency of 1 mc/s. Thus in equation (4) the term due to magnetic field of the wave is negligible. and

---

*$\frac{E}{H}=120\pi=\sqrt{(\mu_0/\epsilon_0)}=\mu_0\frac{1}{\sqrt{(\mu_0\epsilon_0)}}=\mu_0 c$

Therefore, $E=\mu_0 Hc=Bc$

$\frac{dy}{dt}$ *is exclusively due to electric field of the wave.*

Conductivity of ionised media : Out of 10^{17} ions per metre^3, 10^{11} are electrons, the current in plasma is chiefly due to electrons. Suppose N is number density of electrons in plasma, the current density resulting from their motion is given by

$$\mathbf{J} = Ne\,\mathbf{v} \qquad \ldots(9)$$

Since, ion drift velocity is almost exclusively due to electric field intensity **E** of the wave, we should put

$$v = \frac{dy}{dt},$$

so that convection current is

$$\mathbf{J} = Ne\left(\frac{dy}{dt}\right)$$

$$= Ne.\frac{eE_0}{\omega m}\sin \omega t.$$

Also interms of conductivity

$$\mathbf{J} = \sigma \mathbf{E}$$

$$= \sigma E_0 \cos \omega t.$$

Equating the two, we get

$$\sigma E_0 \cos \omega t = \frac{Ne^2 E_0}{m\omega} \sin \omega t$$

$$\sigma E_0\, e^{j\omega t} = \frac{Ne^2\, E_0}{m\omega}\; e^{j(\omega t - \pi/2)}$$

$$\sigma = \frac{-j}{\omega}.\frac{Ne^2}{m}. \qquad \ldots(10)$$

which implies that electron current, **J**, *lags* the electric field intensity, **E**, by $\pi/2$ rad. Therefore electron current is said to be inductive. Product **E.J** will be zero and *consequently, with collisions being absent, there is no energy loss in ionised medium or the oscillating electrons do not absorb energy from the passing electromagnetic wave.*

Conditions of wave propagation : To discuss the propagation of the wave, we write Maxwell's curl equations :

$$\nabla \times \mathbf{E} = -j\omega\mu_0 \mathbf{H} \qquad \ldots(11)$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}$$

$$= -j\,\frac{Ne^2\mathbf{E}}{m\omega} + \frac{\partial}{\partial t}\,(\epsilon_0\, E_0\, e^{j\omega t})$$

$$= j\omega\,\epsilon_0\, E - \frac{jNe^2\mathbf{E}}{m\omega}$$

$$= j\omega\,\epsilon_0 \mathbf{E}\left(1 - \frac{Ne^2}{m\omega^2\epsilon_0}\right) \qquad \ldots(12)$$

Since free electrons and heavy positive particles are distributed with statistical uniformity in the ionised region, the net space charge is zero so that

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} = 0. \qquad \ldots(13)$$

Taking the curl of eq. (11), we write

$$\nabla \times \nabla \times \mathbf{E} = -j\omega\mu_0 (\nabla \times \mathbf{H})$$

$$\nabla (\nabla . \mathbf{E}) - \nabla^2 \mathbf{E} = -j\omega\mu_0 . j\omega\epsilon_0 \mathbf{E} \left(1 - \frac{Ne^2}{m\omega^2\epsilon_0} \right)$$

$$\nabla^2 \mathbf{E} = -\omega^2\mu_0\epsilon_0 \left(1 - \frac{Ne^2}{m\omega^2\epsilon_0} \right) \mathbf{E}, \qquad \ldots(14)$$

on using eqs. (12) and (13).

We now define a quantity known as the plasma frequency, f_N, as

$$f_N = \frac{1}{2\pi} \sqrt{\left(\frac{Ne^2}{m\epsilon_0} \right)} = \sqrt{(80{\cdot}6\ N)}. \qquad \ldots(15)$$

where f_N is in c/s, and N in electrons per cubic metre. Plasma frequency is simply another way of specifying the electron density in the plasma. Putting $N = 10^{11}$/metre^3, we find $f_N \approx 3mc/s$. Putting eq. (15) into eq. (14), we get

$$\nabla^2 \mathbf{E} = -\omega^2\mu_0\epsilon_0 \left(1 - \frac{\omega_N^2}{\omega^2} \right) \mathbf{E}$$

$$= -\omega^2\mu_0\epsilon_0 \left(1 - \frac{f_N^2}{f^2} \right) \mathbf{E}.$$

But from wave equation for E, we know that

$$\nabla^2 \mathbf{E} = \gamma^2 \mathbf{E},$$

we find that propagation constant, on comparing the two, is

$$\gamma^2 = -\omega^2\mu_0\epsilon_0 \left(1 - \frac{f_N^2}{f^2} \right)$$

or

$$\gamma = \alpha + j\beta = j\omega \sqrt{(\mu_0\epsilon_0)} \left(1 - \frac{f_N^2}{f^2} \right)^{1/2},$$

we note that if

$f > f_N$ γ is purely imaginary $(=j\beta)$ and the wave is propagated.

$f < f_N$ γ is purely real $(=\alpha)$ and there is no phase change as the wave moves *i.e.* no wave motion. In other words, the fields are attenuated.

Thus for propagating range of frequencies $f > f_N$ (frequency of the wave should be greater than 3 mc/s if $N \approx 10^{11}$ per metre^3) and the phase constant is

$$\beta = \omega \sqrt{(\mu_0\epsilon_0)} \left(1 - \frac{f_N^2}{f^2} \right)^{1/2}. \qquad \ldots(16)$$

Phase and Group velocities : Phase velocity is given by

$$v_p=\frac{\omega}{\beta}=\frac{1}{\sqrt{(\mu_0\epsilon_0)}\left(1-\frac{f_N^2}{f^2}\right)^{1/2}}=\frac{c}{\sqrt{\left(1-\frac{f_N^2}{f^2}\right)}}$$

$$=\frac{c}{\sqrt{\left(1-\frac{80\cdot 6N}{f^2}\right)}}, \text{ using eq. (15)}$$

...(17)

The group velocity with which the signal is transmitted is given by

$$v_g=\frac{d\omega}{d\beta}$$

From eq. (16),

$$\beta=\omega\sqrt{(\mu_0\epsilon_0)}\left(1-\frac{f_N^2}{f^2}\right)^{1/2}$$

$$=\frac{\omega}{c}\left(1-\frac{\omega_N^2}{\omega^2}\right)^{1/2},$$

so that

$$\frac{d\beta}{d\omega}=\frac{1}{c}\left(1-\frac{\omega_N^2}{\omega^2}\right)^{1/2}+\frac{1}{2}\frac{\omega}{c}\left(1-\frac{\omega_N^2}{\omega^2}\right)^{1/2}\times\frac{2\omega_N^2}{\omega^3}$$

$$=\frac{1}{c}\left(1-\frac{\omega_N^2}{\omega^2}\right)^{1/2}+\frac{\omega_N^2}{\omega^2 c}\left(1-\frac{\omega_N^2}{\omega^2}\right)^{-1/2}$$

$$=\frac{1}{c\left(1-\frac{\omega_N^2}{\omega^2}\right)^{1/2}}$$

or $$\frac{d\omega}{d\beta}=c\sqrt{\left(1-\frac{\omega_N^2}{\omega^2}\right)}=c\sqrt{\left(1-\frac{f_N^2}{f^2}\right)}.$$

Thus $$v_g=\frac{d\omega}{d\beta}=c\sqrt{\left(1-\frac{f_N^2}{f^2}\right)}.$$

Note that $v_p\, v_g=c^2$,

where c is the velocity of light in the free space.

Refractive Index : The refractive index, μ, (not to be confused with permeability of space) is

$$\mu=\frac{c}{v_p}=\sqrt{\left(1-\frac{80\cdot 6N}{f^2}\right)}, \qquad ...(18)$$

N is in per metre^3. Suppose N be in per cm^3 and f be in kc/s., then

$$\mu=\sqrt{\left(1-\frac{80\cdot 6N.10^{-6}}{(f/10^3)^2}\right)}$$

$$=\sqrt{\left(1-\frac{80\cdot 6N}{f^2}\right)}.$$

That is, same relation as eqn. (18) is thus retained. But note that f is in kc/s. Let us express μ in c.g.s. system also. From eqn. (17),

$$\mu=\sqrt{\left(1-\frac{f_N^2}{f^2}\right)}=\sqrt{\left(1-\frac{Ne^2}{m\omega^2\epsilon_0}\right)}$$

$$=\left(1-\frac{4\pi Ne^2}{4\pi\epsilon_0 m\omega^2}\right)^{1/2}$$

We can show that

$$\frac{1}{4\pi\epsilon_0}\cdot\frac{Ne^2}{m} \text{ in } mks=\frac{Ne^2}{m} \text{ in c.g.s.,}$$

so that, we can write

$$\mu=\left(1-\frac{4\pi Ne^2}{m\omega^2}\right)^{1/2} \text{ in c.g.s.} \qquad \text{.. (19)}$$

In next articles, we shall be using this equation quite frequently.

Let us now explain that how the ionosphere is instrumental in achieving long distance comunication : We divided the ionised region into several infinitesimal slabs, in each of which the electron density can be considered to be uniform with height. Let us consider the boundary between free space and first slab, for which plasma frequency is $f_{N,\,i}$ say. The phase velocity for this slab is

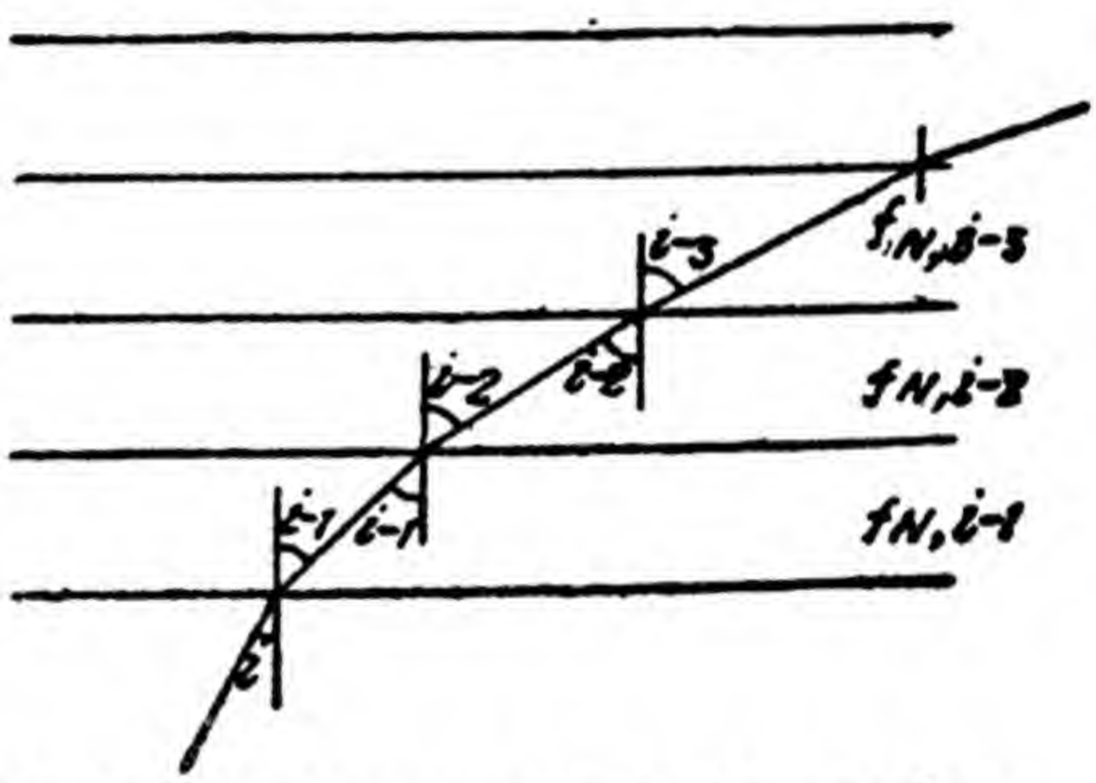


Fig. 9. As there are infinite number of slabs, wave path is no longer a series of straight lines but a continuous curve.

$$v_{p,\,i-1}=\frac{c}{\sqrt{\left(1-\frac{f^2_{N,\,i-1}}{f^2}\right)}}.$$

Also $$\mu_{i-1}=\frac{\sin i}{\sin (i-1)},$$

so that $$\mu_{i-1}\sin (i-1)=\sin i$$

or $$\frac{c}{v_{p,\,i-1}}\sin (i-1)=\frac{c}{c}\sin i.$$

So, in general, we can write for any number of slabs, that

$$\sin i=\frac{c}{v_{p,\,i-1}}\sin (i-1)=\frac{c}{v_{p\ i-2}}\sin (i-2)=\ldots$$

$$\sin i=\mu_{i-1}\sin (i-1)=\mu_{i-2}\sin (i-2)=\ldots \qquad \ldots(20)$$

We know that

$$\mu=\sqrt{\left(1-\frac{80\cdot 6N}{f^2}\right)},$$

so that

$$\mu_{i-2}<\mu_{i-1}<\mu_i \text{ because } N_{i-2}>N_{i-1}>N_i$$

But in order to satisfy equation (2), we must then have

$$\text{angle } (i-2) > \text{angle } (i-1) > \text{angle } (i),$$

which implies that the path of the wave bends gradually away from the normal to the boundary. Finally, a level may be reached at which the electron density is such that angle of refraction is 90°, so that

$$\mu=\frac{\sin i}{\sin 90^\circ}=\sin i,$$

and the path is horizontal. Thus for horizontal path

$$\sqrt{\left(1-\frac{f_N^2}{f^2}\right)}=\sin i$$

or

$$f_N=f\cos i. \qquad \ldots(21)$$

Thus a wave of frequency f incident at an angle i with the normal to the boundary is reflected from a level at which plasma frequency is equal to $f\cos i$. For special case of normal incidence on the ionosphere, $i=0$ (or $\mu=0$), and the condition for reflection is

$$f_N=f=f_c. \qquad \ldots(22)$$

f_c is called the critical frequency. The wave is then reflected from a level at which plasma frequency is equal to the wave frequency. If we put $\mu=0$ in equation (19), we find that

$$f_c=f_N=\frac{Ne^2}{m\pi} \text{ in c.g.s. system.} \qquad \ldots(23)$$

When frequency of the wave is sufficiently high, say light frequencies, then we find that, on taking N equal to 10^{11} per metre^3,

$$\frac{80{\cdot}6N}{f^2}=\frac{80{\cdot}6.10^{11}}{(10^{14})^2}=80{\cdot}6.10^{-17}.$$

and therefore the term is negligible, so that

$$\mu\simeq 1,$$

or on the propagation of such waves, there is no effect of ionised medium. They will penetrate the ionosphere at the same angle at which they are sent.

Effect of Collisions : Let us here introduce the effect of collisions. At the high altitudes prssure is low so that collisions are not frequent, but at the lower edge of ionosphere, where pressure is high, collisions can not be neglected. We have seen that electrons oscillate under the action of electric field of the incident wave ; therefore accounting for the motion in the direction of E vector, we write equation of motion of electron when collisions are present as

$$m\frac{d^2y}{dt^2}=eE-\xi\frac{dy}{dt}, \qquad \ldots(24)$$

where ξ is the collision factor. Minus sign is introduced as their effect is to cause absorption of energy from the electromagnetic wave by destroying in part the direct momentum of vibrating electrons. Further, putting $v=dy/dt$, we get

$$m\frac{dv}{dt}=eE-\xi.v. \qquad \ldots(25)$$

Since electrons are oscillating,

$$v=v_0\, e^{i\omega t},$$

so that

$$\frac{dv}{dt}=j\omega v.$$

Putting in equation (25) we arrive at

$$mj\omega v=eE-\xi v$$

or

$$v=\frac{eE}{(jm\omega+\xi)},$$

so that cor v ction current density will be

$$J=Nev=\frac{Ne^2E}{(\xi+jm\omega)}$$

$$=\frac{Ne^2E\,(\xi-jm\omega)}{(\xi^2+m^2\,\omega^2)}$$

$$=\frac{Ne^2E\,\xi}{(\xi^2+m^2\,\omega^2)}-\frac{jm\omega.Ne^2E}{(\xi^2+m^2\omega^2)} \qquad \ldots(26)$$

For second term on right hand side E.J is zero because both have a phase of $\pi/2$ but for first term E.J will not be zero because both are in phase. This means, unlike to case discussed after equation (10), we find here that oscillating electrons absorb energy from the electromagnetie wave.

In next article, we shall account for the effect of collisions and of earth's magnetic field on the propagation of radiowaves through ionised region.

29·4. PROPAGATION IN THE PRESENCE OF MAGNETIC FIFLD—THE MAGNEIO-IONIC THEORY-APPLETON HARTREE FORMULA :

The actual manner in which the charged particles affect the propagation of radio waves was first pointed out by Eccles in 1912. Some essential details which are lacking in Eccles theory and for which it lay undevelop:d were supplied by Larmor in 1924. The Eccles-Larmor theory is now regarded as the basic theory of radio wave propagation in the ionosphere and has been supplemented by the so called magneto ionic theory developed by Appleton, Hartree, Goldstein and others

We shall now discuss the theory of propagation of electromagnetic waves through an ionised medium in the

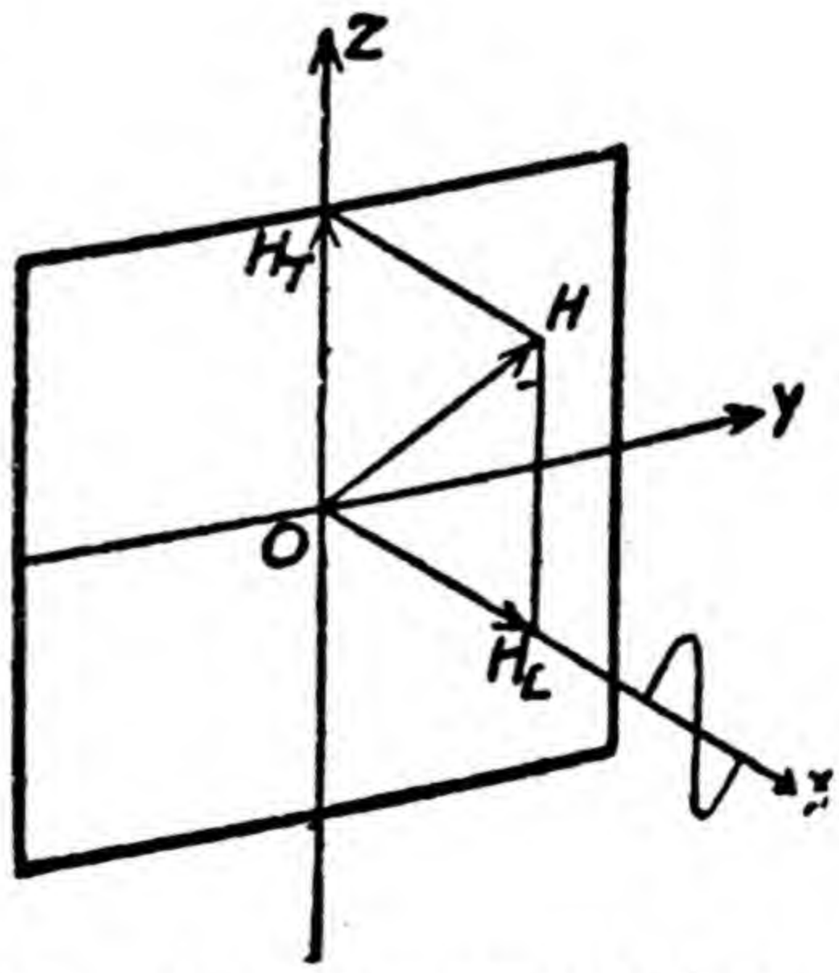


Fig. 10. (*a*) Illustrates direction of propagation of electromagnetics waves.

presence of external magnetic field. This theory is known as magneto ionic theory. Appleton in 1925 first applied it to explain the phenomenon of propagation of radio waves through the upper ionised regions of the earth's atmosphere. Later Hartree derived the formula independently

The earth's magnetic field exerts a *force* on the moving electrons. This *force* is proportional to the instantaneous velocity of the electron and to the component of magnetic field at right angles to the direction of motion†. The direction of this force is right angles to both the direction of motion of electron and component of magnetic field. The path of the electron is thus altered. It is elliptical at higher radio frequencies. As the frequency is reduced ratio of minor axis to the major axis increases until at frequency 1·4 mc/s. called gyromagnetic frequency, $(=eH/mc)$ electron vibrates in spiral. At still lower frequencies, electron vibrates in loops as shown in figure 10 (*b*).

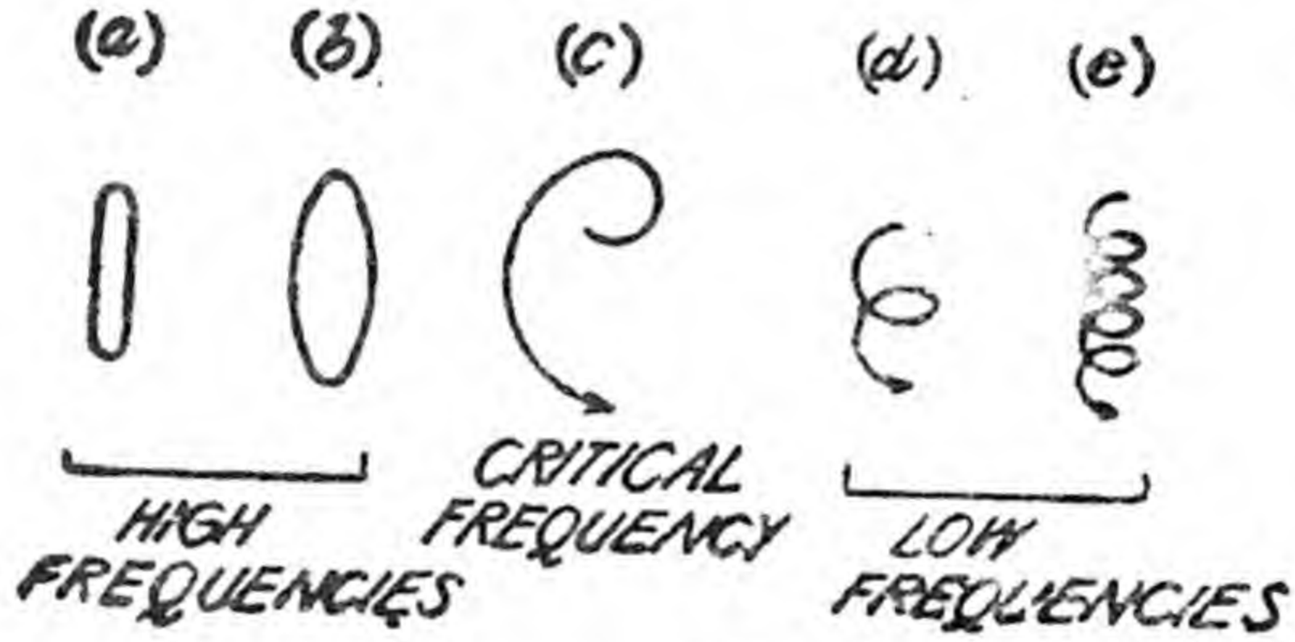


Fig. 10(*b*)

Consider an electro-magnetic wave (plane polarized) moving along the positive direction of X-axis of a right handed co-ordinate system in the ionosphere. There is a possibility of collision of electrons with gaseous atoms or another electrons. The ultimate effect of which is a type of frictional force of coefficient of friction, say ξ. Let the magnetic field be acting in XZ plane. Then we shall have

$$H_X=H_L,\ H_Y=0, \text{ and } H_Z=H_T.$$

Now if m is the mass of electron, and v the velocity of propagation, e the charge on electron, then the equations of motion of the electron in the field of advancing elecromagnetic wave in c.g.s. system are

$$m\frac{d^2x}{dt^2}=eE_X-\xi\frac{dx}{dt}+\frac{e}{c}[v\times H]_X{}^*. \qquad \ldots(1)$$

†It will exert a force equal to $e(v\times H)/c$.

$$^*[v\times H]=\begin{vmatrix} i & j & k \\ dx/dt & dy/dt & dz/dt \\ H_L & 0 & H_T \end{vmatrix}.$$

so that $[v\times H]_X=H_T\frac{dy}{dt}$,

$$[v\times H]_Y=H_L\frac{dz}{dt}-H_T\frac{dx}{dt},$$

$$[v\times H]_Z=H_L\frac{dy}{dt}$$

$$m\frac{d^2y}{dt^2}=eE-\xi_Y\frac{dy}{dt}+\frac{e}{c}[v\times H)_Y. \qquad ...(2)$$

$$m\frac{d^2z}{dt^2}=eE_Z-\xi\frac{dz}{dt}+\frac{e}{c}[v\times H]_Z. \qquad ..(3)$$

or

$$\left.\begin{aligned} m\frac{d^2x}{dt^2}&=eE_X-\xi\frac{dx}{dt}+\frac{e}{c}H_T\frac{dy}{dt}.\\ m\frac{d^2y}{dt^2}&=eE_Y-\xi\frac{dy}{dt}+\frac{e}{c}H_L\frac{dz}{dt}-\frac{e}{c}H_T\frac{dx}{dt}\\ m\frac{d^2z}{dt^2}&=eE_Z-\xi\frac{dz}{dt}-\frac{e}{c}H_L\frac{dy}{dt}\end{aligned}\right\}\ ...(4)$$

If ω is the angular frequency of the wave, then for a wave motion, value of the operator $\frac{d}{dt}=j\omega$ and $\frac{d^2}{dt^2}=-\omega^2$.

Then from equation (4), we have

$$\left.\begin{aligned} -m\omega^2x&=eE_X-j\omega\xi x+j\omega\frac{e}{c}H_Ty.\\ -m\omega^2y&=eE_Y-j\omega\xi y+j\omega\frac{e}{c}H_Lz-j\omega\frac{e}{c}H_Tx\\ -m\omega^2z&=eE_Z-j\omega\xi z-j\omega\frac{e}{c}H_Ly\end{aligned}\right\}\ ...(5)$$

Dividing by $4\pi Ne^2$ and puttting

$$\frac{-m\omega^2}{4\pi Ne^2}=A, \qquad \frac{\omega\xi}{4\pi Ne^2}=B.$$

$$\frac{\omega.\frac{e}{c}.H_T}{4\pi Ne^2}=C_T \quad \text{and} \quad \frac{\omega.\frac{e}{c}.H_L}{4\pi Ne^2}=C_L,$$

we get

$$\left.\begin{aligned} E_X&=(A+jB)\,4\pi Nex-jC_T\,(4\pi Ney), && (a)\\ E_Y&=(A+jB)\,4\pi Ney+jC_T\,(4\pi Nex)-jC_L\,(4\pi Nez) && (b)\\ \text{and } E_Z&=(A+jB)\,4\pi Nez+jC_L\,(4\pi Ney) && (c)\end{aligned}\right\}\ ...(6)$$

Since the plane wave is travelling in positive X-direction, we can arrive at the conclusion that

$$\frac{\partial}{\partial y}=\frac{\partial}{\partial z}=0, \text{ but } \frac{\partial}{\partial x}\neq 0.$$

Maxwell's electromagnetic equation for conducting medium can be written as

$$\text{curl } H=\frac{1}{c}\frac{\partial E}{\partial t}+\frac{4\pi\sigma}{c}E.$$

Since current density $J=\sigma E=Nev$, therefore

$$\text{curl } H=\frac{1}{c}\frac{\partial E}{\partial t}+\frac{4\pi Nev}{c}. \qquad ...(7)$$

The components of curl H in different directions are

$$\left.\begin{aligned}
[\text{curl } H]_X&=\frac{1}{c}\left[\frac{\partial E_X}{\partial t}+4\pi Ne\frac{\partial x}{\partial t}\right]=0.\\
[\text{curl } H]_Y&=\frac{1}{c}\left[\frac{\partial E_Y}{\partial t}+4\pi Ne\frac{\partial y}{\partial t}\right]=-\frac{\partial H_Z}{\partial x},\\
[\text{curl } H]_Z&=\frac{1}{c}\left[\frac{\partial E_Z}{\partial t}+4\pi Ne\frac{\partial z}{\partial t}\right]=\frac{\partial H_Z}{\partial x}.
\end{aligned}\right\} \qquad ...(8)$$

The plane wave can be taken as $e^{j\omega\{t-(x\sqrt{\epsilon})/c\}}$, so that operators are

$$\frac{\partial}{\partial x}=-j\omega\frac{\sqrt{\epsilon}}{c} \text{ and } \frac{\partial}{\partial t}=j\omega;$$

which when substituted in eq. (8) gives

$$\left.\begin{aligned}
&\frac{1}{c}\,j\omega E_X+\frac{4\pi Ne}{c}\,j\omega x=0,\\
&\frac{1}{c}\,j\omega E_Y+\frac{4\pi Ne}{c}\,j\omega y=+j\omega\frac{\sqrt{\epsilon}}{c}H_Z.\\
&\frac{1}{c}\,j\omega E_Z+\frac{4\pi Ne}{c}\,j\omega 2=-j\omega\frac{\sqrt{\epsilon}}{c}H_Y.
\end{aligned}\right\} \qquad ...(9)$$

Above equations become on regarranging

$$\left.\begin{aligned}
E_X&=-4\pi Nex.\\
E_Y&=-4\pi Ney+\sqrt{\epsilon}\,H_Z\\
E_Z&=-4\pi Nez-\sqrt{\epsilon}\,H_Y.
\end{aligned}\right] \qquad ...(10)$$

The other field equation is

$$\text{curl } E=-\frac{1}{c}\frac{\partial H}{\partial t}.$$

Writing components of curl E, we have

$$(\text{curl } E)_X=\frac{\partial E_Z}{\partial y}-\frac{\partial E_Y}{\partial z}=-\frac{1}{c}\frac{\partial H_X}{\partial t},$$

$$(\text{curl } E)_Y=\frac{\partial E_X}{\partial z}-\frac{\partial E_Z}{\partial x}=-\frac{1}{c}\frac{\partial H_Y}{\partial t}.$$

$$(\text{curl } E)_Z=\frac{-\partial E_X}{\partial y}+\frac{\partial E_Y}{\partial x}=-\frac{1}{c}\frac{\partial H_Y}{\partial t}.$$

In our case, $\frac{\partial}{\partial y}=\frac{\partial}{\partial z}=0$, so that

$$\text{curl } E=\begin{vmatrix} i & j & k \\ \partial/\partial x & 0 & 0 \\ E_X & E_Y & E_Z \end{vmatrix}.$$

which can be written as

$$-\frac{1}{c}\frac{\partial}{\partial t}(H_X)=0,$$

$$-\frac{1}{c}\frac{\partial}{\partial t}(H_Y)=-\frac{\partial E_Z}{\partial x}.$$

$$-\frac{1}{c}\frac{\partial}{\partial t}(H_Z)=\frac{\partial E_Y}{\partial x}.$$

Now we know that

$$\frac{\partial}{\partial t}=j\omega \text{ and } \frac{\partial}{\partial x}=-j\omega\frac{\sqrt{\epsilon}}{c},$$

so that above relations become

$$\left.\begin{aligned} -\frac{1}{c}j\omega H_X&=0 \qquad\qquad H_X=0 \\ -\frac{1}{c}j\omega H_Y&=j\omega\frac{\sqrt{(\epsilon)}}{c}E_Z \quad \text{or} \quad H_Y=-\sqrt{(\epsilon)}\,E_Z, \\ -\frac{1}{c}j\omega H_Z&=-j\omega\frac{\sqrt{(\epsilon)}}{c}E_Y \quad \text{or} \quad H_Z=\sqrt{(\epsilon)}\,E_Y. \end{aligned}\right\} \qquad \ldots(11)$$

We can [from eqs. (10) and (11)] show that, components of E, are connected to dielectric constant by the relations

$$\left.\begin{aligned} E_X&=-4\pi Nex, \\ (\epsilon-1)\,E_Y&=4\pi Ney \\ (\epsilon-1)\,E_Z&=4\pi Nez \end{aligned}\right\} \qquad \ldots(12)$$

From eq. (12) and eqs. 6 (*a*), we have

$$4\pi Nex\,(1+A+jB)=jC_T.4\pi Ney,$$

or

$$4\pi Nex=\frac{jC_T 4\pi Ney}{1+A+jB} \qquad \ldots(13)$$

From eq. (12) and eq 6 (*b*), we get

$$\frac{4\pi Ney}{\epsilon-1}-(A+jB)\,4\pi Ney=j\,(4\pi\,Nex)\,C_T-jC_L 4\pi Nez.$$

Putting for x from eq. (13) in above equation, we get

$$\frac{4\pi Ney}{\epsilon-1}-(A+jB)\,4\pi Ney+\frac{C_T^2 4\pi Ney}{(1+A+jB)}=-jC_L.4\pi Nez,$$

or

$$4\pi Ney\left\{\frac{1}{\epsilon-1}-(A+jB)+\frac{C_T^2}{(1+A+jB)}\right\}=-jC_L.4\pi Nez, \qquad \ldots(14)$$

Further from eq (12) and eq. 6 (*c*), we get

Multiplying eqs. (14) and (15), we arrive at

$$\left[\frac{1}{\epsilon-1}-(A+jB)\right]^2+\frac{C_T^2}{(1+A+jB)}\left[\frac{1}{\epsilon-1}-(A+jB)\right]-C_L^2=0.$$

This is quadratic in $\left[\frac{1}{\epsilon-1}-(A+jB)\right]$ **so that we can solve it to give**

$$\epsilon=1+\frac{1}{(A+jB)-\frac{C_T^2}{2(1+A+jB)}\pm\sqrt{\left(\frac{C_T^4}{4(1+A+jB)^2}+C_L^2\right)}}.$$

This is the *Appleton Hartree formula* **for the dielectric constant of medium in presence of magnetic field taking into account the collisions. It is seen that** ϵ **has got two values corresponding to the positive and negative signs involved in the formula and the incident wave will be split up into two waves—*ordinary and extra-ordinary*. Hence the medium acts as *doubly refracting crystal.***

Special Cases :

(*A*) *No external magnetic field* :

When $H=0$ *i.e.* $C_T=C_L=0$, we find

$$\epsilon=1+\frac{1}{A+jB}$$

$$=1+\frac{1}{\frac{-m\omega^2}{4\pi Ne^2}+\frac{j\,B\omega}{4\pi Ne^2}}=1+\frac{4\pi Ne^2}{-m\omega^2+j\omega B},$$

$$\epsilon=1+\frac{4\pi Ne^2}{-m\omega^2+j\omega B}.$$

i.e. dielectric constant .of ionosphere **depends upon the electron number density** N **at a particular frequency.**

(*B*) *Longitudinal propagation* :

When component $H_T=0$ *i.e.* when the magnetic field is along the direction of propagation $H_L\neq0$, $C_L\neq0$ but $C_T=0$, so that

$$\epsilon=1+\frac{1}{(A+jB)\pm C_L}$$

$$=1+\frac{1}{\left(\frac{-m\omega^2}{4\pi Ne^2}+\frac{j\xi\omega}{4\pi Ne^2}\right)\pm\frac{e.H_L\omega}{c.4\pi Ne^2}}$$

If $\xi=0$, neglecting collisions,

$$\epsilon=1+\frac{1}{\frac{-m\omega^2}{4\pi Ne^2}\pm\frac{c.H_L\omega}{c.4\pi Ne^2}}$$

$$=1+\frac{1}{\frac{-m\omega^2}{4\pi Ne^2}\pm\frac{\omega H_L}{c4\pi Ne}}$$

This formula shows that the dielectric constant depends upon the magnetic field, density of electrons and on the frequency of the incident wave. Thus we find that there is dispersion.

(*C*) *Transverse propagation* :

In this case $H_T \neq 0$, $H_L = 0$, $C_L = 0$, $C_T \neq 0$,

$$\epsilon = 1 + \frac{1}{A - \frac{C_T^2}{2(1+A)} \pm \frac{C_T^2}{2(1+A)}}.$$

SPLITTING OF RADIO WAVES IN IONOSPHERE :

The splitting of radio waves, as they travel upwards in the ionosphere, is best explained by the Appleton Hartree ionic formula. According to this formula, refractive index is

$$\mu^{*2} = \epsilon = 1 + \frac{1}{(A+jB) - \frac{C_T^2}{2(1+A+jB)} \pm \sqrt{\left\{\frac{C_T^4}{4(1+A+jB)^2} + C_L^2\right\}}}$$

For simplicity, we assume that $B=0$; then we have

$$\mu^{*2} = 1 \pm \frac{1}{A - \frac{C_T^2}{2(1+A)} + \sqrt{\left\{\frac{C_T^4}{4(1+A)^2} + C_L^2\right\}}}$$

For explaining the splitting of radio waves, we must note that for the condition of vertical reflection $\mu^* = 0$, so that

$$A - \frac{C_T^2}{2(1+A)} \pm \sqrt{\left\{\frac{C_T^4}{4(1+A)^2} + C_L^2\right\}} = -1,$$

$$(1+A) - \frac{C_T^2}{2(1+A)} = \pm \sqrt{\left\{\frac{C_T^4}{4(1+A)^2} + C_L^2\right\}},$$

$$(1+A)^2 - \frac{C_T^2}{2} = \pm \sqrt{\left\{\frac{C_T^4}{4} + C_L^2 (1+A)^2\right\}}.$$

Squaring both the sides,

$$(1+A)^4 + \frac{C_T^4}{4} - (1+A)^2 C_T^2 = \frac{C_T^4}{4} + C_L^2 (1+A)^2,$$

$$(1+A)^4 - (1+A)^2 [C_T^2 + C_L^2] = 0,$$

$$(1+A)^2 [(1+A)^2 - (C_T^2 + C_L^2)] = 0.$$

Either $(1+A)^2 = 0$ ...(1)

or $[(1+A)^2 - (C_T^2 + C_L^2)] = 0.$ ...(2)

Let $C_T^2 + C_L^2 = \gamma^2$;

then we get $1+A=0,$ ...3(*a*)

$1+A = \pm\gamma.$ ...3(*b*) and 3(*c*)

The first condition 3(*a*) of reflection is independent of the earth's magnetic field.

Now let us draw dispersion curves as in figure 11, showing

relation between μ^{*2} and $x=-\frac{1}{A}\left(=\frac{4\pi Ne^2}{m\omega^2}=\text{constant}.\lambda^2\right)$.

A study of dispersion curves discloses the remarkable fact that the value of $x=-\frac{1}{A}$, for which μ^{*2} becomes equal to zero, is independent of the inclination of magnetic field to the direction of propagation. Now for any inclination of magnetic field to the direction of wave propagation, it is seen that the refractive indices are always brought down to zero value at E_1, O and E_2. The electron number density corresponding to the point O is given by $1+A=0$, *i.e.*, the condition of reflection represented by this point is the same as that for the case of no magnetic field. The ray which undergoes reflection at this point is called *ordinary ray*. The ordinary ray of which the refractive index varies along the curves BE_1 and FE_2 and attains zero value at the point $1+A=\pm\gamma$ is called *extra-ordinary ray* since its condition of reflection is affected by the magnetic field. However, the analogy cannot be pursued because there are cases in which both ordinary and extra-ordinary rays may be reflected from the same height.

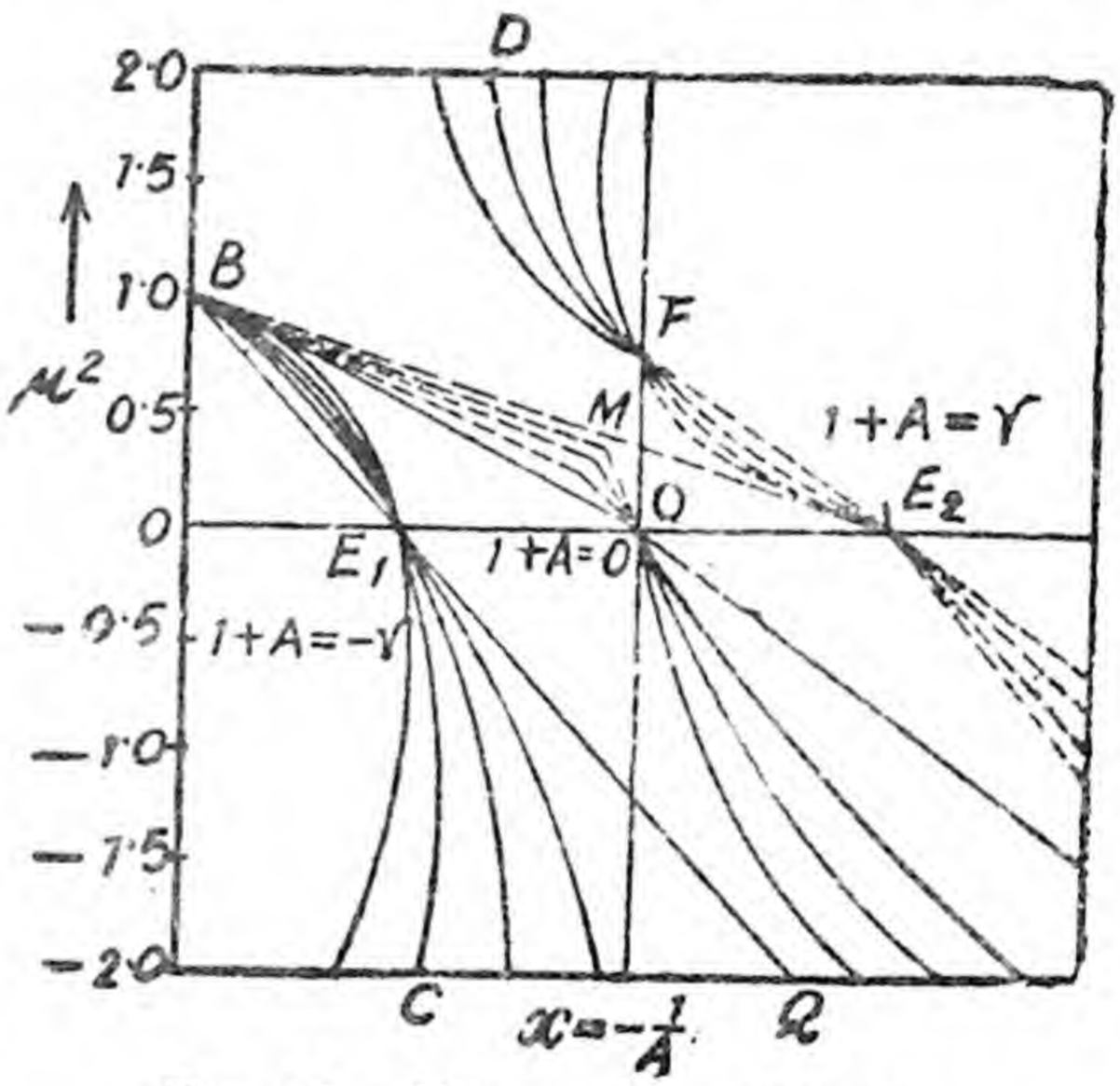


Fig. 11. Dispersion curves.

STATE OF POLARISATION :

According to Appleton Hartree formula,

$$\epsilon=\mu^{*2}=1+\frac{1}{(A+jB)-\frac{C_T^2}{2(1+A+jB)}\pm\sqrt{\left\{\frac{C_T^4}{4(1+A+jB)^2}+C_L^2\right\}}}$$

where μ^{*2} is complex refractive index. From above, the state of polarisation is

$$R=\frac{H_z}{H_y}=\frac{-j}{C_L}\left\{\frac{1}{\mu^{*2}-1}-(A+jB)\right\},$$

where H_z is z-component of magnetic vector of wave and H_y is the y-component of magnetic vector of wave. If $B=0$, then

$$R=-\frac{i}{C_L}\left\{\frac{1}{\mu^{*2}-1}-A\right\}.$$

Again by putting $x=-\frac{1}{A}$,

$$R=-j\{f\pm\sqrt{(1+f^2)}\}. \qquad ...(A)$$

where $f=\beta\frac{1}{1-x}$, $\beta=\frac{y\sin^2\theta}{2\cos\theta}$, θ is the angle between the direction of propagation and the positive direction of the magnetic field.

The sign before the expression for R has been made negative because the charged particles in the ionosphere which influence the propagation of waves are supposed to be electrons. R thus signifies the ratio between the components of the magnetic vector of the wave in and at right angles to the plane containing the wave normal and the direction of the magnetic field hereinafter called the magnetic plane.

When R is zero, H_z is zero and the wave is plane polarised with its magnetic vector perpendicular to the magnetic plane. When R is infinite, H_y is zero, and the wave is again plane polarised, but with its magnetic vector in the magnetic plane. When R is numerically equal to unity, the magnitudes of H_z and H_y are equal and the wave is circularly polarised.

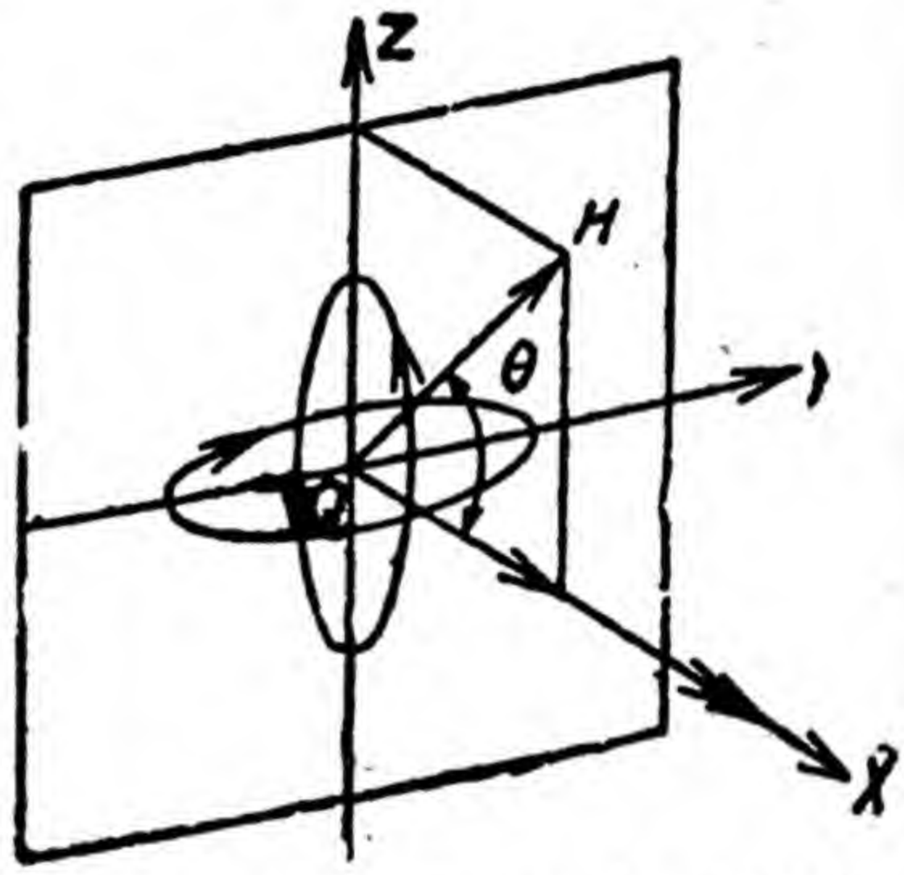


Fig. 12. Illustrating elliptic polarisation of the magnetically split components.

We will describe here the nature of polarization of the various types of waves that are propagated in the ionosphere.

(*i*) *Longitudinal case* : In this case, $\theta=0$ and $R=\pm j$.

For $R=j$ the components (H_y, H_z) are equal and have a phase difference of $\pi/2$, while in other case when $R=-j$ the phase difference is $-\pi/2$, so that two waves into which incident wave splits are both circularly polarised but in opposite sense.

(*ii*) *Transverse case* : Here, putting $\theta=\pi/2$, we find R to be equal to $-j\infty$ or 0 ; these are interpreted as states of plane polarisation. For transverse propagation, therefore, the two component waves into which an incident wave is split up are, in contrast to the longitudinal case, both plane polarised.

(*iii*) *General case* : The state of polarization for any direction of propagation other than longitudinal or transverse is in general elliptic. The two ellipses, corresponding to the two split

components, have the same ratio as of the major to the minor axis. They are, however, oriented at right angles to each other and their senses of rotation are opposite.

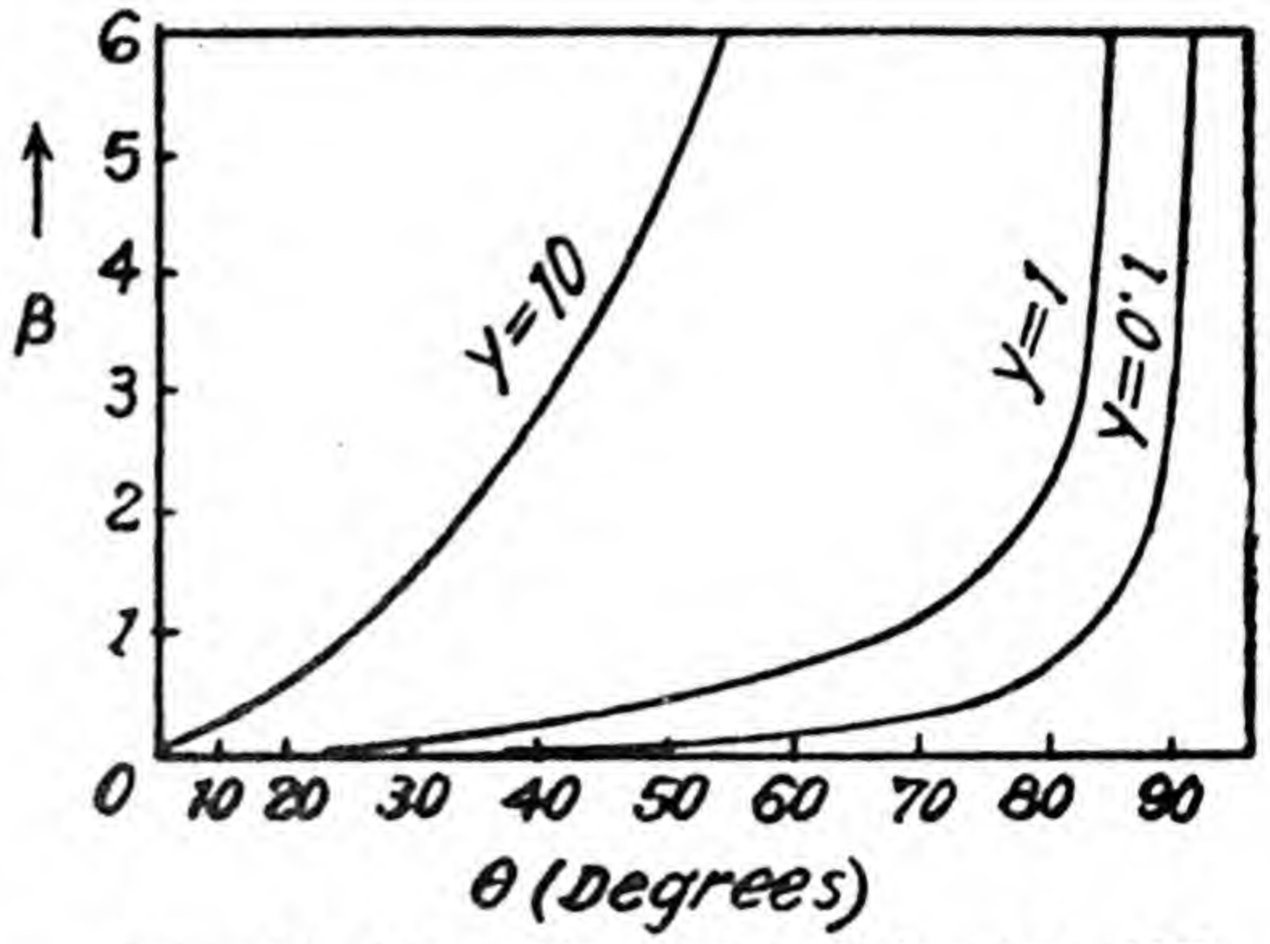


Fig. 13. Variation of β with θ for different values of y.

In order to determine the polarization, we have to take all these factors into account. This may be done first by plotting graphs connecting.

$$\beta=\left(\frac{y\sin^2\theta}{2\cos\theta}\right) \text{ and } \theta$$

for various values of y (fig. 13) and then graphs connecting R and x for various values of β (fig. 14).

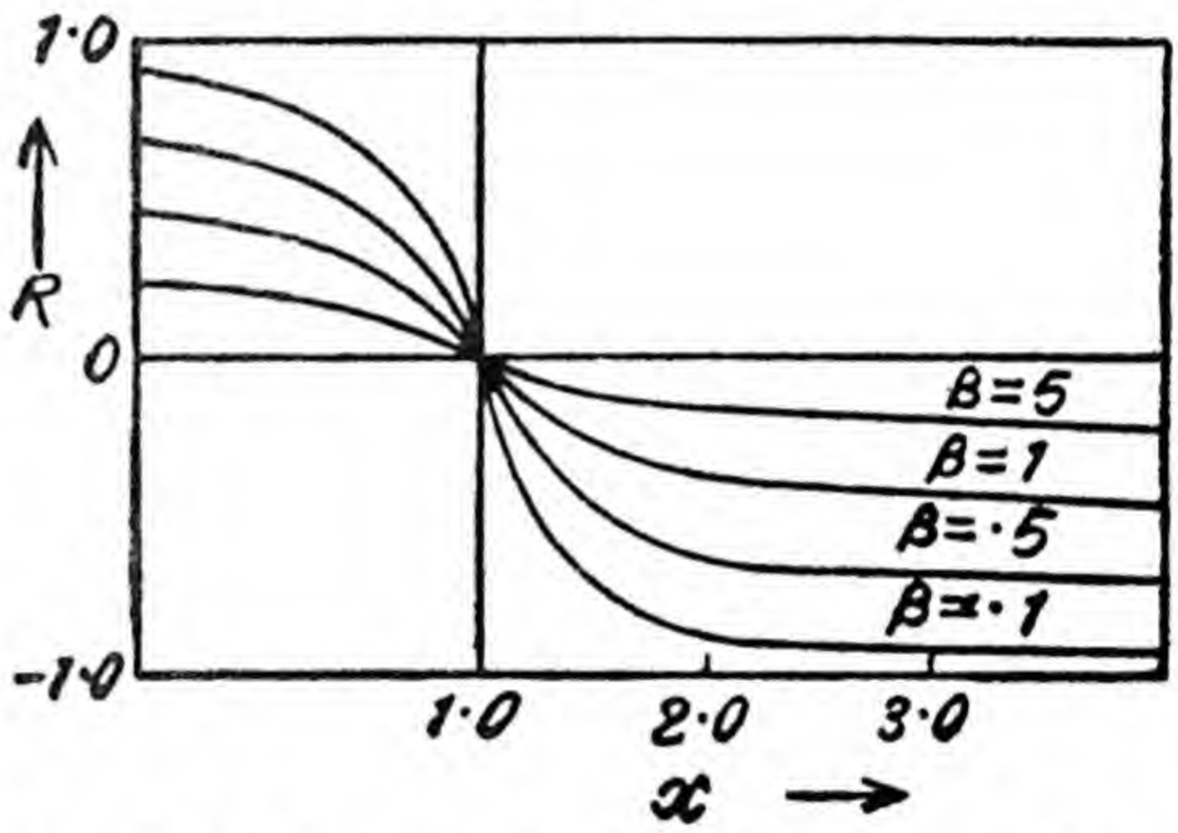


Fig 14. Illustrating changes in polarisation as the wave of a given frequency enters into regions of increasing electron number density.

As is evident from the figure 14, the state of polarization therefore changes as the wave travels through the regions of continually changing electron number density.

If the exploring wave reaches the region of reflection of the ordinary wave, where $N=\frac{m\omega^2}{4\pi e^2}$, *i.e.* $x=1$, the state of polarization changes sign, *i.e.* which was previously right handed becomes left-handed and *vice-versa*. The wave is plane polarized.

29·5. SKIP DISTANCE AND MAXIMUM USABLE FREQUENCY :

Now we shall discuss the considerations, *e.g.*, skip distance and maximum usable frequency which ought to be taken into account in deciding the frequency of transmission for short wave intercontinental communication.

We have already proved that refractive index is

$$\mu=\frac{\sin i}{\sin r}=\sqrt{\left(1-\frac{80\cdot6\ N}{f^2}\right)}$$

According to this relation, as the wave penetrates into regions of greater electron density, the angle of refraction increases corres-

pondingly. When μ has decreased to the point where $\mu = \sin i$, the angle of refraction r, will be 90° and the wave will be travelling horizontally. The highest point reached by the wave is therefore that point at which the electron density N satisfies the relation

$$\sqrt{\left(1-\frac{80{\cdot}6\,N'}{f^2}\right)}=\sin i \text{ or } N'=\frac{f^2\cos^2 i}{80{\cdot}6}. \quad \ldots(1)$$

If the electron density at some level in a layer is sufficiently great to satisfy relation (1), the wave will be returned to earth from that level. If the maximum electron density in a layer is less than that required by relation (1), the wave will penetrate the layer and may be reflected from the higher layer of greater ionisation density.

Critical frequency is the maximum frequency of the wave which is returned from a layer for normal incidence However, a wave having frequency greater than critical frequency can return to earth if it enters ionospheric layer at an oblique angle, say *i*. Thus in case the frequency of the wave exceeds the critical frequency, the influence of the ionospheric layer on the path of propagation depends upon the angle of incidence at the ionosphere. It is noted that as the angle of incidence at the ionosphere decreases, the distance from the transmitter, at which the ray returns to ground, first decreases and at a particular angle of incidence it becomes minimum which is called the *skip distance*. With further decrease in the angle of incidence, the distance to the point of return first increases and then with still smaller angles the wave penetrates the layer. This state of affairs can be explained as follows : with large value of angle of incidence, the relation (1) is satisfied with small value of electron density and therefore the wave is returned to earth only after slight penetration into the ionised media [ray (i)]. When the angle of incidence further decreases [as for rays (ii), (iii)] the relation (1) will be satisfied with some what greater value of electron density ; consequently the wave penetrates more into the ionised media and returns to a point on the ground more close to the transmitter. With further decrease in the angle of incidence, the wave will reach to the regions of higher electron density and near the regions of maximum electron density the change of electron density with height is relatively small ; due to this reason, the bending of the wave away from the normal will be gradually through small angles and the wave will travel almost horizontally (parallel to the layer) for a considerable distance. It will therefore return to a point on the ground which is at greater distance from the transmitter [ray (*iv*)]. Finally, when the angle of incidence is reduced to a very small value then relation (1) is not satisfied because such a high value of electron density is not available in the ionised media; consequently the wave penetrates the ionised media [ray (*v*)].

The minimum distance from the transmitter to the point on the ground at which sky wave of a given frequency will be returned to earth by the ionosphere is called skip distance. We shall show in equation (8) that skip distance depends upon :

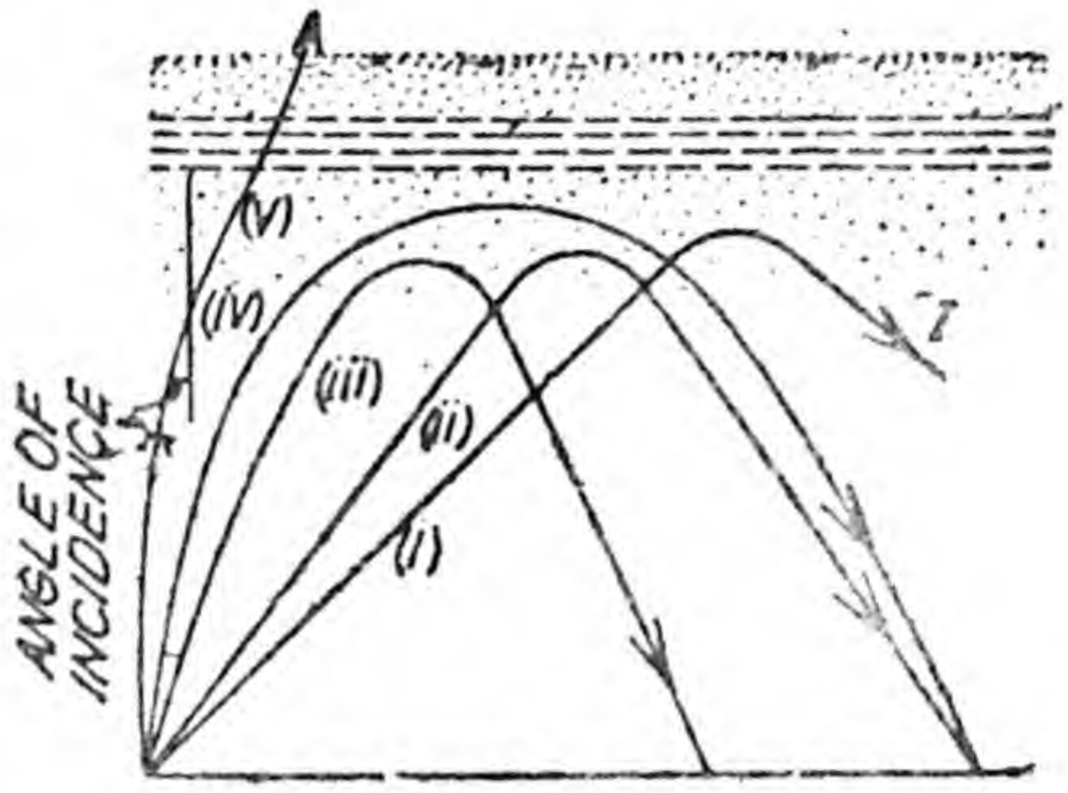


Fig 15. Illustrating the skip distance.

(*i*) f, the frequency of transmission,

(*ii*) f_c, the critical frequency of the layer and consequently on the electron density because $f_c = 80{\cdot}6\ N'$,

(*iii*) h, the height of the layer.

Since f_c, and h are the characteristics of any ionospheric layer and therefore each layer in the ionosphere has, for each frequency f, a skip distance. The smallest skip distance is sometimes for E layer and sometimes for F layer but rarely for F_1 layer.

As is obvious from forthcoming equation (8), skip distance increases with frequency provided height and critical frequency are kept constant. Therefore for a fixed value of skip distance, there would correspond a certain value of frequency. *The frequency which makes a given receiving point correspond to a aistance from the transmitter equal to the skip distance for that frequency is termed as the maximum usable frequency* (MUF) for communicatio to the receiving point. Therefore if the frequency of the wave exceeds the maximum usable frequency for a given receiving point then the layer fails to return the wave to that point. That is, there would.be no reception at that point, for the distance to the receiving point falls short of the skip distance. If on the other hand, the frequency is less than MUF, then the distance to the receiving point exceeds the skip distance and propagation is effected by a relatively more slanting ray. The maximum usable frequency is of significance for two reasons :

(*i*) it is the maximum frequency that can be used in order to receive a sky wave at a given receiving point, and

(*ii*) it is the frequency at which the sky wave signal returned to the ground is, in general, strongest.

The value of MUF for ordinary ray is always less than that for extra ordinary, although this difference is small except for short distances and decreases with increase of frequency and distance.

From the forthcoming equation (4 we note that maximum usable frequency is related to the critical frequency and the angle of incidence by $MUF = f_c \sec i$.

In the case of *F*-layer, the the maximum angle of incidence is of the order of ·74 degrees which occurs for a ray that leaves the transmitting point at the grazing angle. Therefore, for F_2 layer

$$MUF = f_c \sec 74° = 3{\cdot}6\, f_o.$$

The highest frequency signals deliver greatest field strength at the receiver and as the frequency is reduced the field strength becomes less, particularly during the day time. The ionospheric sudden disturbances greatly reduce the maximum usable frequency and the absorption is very much increased, thereby, reducing the signal strength. Therefore frequency of communication is to be altered.

The maximum usable frequency shows day to day variations. Therefore they are selected on the basis of monthly averages. It is customary to choose a frequency atleast 15 percent below the monthly average for maximum usable frequency to take care of day to day and hour to hour deviations. This selected frequency is called *optimum frequency*. The optimum frequency for transmission between any two points is therefore selected as some frequency lying between about 50 and 85 per cent of the predicted maximum unstable frequency between those points.

Figure 16 shows the variation of maximum usable frequency as function of distance for various times of day. Each curve is marked with appropriate local time and with the layer determining the maximum usable frequency.*

Now we shall develop few mathematical relations for skip distance and maximum usable frequency.

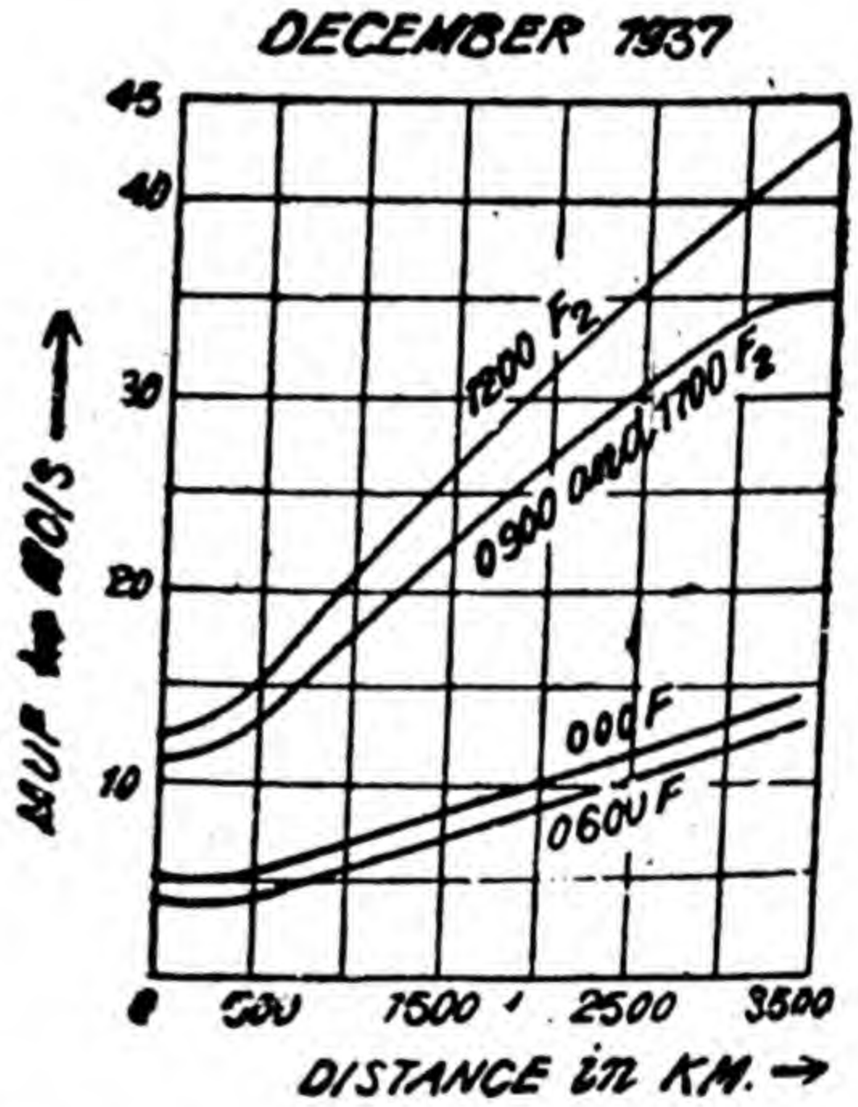


Fig. 16. Variation of *MUF* with distance.

Expression for skip distance and MUF :

In the theoretical determination of the maximum usable frequencies, it is convenient to consider separately two cases :

(*i*) distance within 500 km. for which the earth may be considered as flat, and

(*ii*) distance greater than 500 km. for which the earth's curvature has to be taken into account.

(i) **Flat Earth (thin layer)** : For the case of flat earth and thin layer, the maximum usable frequency and the skip distance can easily be calculated geometrically. This case is practically

*For Washington, D.C.

realised when the distance of transmission is less than 500 km. and reflection takes place from sporadic E region.

Neglecting collisions and the effect of the earth's magnetic field, the maximum frequency f which can be reflected from the layer at an angle of incidence i is given by $f^2=\frac{80{\cdot}6\ N'}{\cos^2 i}$. ...(2)

Therefore critical frequency for which $i=0$, is $f_c^2=80{\cdot}6\ N'$,...(3) because f_c is the frequency for which the refractive index of the medium becomes zero at the top of the trajectory.

From equations (1) and (3), we find that $f_o=f\cos i$. ...(4)

Figure 17 (*a*) shows transmission path for the case of reflection from a thin ionized layer (flat earth).

Now from figure 17 (*a*), we have

$$\tan i=\frac{D}{2h} \quad ..(5)$$

where D is the distance of transmission and h the height of the layer,

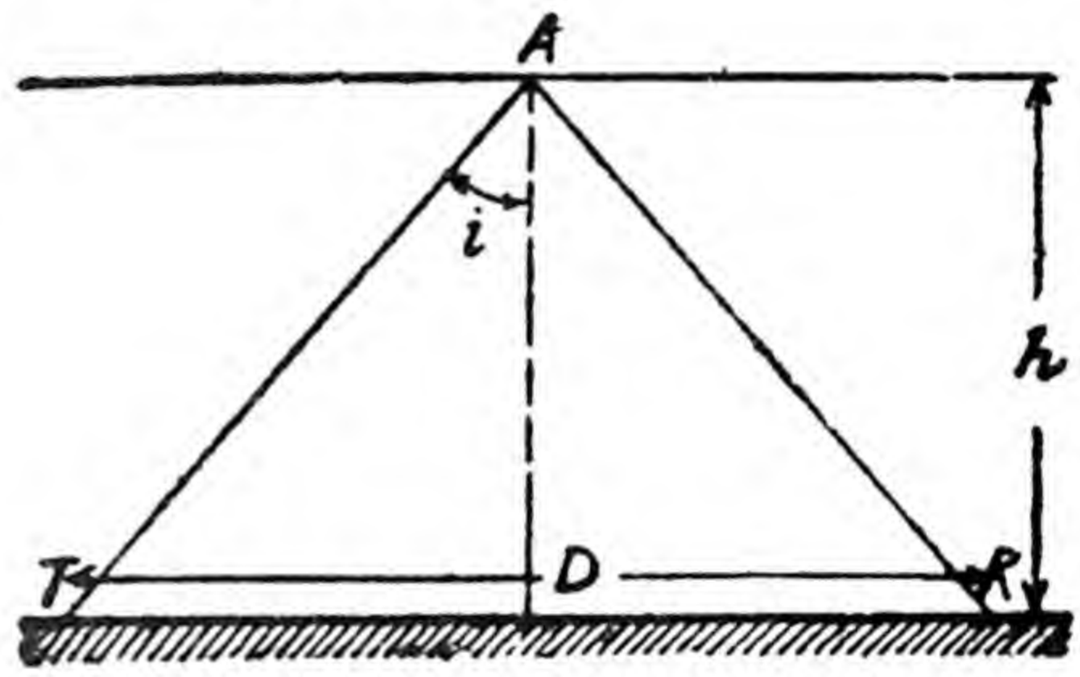


Fig 17 (a) Transmission path for the case of reflection from a thin ionized layer.

or $$1+\tan^2 i=1+\frac{D^2}{4h^2}$$

or $$\sec^2 i=1+\frac{D^2}{4h^2}. \quad ...(6)$$

From equations (4) and (6), we have

$$\frac{f^2}{f_c^2}=1+\frac{D^2}{4h^2}. \quad ...(7)$$

The equation (7) gives the maximum frequency which may be used for the minimum distance *i.e.* the skip distance, upto which a particular frequency can be used. Thus

$$f_{max.}=f_o\sqrt{\left(\frac{D^2+4h^2}{4h^2}\right)}$$

and $$D_{skip}=2h\sqrt{\left(\frac{f^2}{f_c^2}-1\right)}. \quad ...(8)$$

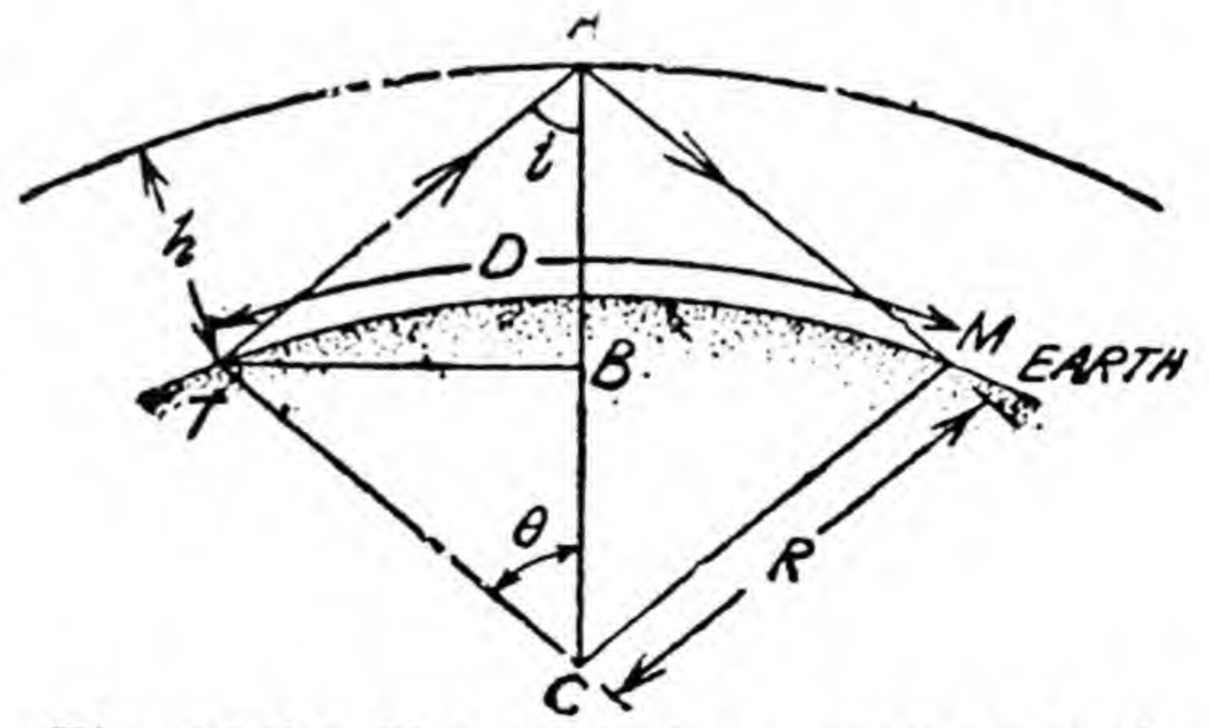


Fig. 17 (b) Transmission over a curved earth by reflection from the ionised layer.

(ii) Curved earth (Thin layer) : If the earth is considered flat, then it is obvious that as the angle of incidence of the wave is increased, the skip distance will also increase until it extends to infinity. If, however, the curvature of the earth is taken into account,

the skip distance cannot extend to infinity ; but attains a maximum value when the ray leaves the transmitter tangentially to the earth (fig. 17*b*).

Let 2θ be the angle subtended at the centre of the earth by transmitting distance D ; we have

$$D=2R\theta, \qquad \ldots(9)$$

$$\tan i=\frac{BT}{AT}=\frac{R\sin\theta}{h+R-R\cos\theta},$$

or $$\tan^2 i=\frac{\sin^2\theta\ R^2}{(h+R-R\cos\theta)^2}.$$

Therefore,

$$\cos^2 i=\frac{f_c^2}{f^2}=\frac{(h+R-R\cos\theta)^2}{R^2\sin^2\theta+(h+R-R\cos\theta)^2}. \qquad \ldots(10)$$

Under the condition of maximum D_{skip}, θ is maximum and is given by

$$\cos\theta=\frac{R}{R+h}=\left(1-\frac{h}{R}\right).$$

Since $\frac{h}{R}\ll 1$, $\cos\theta\cong 1$ or θ is always very small. We may, therefore, write that

$$\cos\theta=\left(1-\frac{\theta^2}{2}\right). \qquad \ldots(11)$$

On putting for $\sin\theta$ and $\cos\theta$ in (10), we get

$$\frac{f_c^2}{f^2}=\frac{\left\{h+R-R\left(1-\frac{\theta^2}{2}\right)\right\}^2}{R^2\theta^2\left\{h+R-R\left(1-\frac{\theta^2}{2}\right)\right\}^2}.$$

Putting $\theta^2=\frac{D^2}{4R^2}$, we get

$$\frac{f_c^2}{f^2}=\frac{\left[h+R-R+\frac{D^2}{8R}\right]^2}{\frac{D^2}{4}+\left[h+R-R+\frac{D^2}{8R}\right]^2}$$

$$=\frac{\left(h+\frac{D^2}{8R}\right)^2}{\frac{D^2}{4}+\left(h+\frac{D^2}{8R}\right)^2}$$

or $$f_{max.}=f_c\left[\frac{D^2+4\left(h+\frac{D^2}{8R}\right)^2}{4\left(h+\frac{D^2}{8R}\right)^2}\right]^{1/2}. \qquad \ldots(12)$$

Variation of maximum usable frequency facto , f_{max}/f_c, with transmission range D for curved earth and thin ionised layer has been plotted in fig. 17 (*c*). for two altitudes *i.e.* for E layer (110km.) and F_2 layer (300 km). It is reasonable to expect that *the optimum short wave frequency will be high at low altitudes.*

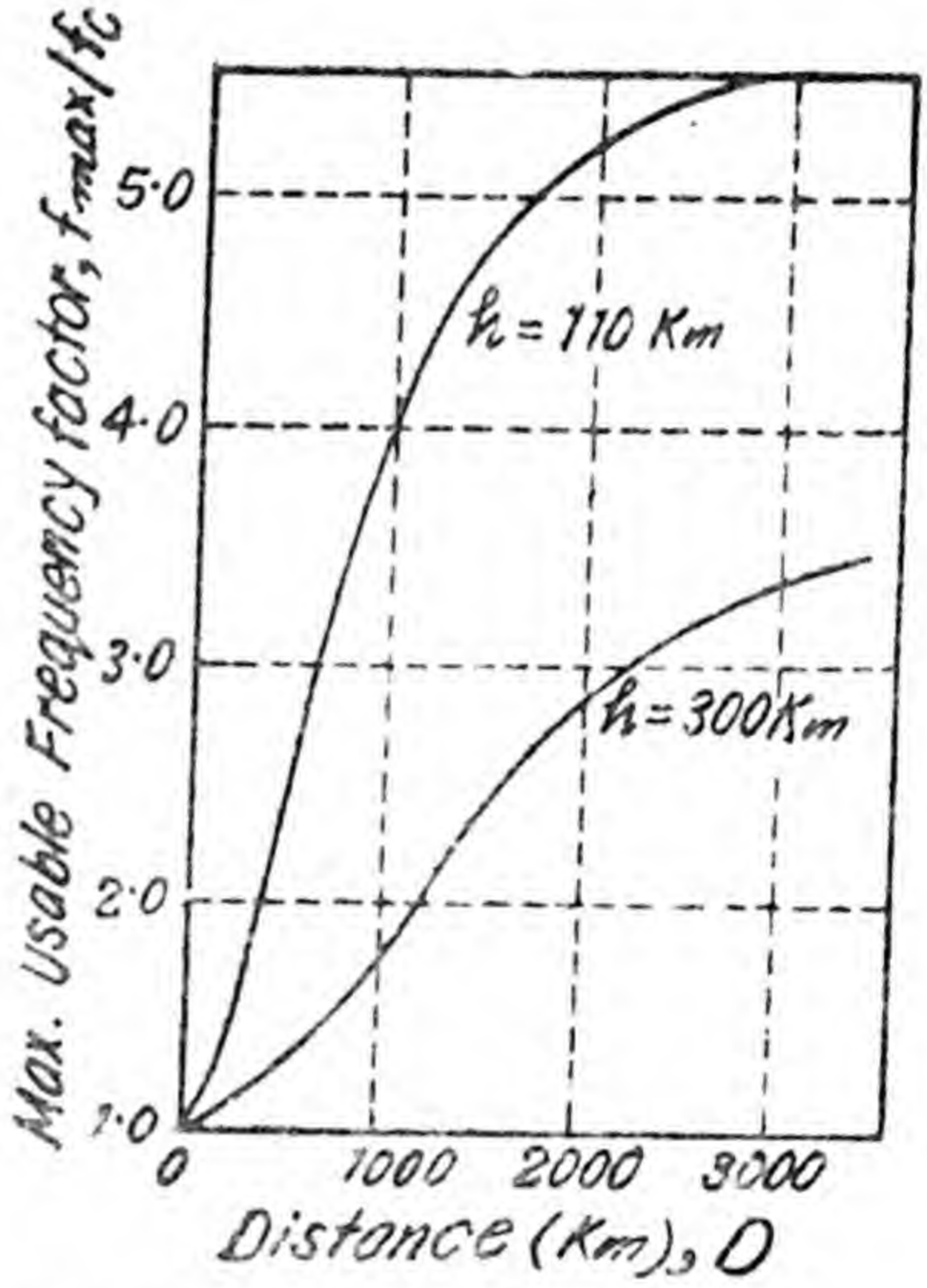


Fig. 17 (*c*) Variation of MUF with distance.

29·6. EQUIVALENCE OF VIRTUAL HEIGHTS OF REFLECTION FOR OBLIQUE AND VERTICAL INCIDENCE : (Group Path and Equivalent Path) :

In this article, we shall prove the following two theorems :

(*i*) *The time of travel of a radio wave around the real curved path in the ionosphere is equal to the time of travel around the virtual triangular path at the velocity of light in free space.*

(*ii*) *If a vertical ray of frequency, f_c, and an oblique ray of frequency, f, are reflected from one and the same region in the ionosphere that is, if the points of reflection are ot the same real height, then virtual height is also the same for both the rays.*

Proof of First Theorem : Let a sky wave, incident obliquely at an angle, i, on the ionised layer, travelling from T to R, follow a path such as TBR (fig. 18*a*). The apex B of the path corresponds to an ionisation density sufficient to reflect back the incident wave. Path TM and NR are traversed in air ($\mu = 1$) with velocity of light in free space, while path MBN lying in ionised region ($\mu < 1$), is traversed by the group velocity*, v_g, of the waves in the ionised region that is *less* than the velocity of light. Path TBR is called *group path* Refer to fig. 18 (*a*). The two waves, one incident obliquely, other vertically are shown. The real height for the two waves is $(h_0 + y_0)$, while virtual height for oblique ray is $\left(h_0 + y_v' \right)$ and for vertical ray is $(h_0 + y_v)$. For the proof of first theorem we are to show that

---

*$v_g = c\left(1 - \dfrac{80 \cdot 6N}{f^2}\right)^{1/2}$

$$t_{TMBR}=t_{TMAR}$$

or $t_{MBN}=t_{MAN}$, (path MAN being considered in free space)

i.e. time along the triangular path & the curved path, taken by the wave to travel is the same. *Actual pat h is TBR* and *equivalent path (virtual path) is TAR*, the latter is considered in free space.

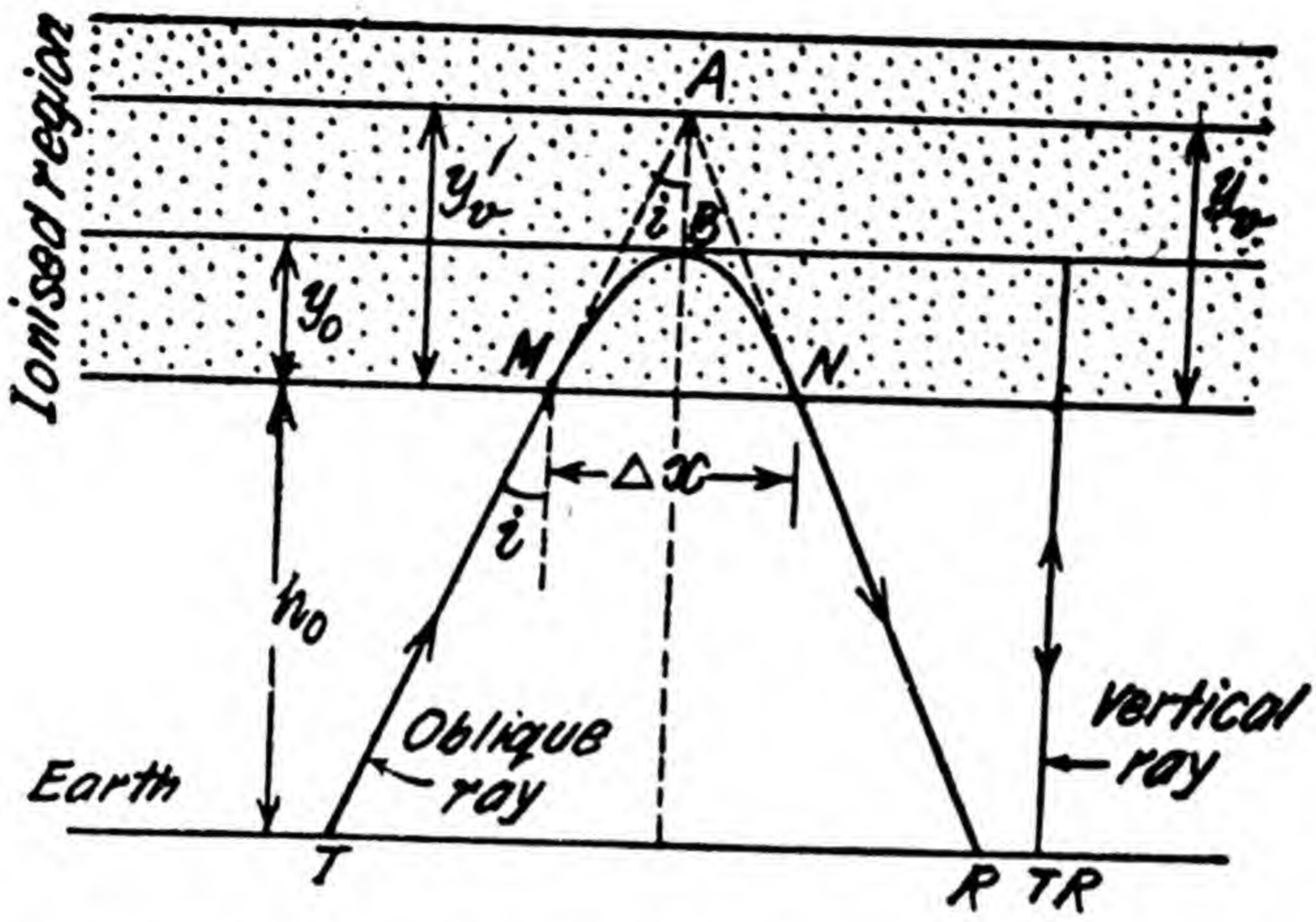


Fig. 18 (*a*). Actual height from where both rays are reflected $=h_0+y_0$

virtual height for oblique ray $=h_0+y_v'$

virtual height for vertical ray $=h_0+y_v$

Consider now that the whole triangular path TAR, obtained by extending straight line portions of the actual path TBR, lies in free space. Then time taken in traversing path MAN will be

$$t_{MAN}=\frac{\text{path}MAN}{c} \quad \text{.. (1)}$$

Let us consider path MBN. Velocity at any point in the ionised region is given by the group velocity,

$$v_g=c\left(1-\frac{80\,6N}{f^2}\right)^{1/2},$$

where N is ionisation density at that point and f is the frequency of wave incident obliquely. Since N varies, v_g will also vary along the path MBN so that time taken by the wave to travel along this path will be

$$t_{MBN}=\int_{MBN}\frac{ds}{v_g}$$

Since $v_g v_p=c^2$ and $\mu=c/v_p$, we have $v_g=\mu c$, so that

$$t_{MBN}=\int_{MBN}\frac{ds}{\mu c}=\frac{1}{c}\int_{MBN}\frac{ds}{\mu} \quad \text{.. (2}a\text{)}$$

We, in fig. (18*b*), are again showing separately the path MBN in ionised region. While the wave is travelling, at any instant, t

let ds' be any element of path, r be the angle made by the direction of propagation of the group at any point in its path with the vertical, dx' be the horizontal component of ds' and μ' be the refractive index at that point then

$$dx' = ds' \sin r$$

so that

$$t_{MBN} = \frac{1}{c}\int_{MBN} \frac{ds'}{\mu'} = \frac{1}{c}\int_0^{\Delta x} \frac{dx'}{\mu' \sin r} \qquad \ldots(2b)$$

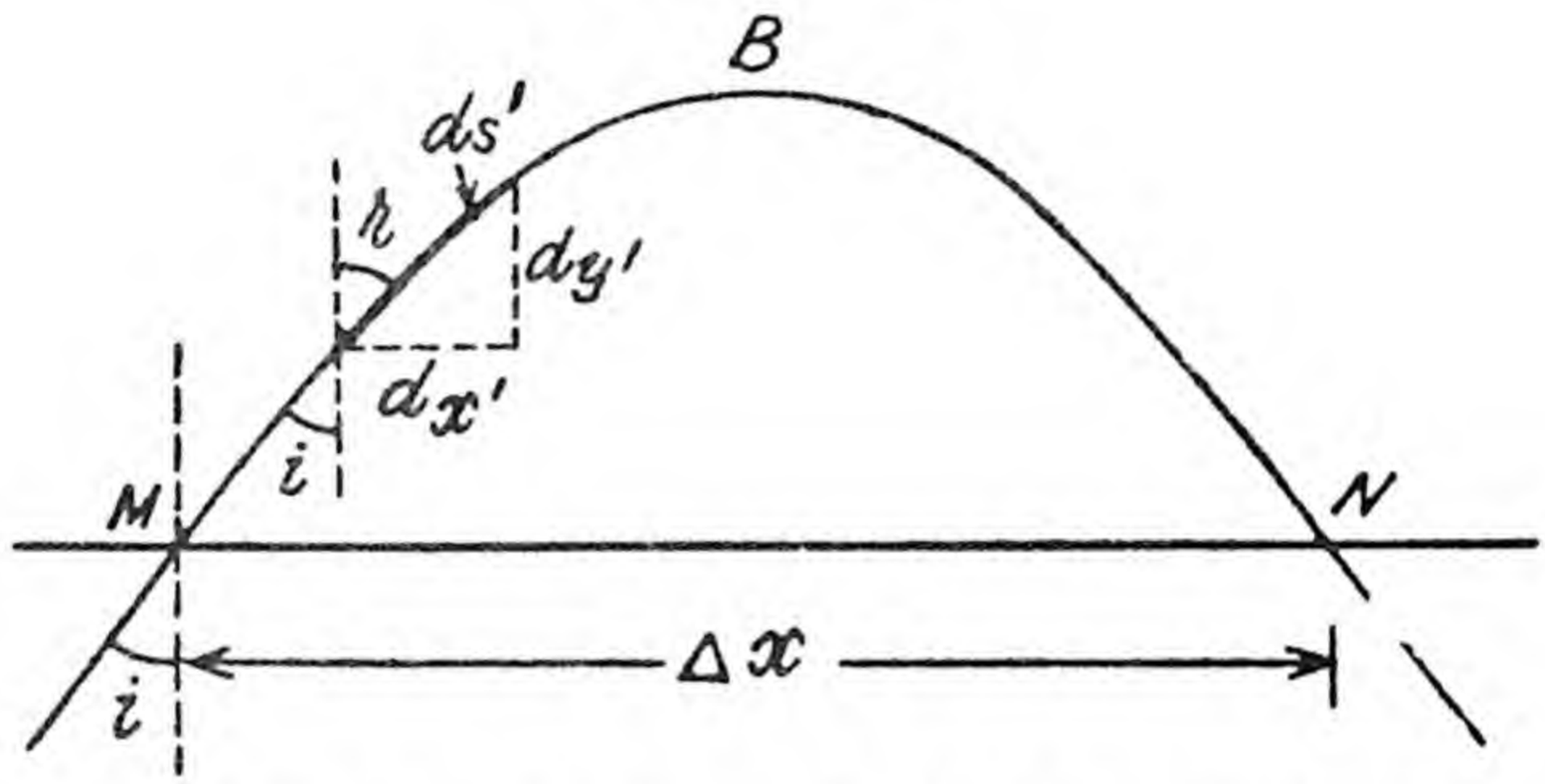


Fig. 18(*b*) Showing the path *MBN* in ionised region.

But according to Snell's law,

$$\mu' = \frac{\sin i}{\sin r},$$

so that

$$t_{MBN} = \frac{1}{c}\int_0^{\Delta x} \frac{dx'}{\sin i} = \frac{\Delta x}{c \sin i} \qquad \ldots(3)$$

But from $\triangle MAN$ (fig. 18*a*), we have

$$\frac{\Delta x}{2} = MA \sin i$$

or $$\Delta x = 2MA \sin i = MAN \sin i$$

Putting it in eq. (3), we get

$$t_{MBN} = \frac{MAN}{c} = t_{MAN} \qquad \text{(see eq. 1)}$$

This proves the theorem that time taken along the path *MBN* in ionised region is the same as the time taken by the wave in travelling along the path *TAR* considered fully in free space. *Thus time of travel along actual path and equivalent path is the same.*

Proof of Second Theorem : In fig. 18 (*a*) we have shown two waves—one incident at an angle, *i*, at the ionised layer and the other incident vertically. Both are reflected from the same real heights. We are to show that

$$h_0+y_v' = h_0+y_v$$

or
$$y_v' = y_v$$

As the wave is reflected at point *B*, we have for *oblique* wave

$$\mu_B = \frac{\sin i}{\sin 90} = \sin i$$

Also

$$\mu_B = \left(1-\frac{80{\cdot}6N_B}{f^2}\right)^{1/2}$$

so that

$$\mu_B^2 = \sin^2 i = \left(1-\frac{80{\cdot}6N_B}{f^2}\right)$$

or
$$80{\cdot}6N_B = f^2\cos^2 i \qquad \ldots(4)$$

For *vertical* wave, $i=0$, so that

$$80{\cdot}6\,N_B = f^2 = f_o^2 \qquad \ldots(5)$$

where f_o is termed as critical frequency (or *equivalent normal incidence frequency corresponding to oblique incidence frequency, f*).

Now refer to fig. (18*b*). For any point, say above *M*, in ionised region, we can write equally well that

$$\mu'^2 = 1-\frac{80{\cdot}6N'}{f^2}$$

$$= 1-\frac{80{\cdot}6N'}{80{\cdot}6N_B}\cos^2 i \qquad \text{(using eq. 4)}$$

$$= 1-\frac{80{\cdot}6N'}{f_c^2}\cos^2 i \qquad \ldots(6)$$

Further

$$\mu' = \frac{\sin i}{\sin r}$$

or
$$\mu'^2\sin^2 r = \sin^2 i$$

or
$$\mu'^2(1-\cos^2 r) = \sin^2 i$$

or
$$\mu'^2 = \mu'^2\cos^2 r + \sin^2 i \qquad \ldots(7)$$

Equating eqs. (6) and (7), we get

$$\mu'^2\cos^2 r + \sin^2 i = 1-\frac{80{\cdot}6N'}{f_c^2}\cos^2 i$$

or
$$\mu'^2\cos^2 r = \cos^2 i\left(1-\frac{80{\cdot}6N'}{f_c^2}\right)$$

$$= \mu_c'^2\cos^2 i \qquad \ldots(8)$$

where μ_c' is the refractive index of the *same* region for vertical wave frequency, f_o, the critical frequency.

From fig. 18 (*a*), we have

$$y_v' = MA \cos i = \cos i . \frac{MAN}{2}$$

$$= \frac{\cos i}{2} . c . \frac{MAN}{c}$$

$$= \frac{\cos i}{2} . c . t_{MBN}$$

$$= \frac{\cos i}{2} . c . \frac{1}{c} \int_{MBN} \frac{ds'}{\mu'} \quad \text{(using eq. 2}b\text{)}$$

$$= \frac{\cos i}{2} \int_{MBN} \frac{ds'}{\mu'}$$

If dy' is the component of path element ds' in the vertical direction (fig. 18*b*), then

$$dy' = ds' \cos r$$

so that

$$y_v' = \frac{\cos i}{2} . 2 \int_0^{y_0} \frac{dy'}{\mu' \cos r} \quad \ldots(10)$$

$$= \cos i \int_0^{y_0} \frac{dy'}{\mu_c' . \frac{\cos i}{\cos r} . \cos r} \quad \text{(using eq. 8)}$$

$$= \int_0^{y_0} \frac{dy'}{\mu_c'} \quad \ldots(11)$$

But μ_c' corresponds to a wave of frequency, f_c, which is incident vertically ($i=0$) at the ionised layer and therefore

$$\int_0^{y_0} \frac{dy'}{\mu_c'} = y_v \quad \ldots(12)$$

Therefore from eq. (11) and (12), we get

$$y_v' = y_v$$

or

$$h_0 + y_v' = h_0 + y_v$$

or virtual height of oblique wave = virtual height of vertical wave

Thus *the equivalent height of reflection of a wave of frequency, f_c, incident vertically on the ionosphere is the same as that of a wave of frequency, f, incident obliquely, provided that the two waves are reflected from the same region of ionisation.*

29·7. METHODS FOR MEASURING THE HEIGHT AND ELECTRON CONCENTRATION OF IONOSPHERIC REGION :

(A) Pulse Method : One of the radio methods developed for

exploring the ionospheric region and which is now universally employed is the pulse method, first introduced by Breit and Tuve. The ionospheric quantities that can be measured by radio method of exploration, in particular by pulse method, are following :

(*a*) Time taken by the radio signal to travel up and down from the ionosphere.
(*b*) Intensity of the reflected signal.
(*c*) The state of polarisation of the received signal.
(*d*) Maximum electron number density of an ionized layer.

(a) Determination of equivalent height :

Principle : In this method, time taken by a series of radio frequency 'pulses' of short duration to travel up to the ionised region and return to the earth is measured, Knowing this time, the equivalent height of the reflecting region is easily calculated. This method is also called 'group retardation method' because the signal in its passage through the ionospheric region travels with the velocity appropriate for a wave group.

Method : An actual sky wave travelling from *T* to *R* will follow a path as *TBR* [fig. 19 (*a*)]. The apex *B* of the path corresponds to a point in the ionosphere where the electron density is such that the wave is reflected.

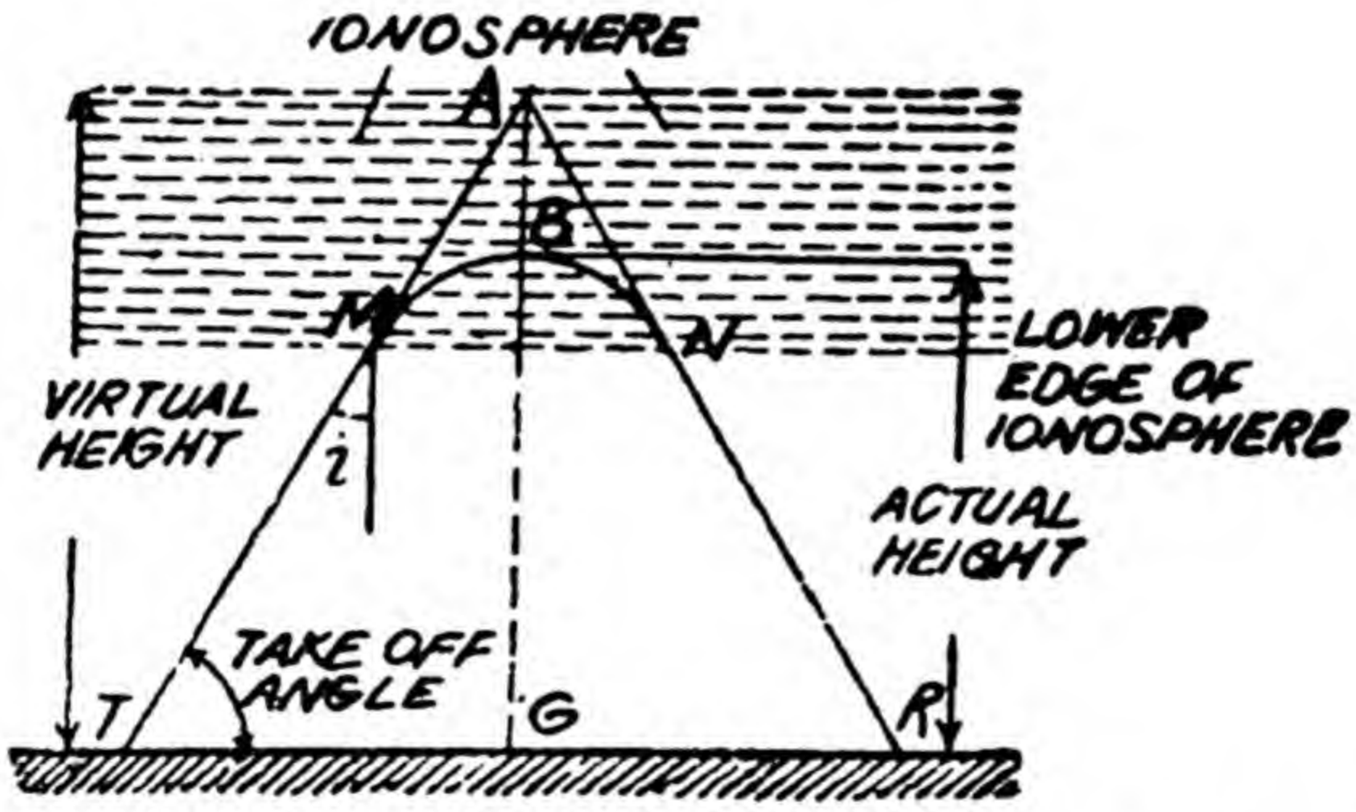


Fig. 19 (*a*) Illustrating the principle of measuring the equivalent height.

The paths *TM* and *NR* are traversed in air ($\mu=1$) with the velocity of light in free space. The length *MBN*, being in the ionized region ($\mu < 1$), is traversed with a group velocity* less

---

*The phase velocity is determined by the rapidity with which the phase changes along the path of a wave and is, therefore, different from signal or group velocity. We have seen that for ionised region, the group velocity is

$$v_g = c\sqrt{\left(1-\frac{f_N^2}{f^2}\right)} = c\sqrt{\left(1-\frac{80\cdot 6N}{f^2}\right)},$$

predicting that

$$v_g < c,$$

for ionised media.

than velocity of light. The time taken by the signal to traverse the path *TMNR* is, therefore, given by

$$t=\int_{TMNR}\frac{ds}{v},$$

where *ds* is an element of path and *v* is the velocity of wave group at any point in its path.

Consider now the triangular path *TAR* obtained by extending the straight line portions of the actual path *TBR*. The time required for a signal to travel with the group velocity over the actual path *TBR* is then the same as that required for a signal travelling with the velocity of light to travel over the equivalent path *TAR**. Thus the ionosphere could be replaced by a perfect reflector located at the height of point *A* (virtual height) without affecting the conditions of propagation. The path *TAR* is therefore, called the *equivalent path P'*. Thus

$$\text{Length of } P'=c\int_{TMNR}\frac{ds}{v}$$

In this way we get height *AG* of the vertex of the equivalent triangle which is called the equivalent or virtual height. The equivalent height is thus always greater than the true height *BG* of the point of reflection.

Technical Details : Transmitting and receiving systems for the pulses sent upwards for the exploration is shown in fig. 19 (*b*).

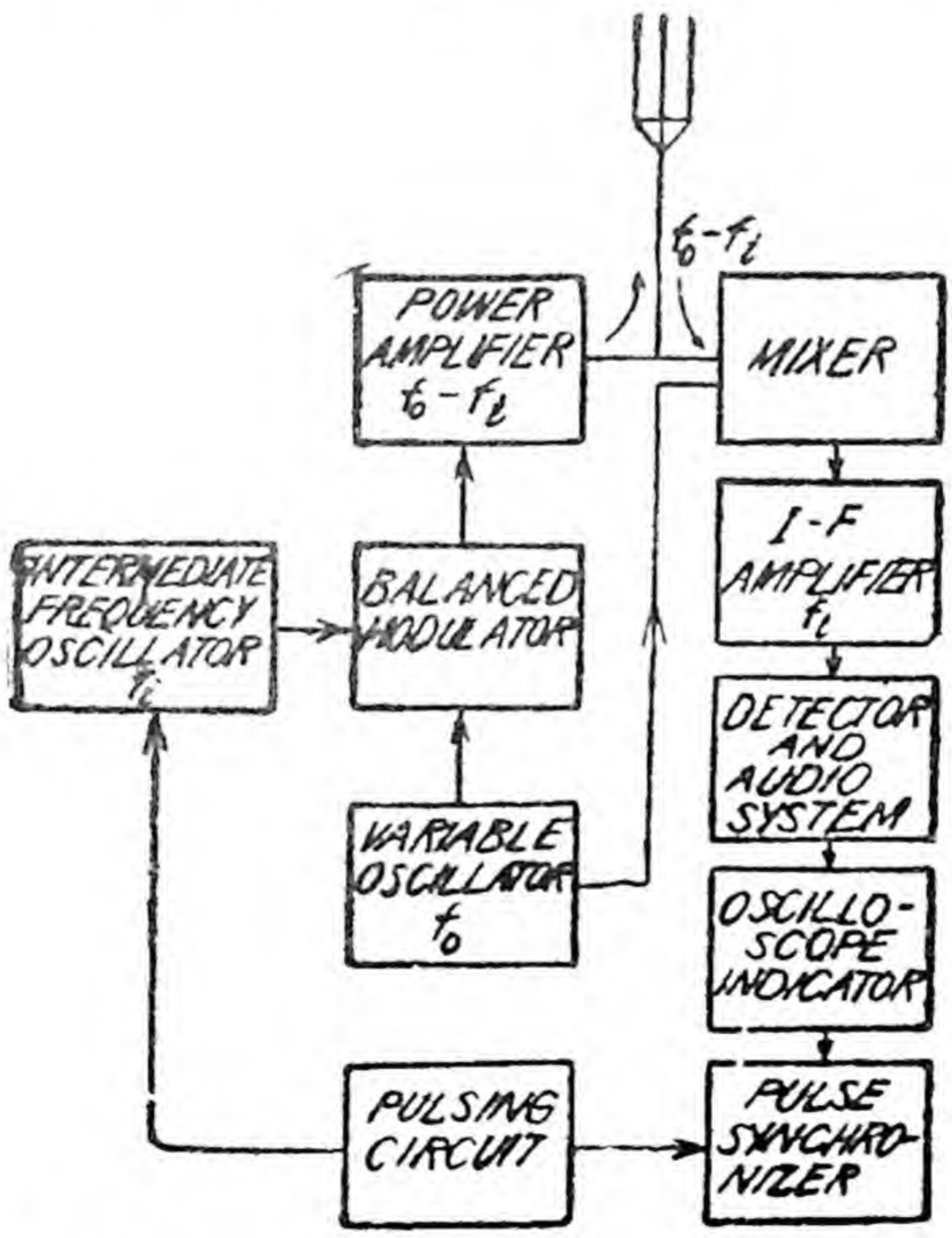
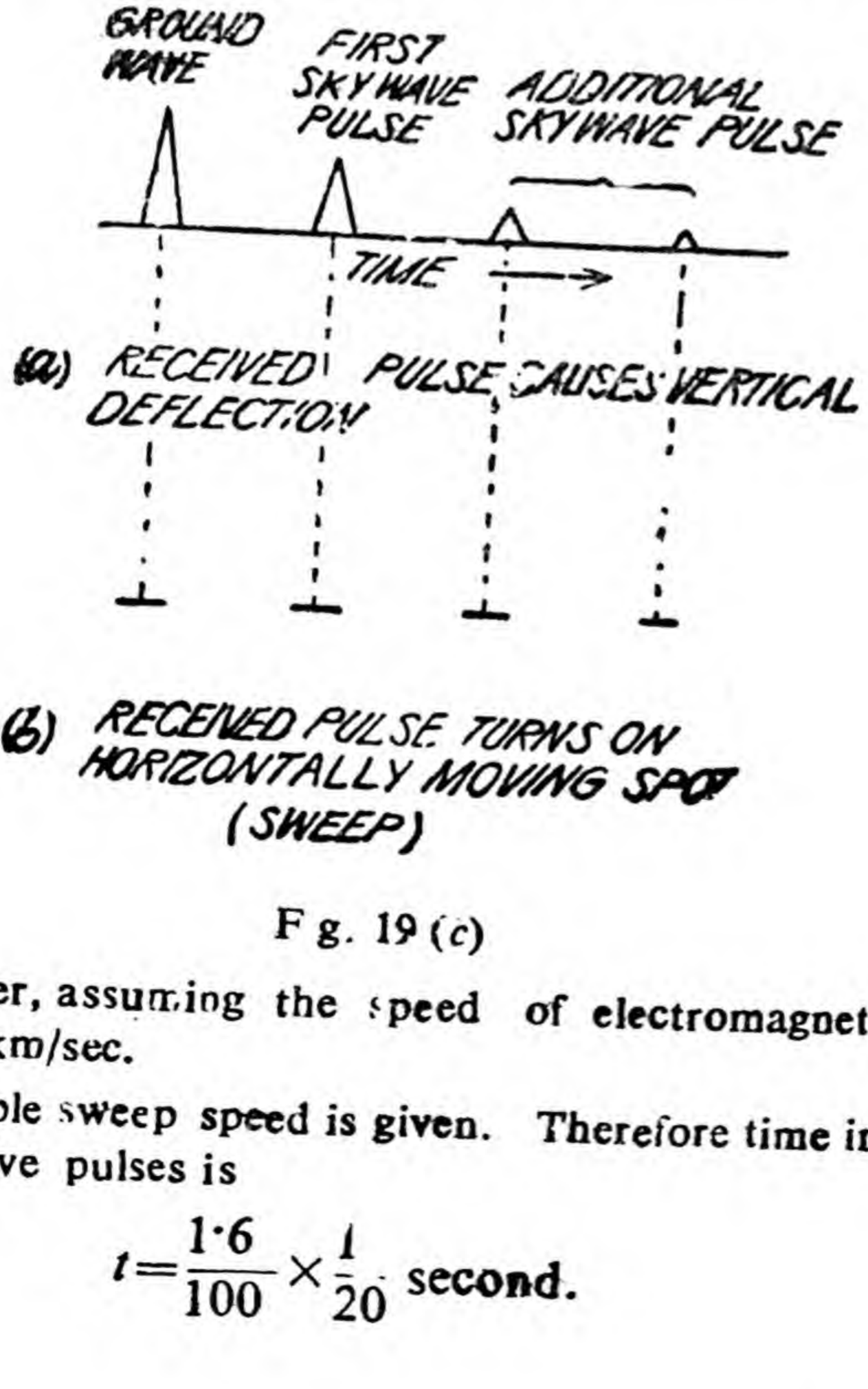


Fig. 19 (b). Assembly for pulse method.

*The group veloeity in the ionosphere is just enough less than the velocity of light to make up for shorter distance.

In this method, short pulses, about 50 to 100 per second are transmitted. The frequency of pulses is changed gradually so that a frequency range of perhaps 0·5 – 2 mc/s is covered in a period of about 15 minutes. Pulses have peak power of several hundred watts. Receiver is located at a distance of few feet from the transmitter and picks up both the direct and reflected signals. The output of the receiver is fed to *y* plates or vertical deflection plates of the oscilloscope. On *x* plate or horizontal deflection plates, no external voltage is applied but internal sweep synchronised with the pulses sent upwards, is used. First the transmitted pulse (ground pulse) and then the reflected pulse (first skywave pulse) produce vertical deflections as shown in fig. 19 (*c*). The time interval between these two pulses is a measure of the path length and suppose it is *t* sec., then the total distance travelled by the pulse will be *ct*, where *c* is the velocity of light. The half of this distance travelled (if the waves are sent velocity upwards) *i.e.* $\frac{1}{2}ct$ will give the virtual height. We calculate virtual height because it has been assumed as if ionised region is not present and the pulse has traversed the entire path with velocity *c* in free space. We illustrate the method by the following example :

Suppose the distance between the ground and sky wave pulse is 6 cms. If the sweep speed is 20 meters per second, then to find the effective height of



F g. 19 (*c*)

the ionospheric layer, assuming the speed of electromagnetic waves in free space to be 3×10^5 km/sec.

In this example sweep speed is given. Therefore time interval between ground and sky wave pulses is

$$t = \frac{1 \cdot 6}{100} \times \frac{1}{20} \text{ second.}$$

The total distance travelled by the sky wave, will be

$$= \frac{1{\cdot}6}{100} \times \frac{1}{20} \times 3 \times 10^8 \text{ meters.}$$

$$= 4 \times 10^4 \text{ meters.}$$

so that the effective heigh will be

$$= 12 \times 10^4 \text{ meters.}$$

Additional pulses, though absent in many cases, represent either waves returned from higher layers or waves that have made more than one round trip between earth and first layer. A second echo pulse would mean the sky wave pulse that has taken one trip more between earth and first layer.

(b) Maximum Electron Density :

(*i*) *When the magnetic field is not accounted* : To determine the maximum electron number density $N_{max.}$ of an ionised layer, exploring waves of gradually increasing frequency are sent vertically upwards to meet the layer. Reflection will occur upto a certain frequency and after that if the frequency is still increased, there will be no reflection. The limiting frequency at which the reflections—the so called 'echos'—just begin to disappear, is called the critical penetrating frequency of the layer.

Condition of reflection from the ionospheric layer is

$$\mu^2 = \sin^2 i = \sqrt{\left(1 - \frac{4\pi Ne^2}{m\omega^2}\right)}.$$

In case of normal incidence, $i=0$, *i.e.* $\mu = 0$

$$N = N_{max.} \quad \text{and} \quad \omega = \omega_c = 2\pi f_c.$$

We get $$1 - \frac{4\pi N_{max}\ e^2}{m\omega_c^2} = 0,$$

$$N_{max.} = \frac{\pi m}{e^2} f_c^2 = 1{\cdot}24 \times 10^4 f_c^2,$$

where f_c is expressed in megacycles per second The determination of f_c, therefore, leads to the calculation of the maximum electron number density of the ionospheric layer penetrated.

(*ii*) *When the magne ic field is accounted* : When magnetic field is present, the wave is split up into two components and the equivalent heights of reflection and the penetration frequencies of the split components—the ordinary and extra ordinary—have to be distinguished from each other.

As the frequency is increased further above the critical frequency for the lower layer, a second and sometimes third critical frequency may be reached corresponding to the critical frequencies for the higher layers. Virtual height and frequency curves are shown in fig. 19 (*d*). In figure, *O* indicates ordinary ray and *x* indicates extra-ordinary ray. At higher frequencies, the earth's magnetic field causes the virtual height trace to split into two parts corresponding to ordinary and extra ordinary rays which have different phase and group velocities.

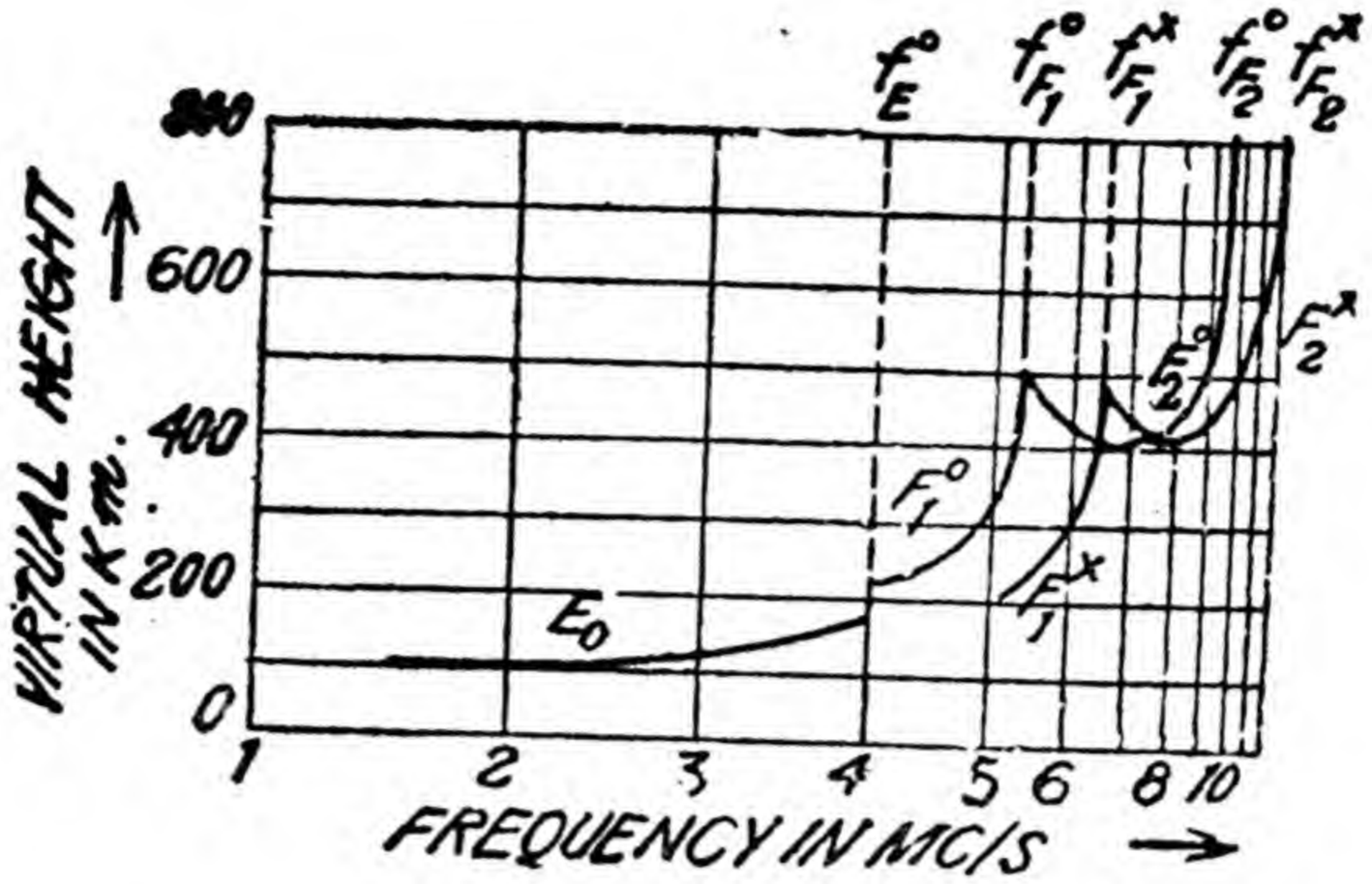


Fig. 19. (d) f_E^0 represents critical frequency of E layer for ordinary ray

$f_{F_1}^x$ represents critical frequency of F_1 layer for extra ordinary ray etc.

Assuming that the echoes corresponding to the ordinary and extra-ordinary components have been identified, the maximum electron number density can be calculated by noting critical penetration frequencies of these components. Thus for ordinary wave, if the critical penetration frequency be f_c, then

$$N_{max.} = \frac{\pi m}{e^2} f_c^2. \qquad \ldots(i)$$

For the extra-ordinary wave following cases are to be distinguished :

(a) If the exploring frequency is greater than the gyrofrequency ($\omega > \omega_H$), then the extra-ordinary ray may be reflected from two distinct regions of electron density given by

$$\omega_c^2 = \omega^2 - \omega\,\omega_H. \qquad \ldots(ii)$$

$$\omega_c^2 = \omega^2 + \omega.\omega_H. \qquad \ldots(iii)$$

These can be derived from conditions of wave reflection from art. 29·4 that are

$$1 + A = 0 \qquad \text{for ordinary ray}$$

$$1 + A = \pm\gamma \qquad \text{for extra ordinary ray}$$

Putting $A = \dfrac{-m\omega^2}{4\pi Ne^2}$ and $\gamma = \sqrt{(C_L^2 + C_T^2)}$

with $$C_L = \frac{\frac{e}{c}.H_L.\omega}{4\pi Ne^2} \quad \text{and} \quad C_T = \frac{\frac{e}{c}.H_T.\omega}{4\pi Ne^2},$$

we get $$1 - \frac{m\omega^2}{4\pi Ne^2} = \pm \frac{\frac{e}{c}\omega}{4\pi Ne^2}(H_L^2 + H_T^2)^{1/2}$$

$$= \pm \frac{\frac{e}{c}.\omega.H}{4\pi Ne^2}.$$

Putting

$$\omega_c^2 = \frac{4\pi Ne^2}{m}, \text{ called critical frequency}$$

and $$\omega_H = \frac{eH}{mc}, \text{ called gyromagnetic frequency}$$

we arrive at

$$1 - \frac{\omega^2}{\omega_c^2} = \pm \frac{\omega.\omega_H}{\omega_c^2}$$

or $$\omega_c^2 = \omega^2 \pm \omega.\omega_H$$

or $$f_c^2 = f^2 \pm f.f_H.$$

Since electron density increases with height, the region defined by (*ii*) is at a lower level than defined by (*iii*). Generally, most of the energy will be reflected at the lower level. If we denote the critical frequency for this layer $f_x = \frac{\omega_x}{2\pi}$, then from second condition,

$$\frac{4\pi N_{max}.e^2}{m} = 4\pi^2 f_x^2 - 4\pi^2 f_x.f_H$$

or $$N_{max} = \frac{\pi m}{e^2}(f_x^2 - f_x f_H). \quad \ldots (iv)$$

(b) If the exploring frequency is less than the gyrofrequency, then the extra-ordinary component is reflected from the level given by

$$\omega_c^2 = \omega^2 + \omega\,\omega_H$$

from which we get $$N_{max.} = \frac{\pi m}{e^2}(f_x^2 + f_x.f_H), \quad \ldots (v)$$

where $f_H = \frac{He}{2\pi mc}$, called gyromagnetic frequency.

(B) Automatic Multifrequency Ionospheric Recorder : Owing to the rapid changes in ionospheric characteristics, it becomes necessary that frequency range of pulses to be sent upward be covered in as short a time as possible. The main characteristics of automatic apparatus, as distinguished from manually operated ones, are that

(i) the transmitter and receiver tuning systems are always kept tuned to each other,

(ii) the oscillator frequency is slowly varied automatically over the entire range, and

(iii) the operation is repeated automatically at predetermined intervals.

Main construction details should, therefore, include the following :

(*a*) *Transmitter* : It consists transmitting valves in push pull arrangement for high power output. The entire frequency range to be covered say 1—13 mc/s, is divided into bands and for each band, separate oscillator tank circuits are employed. Transmission line feeds the power to antenna having broad band and vertical directivity.

(*b*) *Receiver* : It is of superheterodyne type having 2 or 3 intermediate frequency stages. The tuning range of the receiver is divided into a number of bands corresponding to the number of bands of the transmitter. Each tuning unit has its own r.f. and mixer stages and is interlocked with the corresponding band of the transmitter. The selectivity of the receiver is adjusted to accept the bandwidth of the pulses without serious loss in sensitivity. It is also designed in way as to be able to recover quickly from the paralysing effect of the direct transmitter pulse.

(*c*) *Recording System* : It is a C.R.O. with a film camera. The output of receiver, which always remains in tune with the transmitter, is applied to the vertical deflecting plates of the oscilloscope. The time base of the oscilloscope is generated by a circuit driven by the main multivibrator pulses. The echo pattern produced on the screen is photographed continuously on a moving film by the camera. From the calibration markings on each record, the heights and the penetration frequencies can be read directly with accuracy specified.

(*d*) *The variable frequency tuner* : It is responsible for controlling the tunings of the transmitter and receiver, and for the repetition of the frequency sweep at predetermined intervals. The variable condensers, of both the transmitter and receiver circuits, are driven mechanically so that transmitter and receiver remain in tune.

(C) **Panoramic Recorder** : This recorder depicts a delineated virtual height *v/s* frequency curve on the oscillograph screen Thus a standing picture of the reflections from the various ionospheric layers is readily available and we can study the rapid ionisation changes in different layers occurring during radio fade-outs and magnetic storms. It also displays the echoes of sporadic *E* layer (whose height and location rapidly change).

For recording arrangement, a large cathode ray tube with *long after glow* is used. The ground pulse (pulse which is sent upwards by the transmitter is also fed to the receiver) and the echoes appear as bright spots and the time base sweep moves horizontally as the frequency is increased. *Thus if frequency sweep is sufficiently quick and the afterglow is sufficiently long*, the virtual height *v/s* frequency curve will be directly traced and observed on the screen as a bright line. The special refinements of this type of recorder are

(i) bright spot records
(ii) fast frequency sweep (5—10 secs.)
(iii) arrangement for slow movement of the time base as the frequency is gradually increased.

Its principle of construction is depicted in fig. 19 (*e*). Details can be listed as follows :

(i) A correction wheel tunes together the transmitter and the the receiver,
(ii) C.R O. screen has an afterglow of about 10 seconds,
(iii) the size of panoramic picture is 18 cm × 9 cm
(iv) record for each picture takes 3—5 seconds,
(v) frequency range is from 1 4 - 14 mc/s,
(vi) the vertical deflection line moves from left to right in synchronisation with the rotation of the high frequency tuning system, producing the afterglow pattern

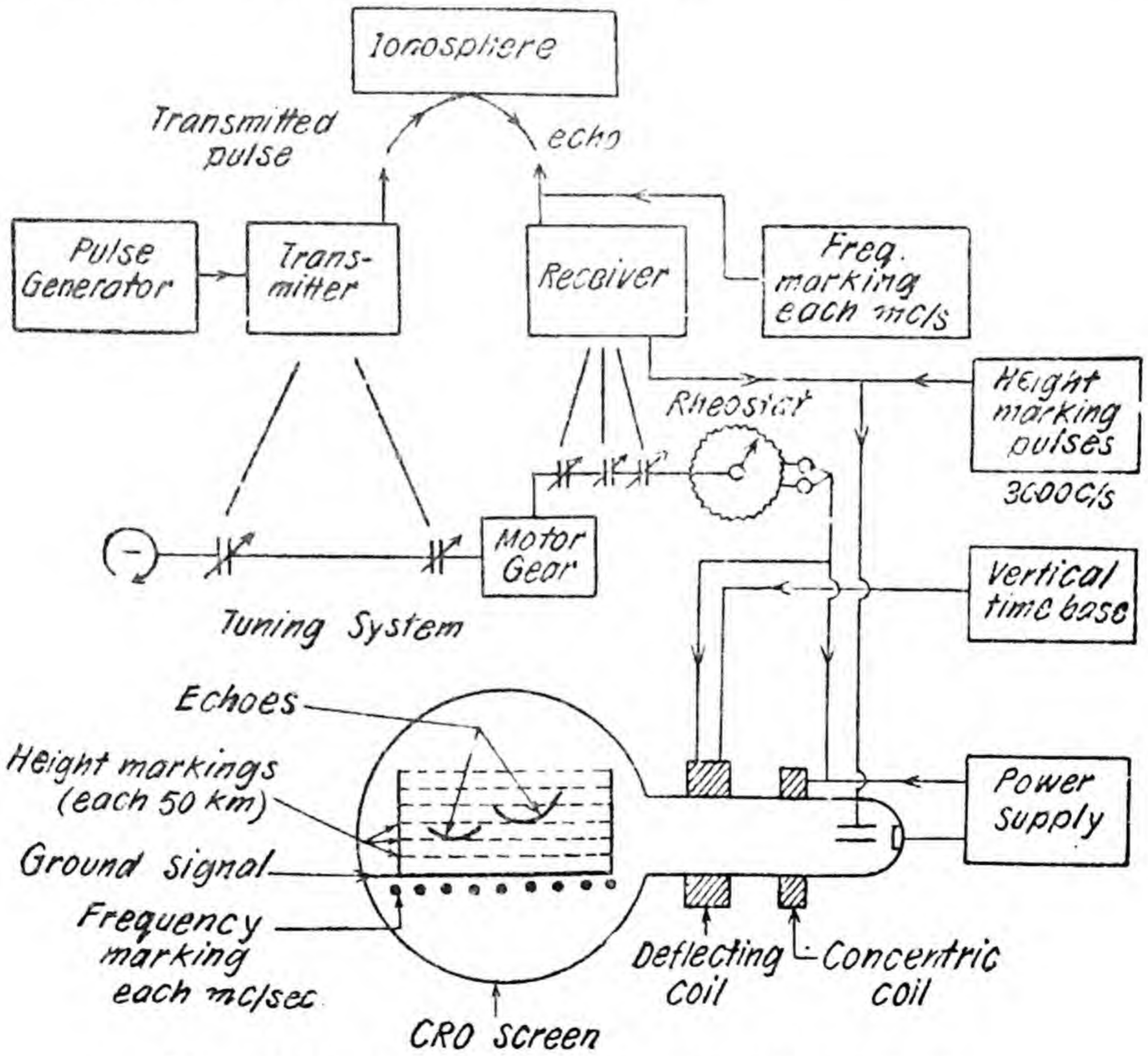


Fig. 19 (*e*). Panormic ionosphere recorder scheme.

29·8. STRATIFICATION OF IONOSPHERE :

The upper part of the earth's atmosphere absorbs large quantities of radiant energy from the sun, which not only heat the atmosphere but also produce some ionisation in the form of free electrons, and positive and negative ions The part of the

upper atmosphere, where the ionisation is appreciable, is called ionosphere.

The ionisation is apparently caused by ultraviolet radiation from the sun and it varies with the time of the day, the time of the year, and the longitude. At the earth's surface the ion density is very small, increasing to appreciable values at altitudes between 100 and 150 km. While the ions are distributed continuously in the atmosphere the concentrations vary and several maxima are reached at various latitudes. The regions near these maxima are called *layers*. For each such layer there is a critical frequency above which an electromagnetic wave, directed vertically, will not be returned to earth. These layers are characterised in terms of their critical frequencies and virtual heights. The critical frequency is a direct measure of the maximum electron density of the layer and virtual height is an indication of the height at which maximum electron density occurs.

The three ionospheric regions of great importance as far as radio communication is concerned, are denoted by the *E*, F_1, F_2 layers. The range of altitude over which these layers may vary is shown in the fig 20

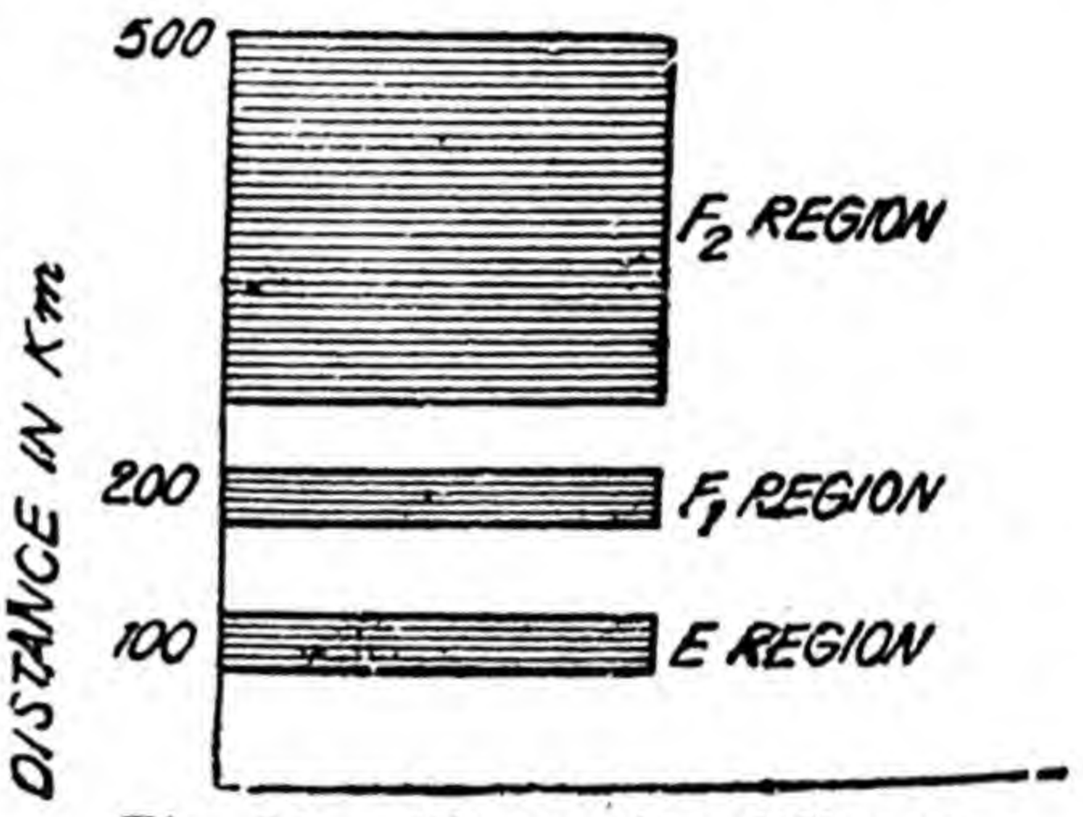


Fig 20. Illustrating different ionospheric layers with altitude.

In addition to three regular layers, there is a region below the *E* layer, which is responsible for much of the day time attenuation of high frequency radio waves, called the *D* region. It lies between heights about 53 and 90 km. Its detailed structure is not known with certainty.

An anomalous ionisation, termed* sporadic *E*, is often present

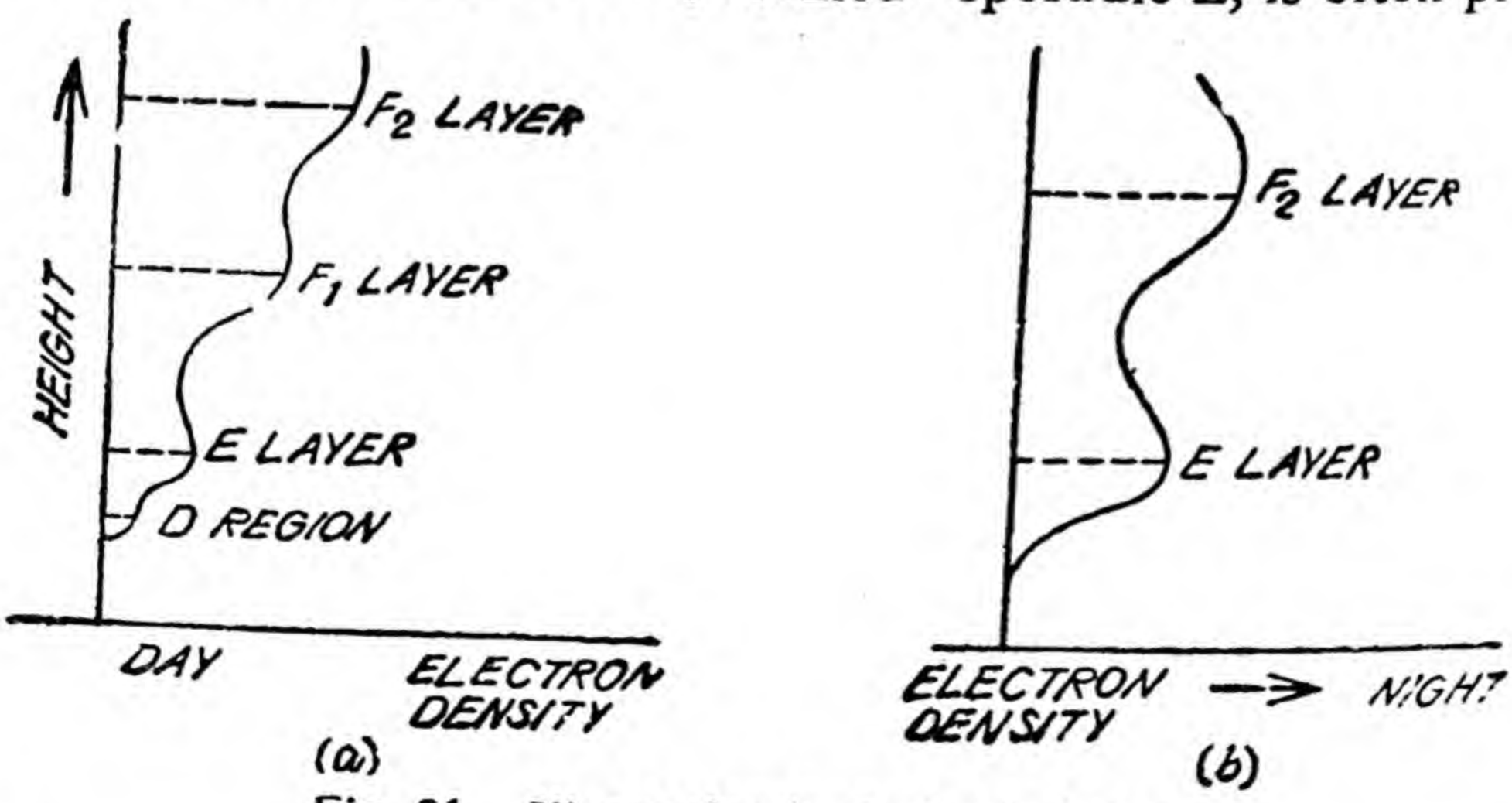


Fig. 21. Illustrating ionospheric layers.

*Presence of this region in *E* layer is quite irregular and hence it is termed as sporadic *E* layer.

in the E region in addition to the regular E ionisation. It often occurs in the form of clouds, varying in size from roughly 1 km. to several hundred kilometers across. Such sporadic E ionisation occasionally will reflect v.h.f. waves upto 60 or 70 mc/s, giving very long distance transmission on such frequencies. The occurrence of sporadic E layer is quite unpredictable. It may be observed both in the day and night. The cause of sporadic E ionisation is still uncertain.

Variation of critical frequency and virtual height of layers with time of the day and of the year :

Refer to fig 22. in which monthly averages of critical frequencies and virtual heights as a function of time of day for winter and summer, and for two parts of the 11 year sunspot cycle are shown, The following predictions can be laid for various layers :

D layer : Virtual height of reflection lies between 70 and 80 km. It is able to reflect back very long waves (> 10 km) incident on it. Reflection of waves is not due to bending but corresponds to the process of reflection from a boundary discontinuity. The maximum electron number density is few thousand electron per cm^3.

E layer : The virtual height of E layer remains substantially constant between 110 and 120 km. throughout the day and from season to season. However, the critical frequency of the E layer shows diurnal and seasonal variation. It increases with the altitude. of the sun and is maximum at noon. It has been shown that f_E, the critical frequency of E layer is given by

$$f_E = k \sqrt[4]{(\cos \psi)},$$

where ψ is the zenith angle of the sun and k is a factor that depends upon the intensity of radiation from the sun. The relation is not valid when $\psi \approx 90°$. In the night, f_E falls below 1 mc/s.

F_1 layer : It exists only in day time. The virtual height is of the order of 225 km during the middle of the day and is consistently maintained from season to season. Variations in critical frequency of F_1 layer are the same as in case of E layer, *i.e.* the maximum occurs at noon, and summer values are greater than winter values. However, it does not obey any such relation as f_E follows.

F_2 layer : The virtual height and critical frequency of this layer show quite a large variation during the day and from season to season. Virtual height in summer ranges from 300 to 400 km., whereas in winter it goes down to 225 km. The critical frequency is much higher in winter than in summer.

In the night F_1 layer disappears whereas F_2 layer continues throughout the night at a height of 300 km. approximately.

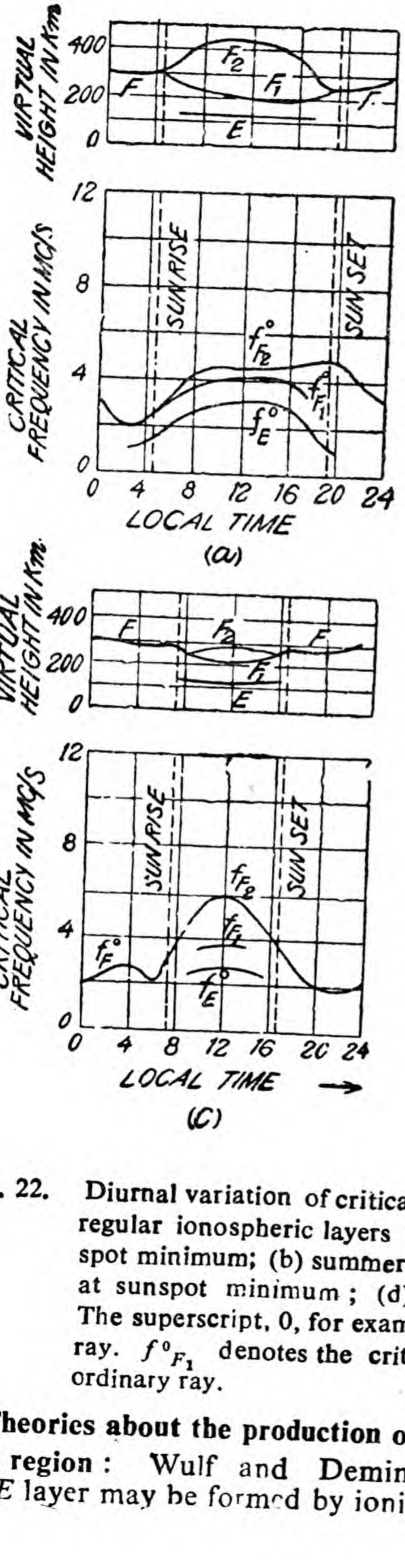


Fig. 22. Diurnal variation of critical frequency and virtual height of regular ionospheric layers (a) summer at period of sunspot minimum; (b) summer at sunpot maximum ; (c) winter at sunspot minimum ; (d) winter at sunspot maximum. The superscript, 0, for example in $f^0_{F_1}$, refers to ordinary ray. $f^0_{F_1}$ denotes the critical frequency of F_1 layer for ordinary ray.

Theories about the production of layers :

E region : Wulf and Deming independently pointed out that *E* layer may be formed by ionisation of O_2, at the first ioni-

sation potential, in the region of transition where molecular oxygen is dissociated and where, as a consequence, its density falls of with height much more rapidly than for a simple isothermal atmosphere. Bhar assumed the ionisation at second ionisation potential. Objections, however, were raised to both these assumptions.

Nicolet pointed out the existence of a number of strong absorption bands, superimposed on the feeble absorption continuum of O_2 in the energy range (12·20 to 13·55 eV). According to him, *E* layer at the observed height is produced by pre-ionisation of O_2 in transition region, which is a result of absorption in the strong bands, and this view is now generally accepted.

D region : The virtual height of reflection of these waves is found to be between 70 and 80 Km. Thus D region ionisation lies between 70 and 90 km. The most comprehensive view is due to suggestions by Nicolet, which follows :

(*a*) a normal layer due to the ionisation of O_2 at its first ionisation potential,

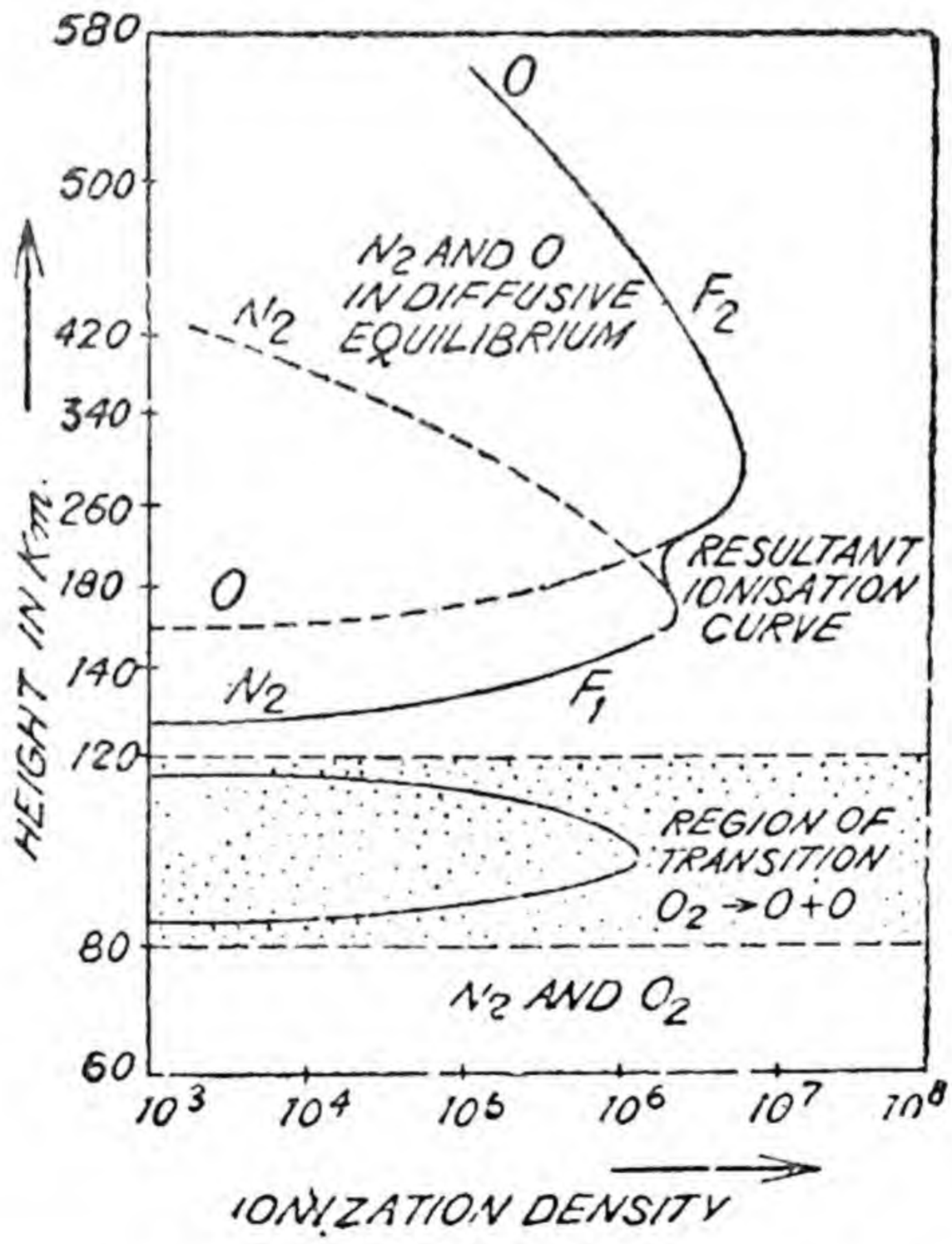


Fig. 23. Curves illustrating ionisation distribution with height.

(*b*) a layer more or less sporadic due to ionisation of sodium, and

(*c*) an extra ionisation during fade out due to ionisation of *NO* by $\lambda = 1300°$ A.

Regions F_1 and F_2 : Bhar in his work attributed region F_1 to the ionisation of N_2 at its second ionisation potential and region F_2 to the ionisation of *O* at its first ionisation potential. The theo-

TABLE : Characteristics of various regions in the Ionosphere :

Particulars	D-region	E-region	F_1-region	F_2-region
Level of maximum ionisation (km)	60—90 by day; disappears at night	100	200	300
Molecular density per c c	$10^{14}-10^{16}$	$5\times10^{11}-5\times10^{13}$	about 10^{11}	about 10^{10}
Electron density per c.c.	$10^{2}-10^{3}$ electrons $10^{6}-10^{8}$ for ions	5×10^{3} to 5×10^{4} at night	$2\times10^{5}-4{\cdot}5\times10^{5}$	max. 2×10^{6} by day in winter max. 2×10^{5} by day in summer
Collisions per second	10^{7} at lower edge	10^{5}	10^{4}	$10^{3}-10^{4}$
Recombination Coeff. cm.³/sec.	$10^{-5}-10^{-7}$	10^{-8}	4×10^{-9}	8×10^{-11} by day 3×10^{-11} by night
Origin	(i) O_2, first ionisation potential (12·2 eV) (Mitra, Bhar and Ghosh) (ii) Ionisation of metals, sodium in particular (iii) Ionisation of NO.	(i) O_2, first ionisation potential, (Wulf and Deming) (ii) O_2, second ionisation potential (Bhar, Mitra and Ghosh) (iii) $O(^1S)+O(^1S)\rightarrow O_2^{+}+e$ (Gauzit) (iv) Pre-ionisation of O_2 due to strong absorption bands in the energy range 12·2 to 13·55 eV (Nicolet) (v) High energy photon theory (Bates and Hoyle)	(i) N_2, second ionisation potential (Mitra, Bhar and Ghosh, Wulf and Deming) (ii) Ionization of O, at its first ionisation potential (13·5 eV) giveng rise to F_1 layer at the height of the maximum ion production and to F_2 layer as a result of height dependent recombination coefficient (Mohler, Bates and Massey: Bates, Martyn)	(i) N_2, first ionisation potential (15·5 eV), (Wulf and Deming) (ii) O, first ionisation potential (13·5 eV) (Bhar, Mitra, Ghosh) (iii) due to (ii) point of F_1 region. (iv) N_2, first ionisation potential, and *O* second and third ionisation potentials (Massey)

retical ionisation distribuions are shown in the figure 23. It was suggested and advanced by Bates and Massey that F_1 region is produced not by ionisation of N_2 but by ionisation of atomic oxygen. Hence according to the accepted view, it is produced by ionisation of atomic oxygen. The new conception obtained exceptionally agrees well with Chapman's theory.

It is further proposed that F_1 being produced in the normal manner by photo ionisation of O_2, the F_2 region is produced by a sort of bifurcation of the same according to one of the two possible prcocsses :

(*i*) The one. favoured by Mohler, and by Bates and Massey is that the effective recombination coefficient decreases with height so rapidly that a maximum of equilibrium electron density is formed at some height above the F_1 region maximum. The region appears as double, maximum observed at level of F_1. another at the level of F_2 with a depression at an intermediate level.

(*ii*) The other process advocated by Martyn is that the bifurcation is caused by the motion of the ions in the earth's magnetic field.

Refer to Table for comparative study on page 1372.

29·9. CHAPMAN THEORY OF LAYER FORMATION :

The regular ionised regions of upper atmosphere owe their origin primarily to solar ultra violet radiation. The intensity of ionising radiation increases with height, whereas the density, ρ, of atmospheric particles decreases with height, and its variation with height is given by

$$\rho = \rho_0 e^{-h/H}, \qquad \ldots(1)$$

where ρ_0 is the density at the reference level $h=0$ and H is a constant equal to $\frac{KT}{mg}$, called scale height (value of H is about 10 km. In this expression K is Boltzmann constant). Thus, the two opposite processes—one increase in the intensity of solar radiation with height, and the other decrease in the density of ionisable particles with height lead to the formation of strata of maximum ionisation at different heights These strata are called ionospheric layers.

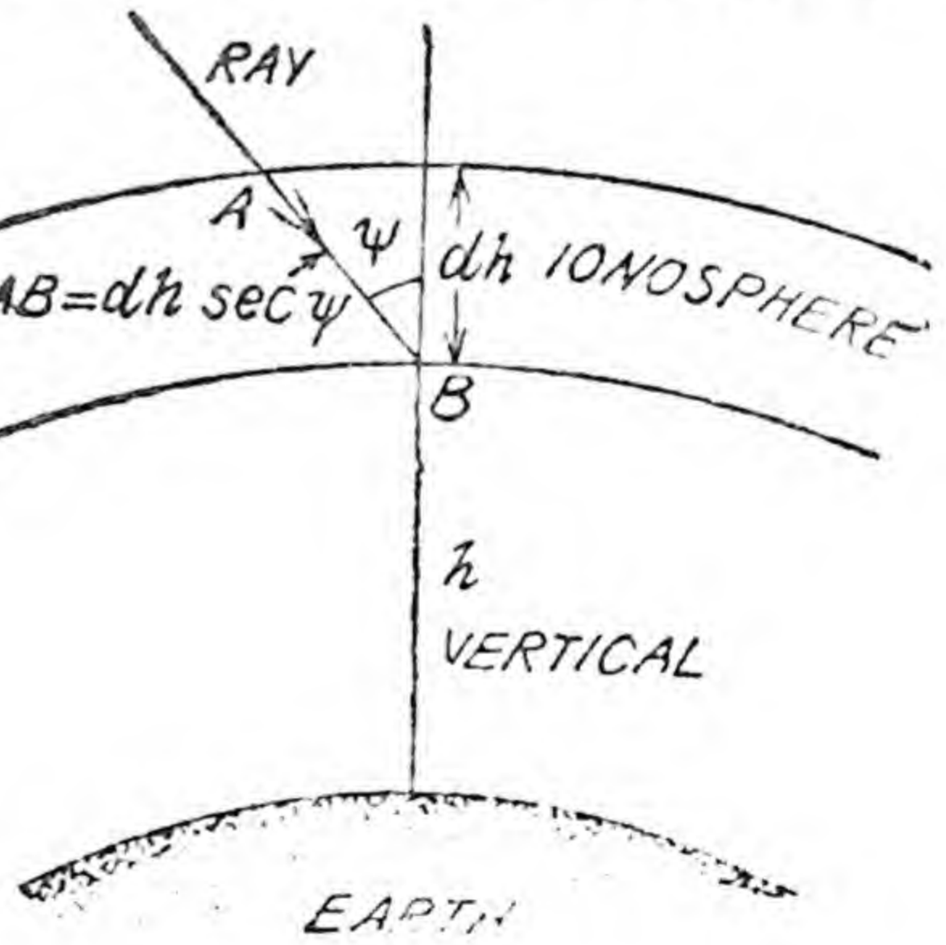


Fig. 24.

Let us consider solar monochromatic radiation arriving at an angle ψ to the vertical and fall in the intensity I of radiation in traversing the distance AB ($=dh \sec \psi$) is dI per unit area of cross section of the atmosphere.

Let A be the absorption coefficient per unit mass of the gas per unit intensity of radiation, then absorption of intensity in travelling the distance AB will be

dI = Absorption coefficient × Mass of atmospheric column of height $dh \sec \psi$ and unit area of cross section × Intensity

$$= A \times \rho\, dh \sec \psi \times I$$
$$= \rho\, AI\, dh \sec \psi. \qquad \ldots(2)$$

Putting the value of ρ from equation (1) and integrating between the limits ∞ to h, we get

$$\int_{\infty}^{h} \frac{dI}{I} = \int_{\infty}^{h} A \sec \psi\, \rho_0 e^{-h/H}\, dh.$$

But on left hand side we should impose the limits of intensity *i.e.* from I_0 to I, where I_0 is the intensity of radiation before it enters the atmosphere, *i.e.* value of I at $h=\infty$. Therefore,

$$\int_{I_0}^{I} \frac{dI}{I} = \left[- A\rho_0 H \sec \psi\, e^{-h/H} \right]_{\infty}^{h}$$

or $$\log (I/I_0) = - A\rho_0 H \sec \psi\, e^{-h/H}$$

or $$I = I_0 \exp \left[- A\rho_0 H \sec \psi\, e^{-h/H} \right]. \qquad \ldots(3)$$

Let β be the number of ions produced due to absorption of unit quantity of radiation. For an amount dI fall in the intensity of radiation, the total ions produced will be $\beta\, dI$. The rate of ion production will be governed by the rate of fall of dI, with distance AB ($=dh \sec \psi$). So the rate of ion production, q, at the height h, is given by

$$q = \beta \frac{dI}{dh \sec \psi}$$
$$= \beta \cos \psi \frac{dI}{dh}. \qquad \ldots(4)$$

Differentiating equation (3) with respect to h,

$$\frac{dI}{dh} = I_0 \left[- A\rho_0 H \sec \psi . e^{-h/H} \left(- \frac{1}{H} \right) e^{- A\rho_0 H \sec \psi\, e^{-h/H}} \right], \qquad \ldots(5)$$

Putting this in equation (4), we get

$$q = \beta A \rho_0 I_0\, e^{- h/H - A\rho_0 H \sec \psi\, e^{-h/H}} \qquad \ldots(6)$$

To find the height at which q is maximum, we have the condition

$$\frac{dq}{dh} = \beta A I_0 \rho_0 \left[e^{- h/H - A\rho_0 H \sec \psi . e^{-h/H}} \right] \times \left[- \frac{1}{H} + A\rho_0 H \sec \psi . e^{-h/H} . \frac{1}{H} \right] = 0$$

or $$1 = A\rho_0 H \sec \psi . e^{-h/H}$$

or $$e^{h/H} = A\rho_0 H \sec \psi.$$

Denoting this height as $h_{max.}$, we write

$$\exp\left(\frac{h_{max.}}{H}\right)=A\rho_0 H \sec\psi. \qquad ...(7)$$

Now the maximum rate of ion production will be

$$q_{max}=\beta A\rho_0 I_0\left(\frac{1}{A\rho_0 H\sec\psi}\right)\exp\left[-A\rho_0 H\sec\psi.\frac{1}{A\rho_0 H\sec\psi}\right]$$

after substituting equation (7) in equation (6). Therefore

$$q_{max}=\beta A\rho_0 I_0\left(\frac{1}{A\rho_0 H\sec\psi}\right).e^{-1}$$

$$=\frac{\beta I_0}{H\sec\psi}.e^{-1}. \qquad ...(8)$$

At mid day when $\psi=0$, equations (7) and (8) give respectively

$$\left.\begin{aligned} e^{h/H}&=A\rho_0 H\\ \text{and}\qquad q_0&=\frac{\beta I_0}{H}e^{-1}, \end{aligned}\right\} \qquad ...(\ )$$

where q_0 and h_0 are the values of q_{max} and h_{max} for $\psi=0$.

Equation (6) can now be written in the form

$$q=\frac{\beta I_0}{H}e^{-1}.He^{+1}.A\rho_0.\exp\left[-\frac{h}{H}-A\rho_0 H\sec\psi.e^{-h/H}\right],$$

which on using equation (9) becomes

$$q=q_0\,e^{h_0/H}.e^{+1}\exp\left[-\frac{h}{H}-A\rho_0 H.\sec\psi.e^{-h/H}\right]$$

$$=q_0\exp\left[1+\frac{(h_0-h)}{H}-A\rho_0 H\sec\psi.e^{-h/H}\right]$$

Again putting $A\rho_0 H=e^{h_0/H}$ in exponent, we arrive at

$$q=q_0\exp\left[1+\left(\frac{h_0-h}{H}\right)-e^{h_0/H}\sec\psi.e^{-h/H}\right]$$

$$=q_0\exp\left[1+\left(\frac{h_0-h}{H}\right)-\sec\psi\,e^{(h_0-h)/H}\right]. \qquad ...(10)$$

If we take h_0 as the reference level from which the height h is measured and the reduced height z is expressed as

$$z=\frac{h-h_0}{H},$$

then equation (10) can be expressed as

$$q=q_0\,e^{1-z-\sec\psi.e^{-z}}, \qquad ...(11)$$

which is the desired equation that gives the rate of ion production as a function of reduced height z and the zenith angle ψ.

Deductions from Chapman's Theory :

(*A*) ***Variation of electron number density with zenith angle of the sun :***

For the ionospheric regions the electron number density changes slowly with time during full day light hours. It means that at any instant in day light hours rate of ion production q and ion decay approximately balance each other *i.e.* variation of ion number density N with time is zero $\left(\frac{dN}{dt}=0\right)$. Now if N be ion number density at any instant t and at a height h, then assuming that recombination is the only process of decay, number of ions lost per unit volume per second $\propto N^2=\alpha N^2$.

At the balancing condition occuring at some instant during day light hours, rate of production must be equal to the rate of recombination *i.e.*

$$q=\alpha N^2$$

or
$$q-\alpha N^2=0=\frac{dN}{dt}.$$

Therefore,
$$N^2=\frac{q}{\alpha}=\frac{q_0}{\alpha}e^{1-z-\sec\psi\, e^{-z}}$$

Variation of electron number density (N) with reduced height (z) for various values of zenith angle (ψ) is shown in fig. 25.

(*B*) *Particle concentration of active constituent at the height of maximum of layer* :

The density of the active constituent at any height is given by

$$\rho=\rho_0 e^{-h/H}.$$

From equation (7), we have

$$\rho=\frac{\rho_0}{AH\rho_0\sec\psi}\approx\frac{\cos\psi}{AH}.$$

This gives the particle concentration of the active constituent at the height maximum of layer (h_{max}).

In figure 25 variation of electron number density, N, with height h according to the simply Chapman theory of a flat earth, for various values of the sun's zenith angle ψ, is depicted; while in figure 26 dependence of electron collision frequency, ν, with height, h, is shown.

From figure 25, it is obvious that for curve $\psi=0$, maximum occurs at $\frac{h-h_0}{H}=0$ *i.e.* $h=h_0$ at which N should be N_0. Therefore at this point $N/N_0=1$.

It is found that changes in the value of ν affect the propagation of radio waves far less than changes of the electron number density N. For many purposes, it is therefore permissible to treat ν as constont over a small range of height h. This is especially

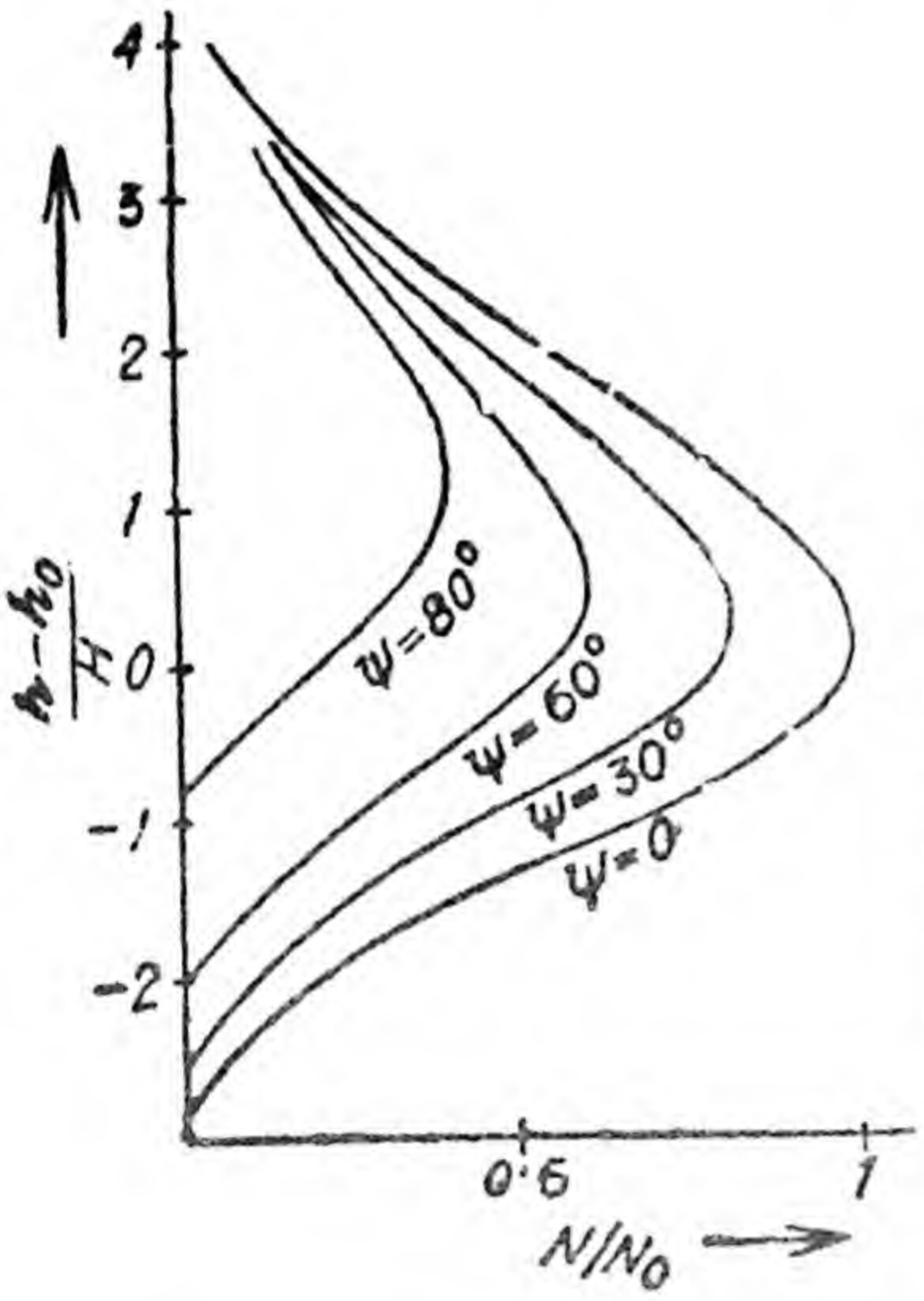


Fig. 25. Variation of electron number density with height. N_0 corresponds to height h_0 with $\psi=0$.

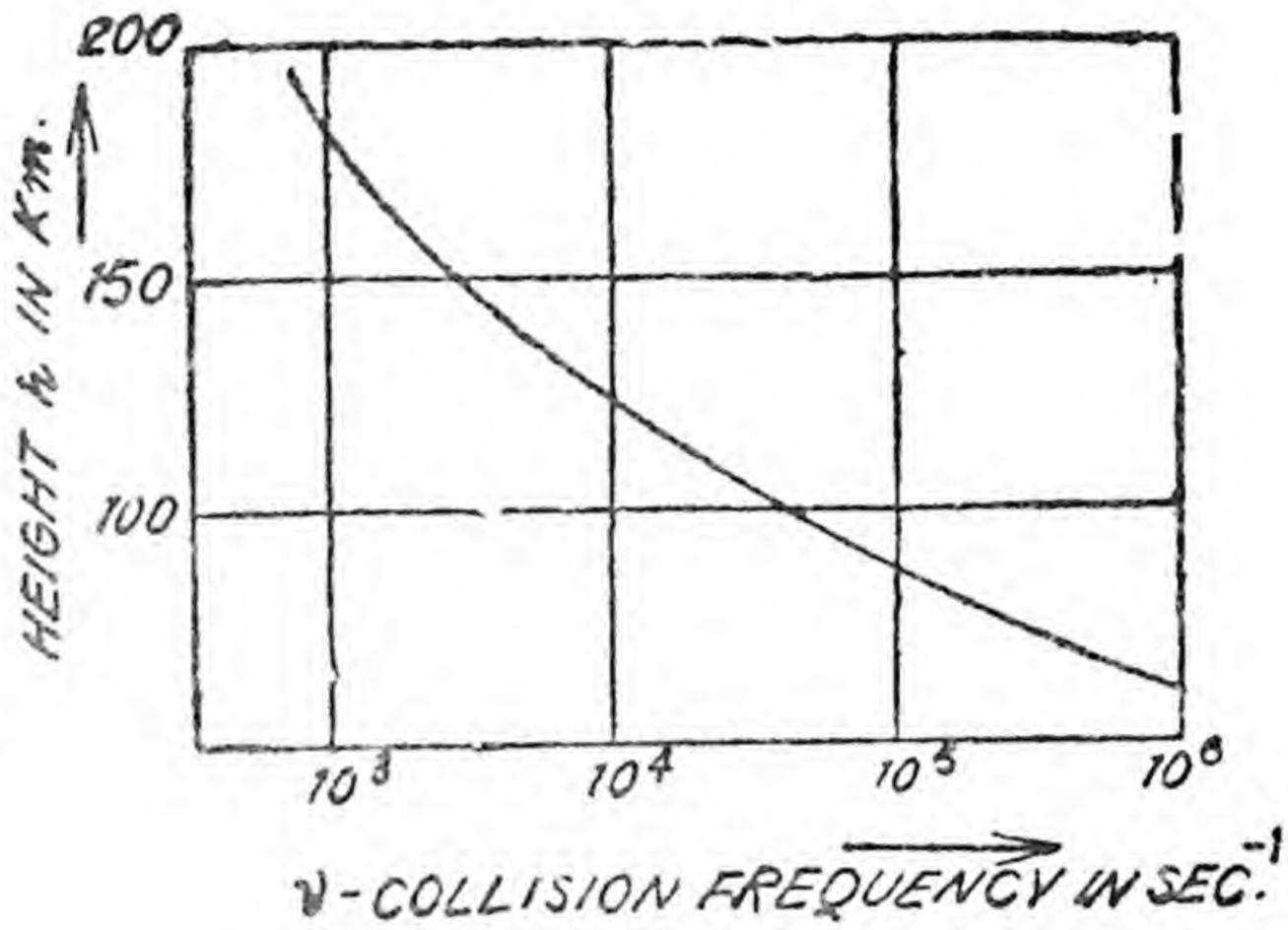


Fig. 26. Variation of collision frequency with height h.

true at high frequencies (greater than about 1 Mc/s.) where the wavelength is small compared with the scale height H, which is about 10 km.

29·10. CONSIDERATIONS IN THE PROPAGATION OF SHORT WAVES :

In this article, we shall account for the considerations which ought to be taken into account in deciding the frequency of transmission for short wave inter-continental communication, and also why and when is it necessary to alter the frequency for such a communication.

The term short waves applies to wavelengths from 100 down to 10 metres. The main mode of propagation of short waves is sky wave due to two reasons :

(i) at these frequencies the ground wave is attenuated rapidly and its range is short, just a few tens of kilometers.

(ii) the sky wave signal is capable of multiple reflections from the ionosphere and can cover long distances because they experience an insignificant attenuation which *decreases with the increase of frequency*. Under normal conditions, D and E regions absorb the energy of the sky wave and F region reflects it back to the earth. The absorption in F_2 layer can be neglected in comparison with that in the E region.

For short wave inter-continental communication two conditions must be satisfied simultaneously :

(*i*) *the frequency used must be lower than maximum usable*

frequency value determined for a propagation path and for a given ionisation of the reflecting layer.

This sets the upper limit of usable frequency. If the frequency of the wave is greater than the maximum usable frequency for that path and ionisation then the waves will fail to reach the desired point of reception, despite any increase in transmitter power.

(*ii*) *signal strength should be sufficient as to overcome the absorption it will meet in D and E regions and yet strong enough for reliable reception.*

This sets the lower limit of usable frequency but this condition is not so critical as the first because the use of a lower frequency may be compensated by an increase in the transmitter power.

Alteration of frequency in short wave communication :

During day time, ionisation density of F_2 layer is high and it can reflect the higher frequencies whose absorption in *D* and *E* regions will be insignificant. But in night, ionisation density of F_2 layer goes down and therefore, for reflection from this layer, usable frequency of the waves must be lowered so as to be able to reach the desired point of reception. Though, at lower frequencies, absorption is increased but, as in the night ionisation density of *E* is reduced and *D* layer disappears, the absorption of the waves becomes insignificant.

Suppose a station keeps on using '*day time*' frequencies at night then a situation may arise when the ionisation density, *N*, of F_2 layer drops to a value that is unable to reflect the wave *i.e. the transmitter frequency exceeds the critical frequency*. Consequently, the wave penetrates the F_2 layer (fig. 27 *a*).

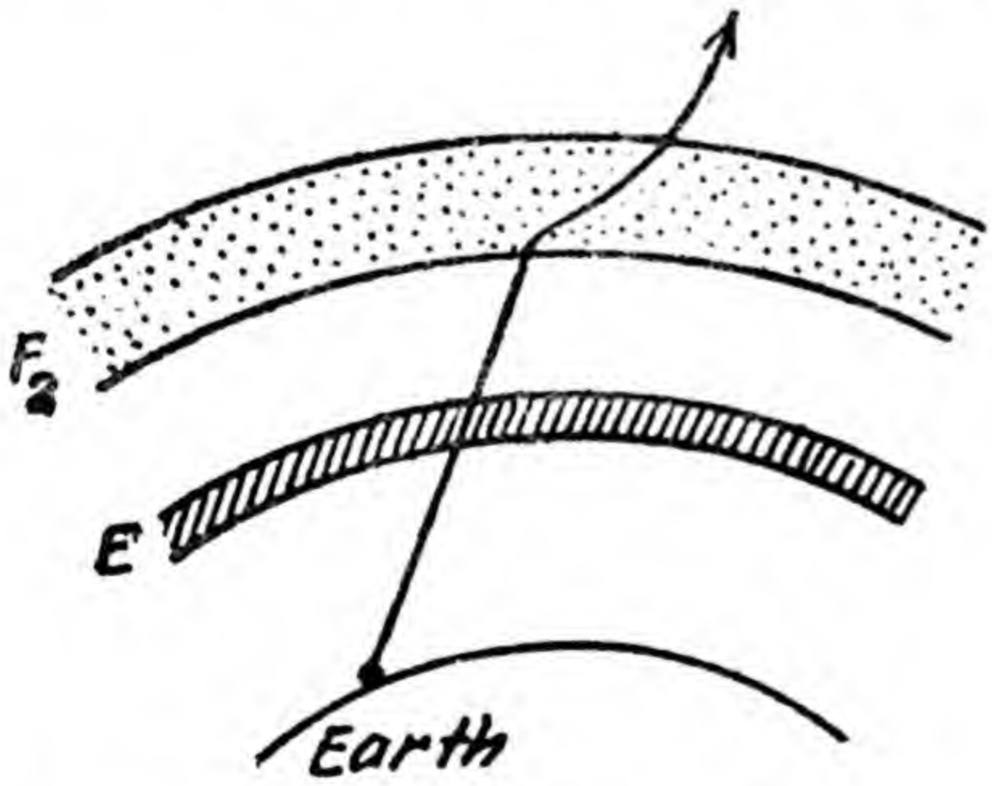


Fig. 27. (*a*). At night, when wave frequency is greater than critical frequency ($f > f_c$).

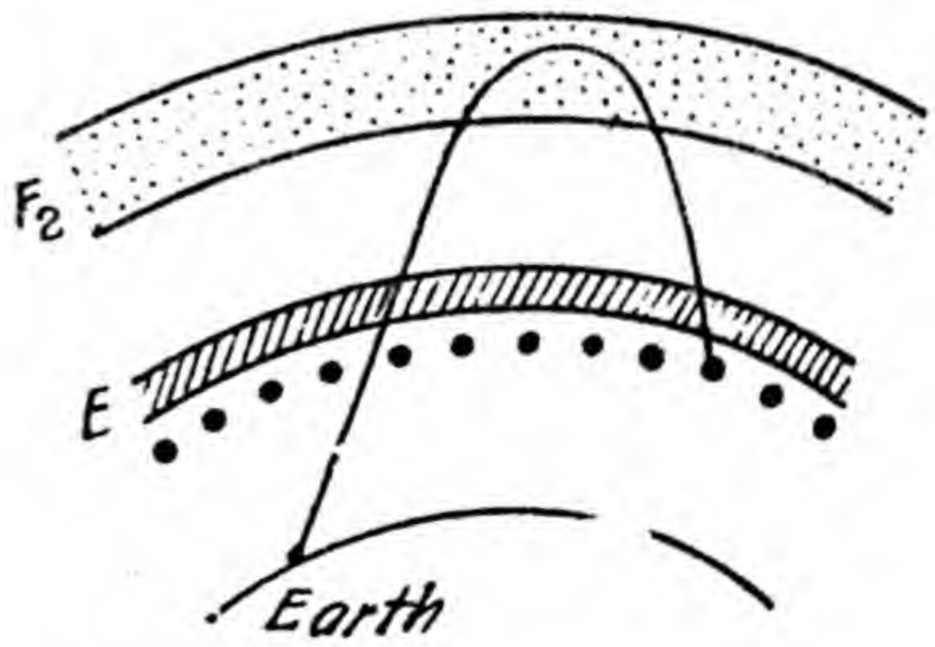


Fig. 27. (*b*). At day time *D* region appears, *E* is highly ionised and so they absorb night time frequency wave.

Suppose a station keeps on using '*night time*' frequencies by day. Since this frequency is below the critical frequency, it will be reflected from F_2 layer but will be absorbed during its passage

through *D* and *E* regions. The strength of the wave is so reduced as to be useless for reliable reception.

It thus becomes obvious that during night, wave frequency must be brought down for faithful communication.

Other factors affecting the short wave communication :

(*i*) *Solar activity* : The earlier discussion applies to normal conditions of short wave propagation. But during periods of high solar activity in summer, situation becomes different. Then *E* region ionisation increases to such an extent that it often acts as a reflecting layer, instead of F_2, for short waves and only *D* region absorption remains.

(*ii*) *Sporadic E layer* : Normal conditions of short wave propagation are also upset by sporadic *E* layer which sometimes has fairly high ionisation density. As it lies below F_2 layer, it then acts as a reflecting layer for short waves.

Major advantage of short wave communication compared to long and medium waves is that they suffer insignificant attenuation in the ionosphere.

29·11. PROPAGATION OF RADIO WAVES OF DIFFEPENT FREQUENCIES :

The propagation mode of a radio wave closely depends upon its frequency. In the present article, we shall discuss the propagation of different frequencies :

(1) Very low frequencies (10 to 30 kc) : The primary mode of propagation in this frequency range is the surface wave. Because of the relatively low attenuation, propagation over distance of several thousand miles is affected by this wave. The main disadvantage with this mode of propagation is its requirement of high transmitter output and huge antenna installations.

There is a possibility of sky wave propagation too. Skywaves are reflected back to earth with very little absorption after a slight penetration into the ionosphere (or from its lower edge). Since the ionisation density, on which attenuation caused by the reflection at the ionosphere depends, changes with diurnal and seasonal variations, the received signal intensity will also exhibit the similar variations being strongest at night and in winter. This variation for the signal for low frequency end is least and practically zero near 10 kc/s.

(2) Low Frequencies (30 to 300 kc) : Because of the increased ground wave attenuation, the range of surface wave propagation decreases and long distance communication becomes entirely dependent on sky waves. Also, with the increase in frequency, attenuation at the lower edge of the ionosphere becomes more appreciable in day time and in summer than during the night and winter. This causes the received signal strength to vary apprecia-

bly and hence long distance communication is not possible in the day particularly in summer.

(3) **Medium Frequencies (300 to 3,000 kc) :** This range of frequencies is used for broadcast purposes. For better reception, signal received should be strong, free from fading and interference.

Skywaves at broadcast frequencies are completely absorbed in day time and consequently propogation in the day is entirely due to ground waves. Since the ground wave attenuation is inversely proportional to the earth's conductivity and directly to the square of the frequency, signals in day time due to this wave die out very rapidly if earth conductivity decreases and signal frequency increases.

It is necessary to deliver the receiver a signal stronger as compared to ordinary interference or noise level The area about a transmitter, in which signal strength in day time is sufficient to override ordinary interferences is called *primary service area* or *primary coverage area*, or simply the distance upon which the ground wave field strength is above noise level is called primary coverage. The area outside the receiving primary coverage, where the field strength does not override interference fully but still gives a useful signal, is called *secondary coverage area* or simply the distance upon which the field strength of sky wave is above the noise level is called secondary coverage. Primary coverage area depends upon transmitter power, directivity of transmitting antenna, earth's conductivity and the frequency. The same factors determine the secondary coverage area. Ground wave primary coverage area does not exceed 100 to 150 km.

In the region over which the sky wave field intensity is comparable to that of the ground wave (at night) increased distortion occurs and fading is different at different frequencies in a frequency band occupied by the modulated wave (carrier and side band frequencies). Such type of fading is called *selective fading*.

At night a strong skywave is returned to the earth by the layer and serves to extend considerably the propagation range.

(4) **High Frequencies (3 to 30 mc/s) :** For such high frequencies, ground wave attenuation is very large and therefore, this mode is limited to very short propagation range. Sky waves are used for long distance communication the range being a function of signal frequency, ionospheric conditions and the angle of incidence of transmitted wave at the ionised layer (skip distance and maximum usable frequeucy, which clearly specify the conditions of skywave propagation depend on these factors).

Energy absorption by the ionosphere at this frequency range is very low due to which ranges of thousands of miles are obtained. The waves are reflected back from *E* or *F* layers. Appearance of *D* layer in day time causes appreciable attenuation for lower frequencies

(5) Very High Frequency (30 to 300 mc/s) : Surface waves, because of their high attenuation, and sky waves because of their failure to be reflected, almost fail to propagate signals of this frequency range. Space wave or line of sight propagaion is practically the only mode of propagation in this frequency range. Due to the refraction in earth's atmosphere, propagation beyond the line of sight is possible. This frequency range is adopted in short distance communication, radio relay systems, propagation of frequency modulated signals and television signals.

(6) Ultra High Frequencies and Microwaves (above 300 mc/s) : Only space wave propagation is possible. With the formation of inversion layers in troposphere, duct propagation is also possible. Tropospheric absorption is high for 3.10^{10} c/s , *i.e.* wavelength 1 cm. or less. Microwaves are used in radar systems.

29·12. IONOSPHERIC STORMS

Sometimes it is noted that the absorption of skywaves increases very much and also the critical frequencies of E and F_2 layers suffer an abnormal change. This all is the result of the ionospheric disturbances called ionospheric storms which last for days together.

All ionospheric storms are associated with severe disturbances of the terrestial magnetic fields called magnetic storms. Due to the violent turbulence of ionospheric storms, layers with definite virtual heights and critical frequencies tend to disappear and the normal stratification of the ionosphere is destroyed. Small clouds of ionisation are produced which move in an irregular manner. The effects of such storms are appreciable near the polar regions and almost negligible at the equator.

NUMERICAL EXERCISES

Ex. 1. *Compute the values of relative permittivity of D, E and F regions of the ionosphere in the case of an electromagnetic wave with a frequency of* 50 *mc/s. Given that electron density (no. of electrons/c.c.) for D, E, F regions is* 400, 5×10^5 *and* $2{\cdot}02 \times 10^6$ *electrons/c.c. respectively.*

The relative permittivity is given by the relation

$$\epsilon_r = 1 - \frac{80{\cdot}6N}{f^2},$$

where if N is in electrons/c.c. then f is in kc/s.

For D region :

$$\epsilon_r = 1 - \frac{80{\cdot}6 \times 400}{(50 \times 10^3)^2},$$

because we are to put f in kc/s.

Therefore

$$\epsilon_r = 1 - 1{\cdot}29 \times 10^{-5} \approx 1.$$

For E region :

$$\epsilon_r = 1 - \frac{80{\cdot}6 \times 5 \times 10^5}{(50 \times 10^3)^2}$$

$$= 1 - 1{\cdot}61 \times 10^{-2} \approx 0{\cdot}9838.$$

For F region :

$$\epsilon_r = 1 - \frac{80{\cdot}6 \times 2{\cdot}02 \times 10^6}{(50 \times 10^3)^2}$$

$$= 1 - 0{\cdot}651 \times 10^{-1} \approx 0{\cdot}9348.$$

Ex. 2. *Calculate the value of frequency at which an electromagnetic wave must be propagated for the D-region to have an index of refraction of* 0·5.

We know that index of refraction of the ionised region is given by

$$\mu = \sqrt{\left(1 - \frac{80{\cdot}6N}{f^2}\right)}$$

where if N is in electrons/c.c. then f is in kc/s.

For D region, $N = 400$ electrons/c.c. Therefore we have the relation

$$0{\cdot}5 = \sqrt{\left(1 - \frac{80{\cdot}6 \times 400}{f^2}\right)}$$

or $$0{\cdot}25 = 1 - \frac{32240}{f^2}$$

or $$f^2 = \frac{32240}{0{\cdot}75} = 4{\cdot}2986 \times 10^4$$

or $$f = 207{\cdot}33 \text{ kc/s.}$$

Ex. 3. *A skywave with a frequency of* 50 *mc/s is incident on the D-region at an angle of* 30°. *Find the angle of refraction. Given electron number density for D-region (per c c.) is* 400.

We know the relation,

$$\mu = \frac{\sin i}{\sin r} = \sqrt{\left(1 - \frac{80{\cdot}6N}{f^2}\right)},$$

where N is in electrons/c.c., f is in kc/s, i and r are the angle of incidence and refraction respectively.

Putting

$N = 400$ electrons/c.c.

$f = 50.10^6 = (50.10^3)$ kc/s.

We get

$$\mu = \frac{\sin i}{\sin r} = \sqrt{\left(1 - \frac{80{\cdot}6 \times 400}{(50 \times 10^3)^2}\right)}$$

$$= \sqrt{(1 - 1{\cdot}29 \times 10^{-5})}$$

$$= 1$$

or $$\sin i = \sin r$$

or $$i = r = 30°.$$

EXERCISES AND PROBLEMS

1. Give a brief account of the effect of the ionosphere on the sky waves. Show that the ionosphere behaves as a medium of refractive index

$$\mu = \left\{ 1 - \left(\frac{80\ 6N}{f^2} \right) \right\}^{1/2},$$

where N is electron density per c.c. and f is frequency in kc/s. The effect of the earth's magnetic field and collisions in the ionosphere may be neglected.

2. Show that the ionosphere under the action of the earth's magnetic field behaves like a birefringent crystal.

3. Write short notes on :
(i) Skip distance, and
(ii) Maximum usable frequency.
Obtain an expression for D_{skip} in terms of critical frequency of the layer.

4. Determine the effective dielectric constant of a medium containing electrons subjected to friction. Apply the result to explain the reflection of radio waves from the ionosphere. Why do the waves of very high frequencies not get reflected from the ionosphere.

5. Discuss in detail the pulse method of measuring the height and electron concentration of an ionospheric region.
Give the block diagram of a suitable pulse transmitter that may be used for ionospheric measurement. What is the effect of gradually varying the wave frequency of the pulse used ?

6. How is the ionosphere instrumental in achieving long distance radio communication ? For what frequencies in general is the ionosphere useful and why ?

What are the modern views regarding the stratification in the ionosphere.

7. State and discuss the conditions of total reflection of radio waves from the ionosphere.
Prove that the skip distance D for a given frequency is given by

$$D = 2h \sqrt{\left\{ \left(\frac{f}{f_c} \right)^2 - 1 \right\}},$$

where h is the effective height of the reflecting layer and f_0 the critical frequency of the layer.

8. In a simple ionopheric height oscillograph record, the distance between the ground and the echo pulse is 1·6 cms. If the sweep speed is 20 metres/second, find the effective height of the ionospheric layer assuming the speed of electromagnetic waves in free space as 3×10^6 km/second.

9. Describe the structure of ionosphere. Discuss the theory which explains the formation of ionosphere.

10. Derive formulae for magneto-ionic refractive indices and absorption coefficient. Discuss their application to ionosphere.

11. How does the presence of electrons affect the dielectric constant of the ionosphere ? Derive the expressions concerned. Why do the waves of very high frequencies not get reflected from the ionosphere ?
12. Discuss the considerations which ought to be taken into account in deciding the frequency of transmission for short-wave inter-continental communication. Why and when is it necessary to alter the frequency for such a communication ? (Art. 29·10).
13. Show that, if a vertical ray of frequency f_v and an oblique ray of frequency f_0 are reflected from regions which are at the same real heights of the ionisphere, the virtual height (of reflection) is also same for both the rays. (Art. 29·6).
14. Describe briefly the essential components of an ionospheric recorder. Draw a typical ionogram indicating parameters normally scaled.
15. Explain why is the ratio of lower frequency limit to MUF for carrying on radio communication at short wave frequencies is generally smaller at night than in day time (Art. 29·10).
16. Explain why is it reasonable to except that the optimum short wave frequency will be high at low altitudes.
17. Discuss theoretically the method of transmission of radio waves through the ionosphere at oblique incidence and show how the MUF for communication between two distant stations can be calculated for thin ionised layers. (Art. 29·5).

30 MODULATION

For the transmission of message to distant parts of the globe, sound waves are first converted into electrical signals with the help of microphone. To release this audio signal into space, antenna is employed. Refer to chapter 28 (antenna action), where it has been pointed out that for effective radiation of energy, antenna should be fed with high frequency electrical signal. But the frequency of audio signal is always quite low and consequently, it can not be fed as such to the antenna for communication. Therefore, before applying the audio signal to the antenna, a process is performed in which the audio signal is superimposed on a high frequency wave, called carrier wave. In this process, as mentioned, a separate high frequency wave is needed because we can not change any of the characteristics *e.g.*, amplitude, frequency and phase of the audio signal as it would amount to a change in the message to be communicated. But the amplitude, or frequency or phase of the high frequency wave (called carrier wave) is modified inaccordance with the audio signal (called modulating signal), so that the resultant wave (called modulated wave) inherits the signal. This process of superimposition of audio signal over carrier wave is called modulation. Since the frequency of modulated wave is quite high, effective radiation from antenna takes place.

Thus modulation is the process in which the frequency level of intelligence is altered. There are thus two main reasons for this alteration in frequency level :

(*i*) at high frequencies, the intelligence practically can be transmitted by radiation, and

(*ii*) different messages having different frequency level can be transmitted simultaneously without any interference.

30·0. DEFINITIONS :

Modulation is defined as the process by which some characteristic of a high frequency wave such as amplitude, frequencyor phase, is altered inaccordance with the instantaneous value of some other voltage called the modulating voltage.

Modulating voltage is the signal while the voltage undergoing the alteration is a high frequency wave and is called the carrier wave.

Let the carrier voltage be expressed as

$$e_c = E_c \cos(\omega_c t + \theta),$$

where E_c is peak amplitude of carrier voltage, ω_c is angular frequency of carrier wave and θ is the phase angle.

Three cases may arise :

(*i*) When amplitude E_c is varied inaccordance with the modulating wave, the process is known as amplitude modulation; ω_c and θ remain constant.

(*ii*) When ω_c varies inaccordance with the modulating wave, while E_c and θ remain constant, the process is known as frequency modulation.

(*iii*) When θ is varied inaccordance with modulating wave, while ω_c and E_c remain constant, the process is known as phase modulation.

30·1. AMPLITUDE MODULATION :

In amplitude modulation, the magnitude of carrier wave is varied inaccordance with the amplitude and frequency of the modulating voltage.

Let the signal voltage be represented as

$$e_m = E_m \cos \omega_m t, \qquad \ldots(1)$$

and unmodulated carrier be written as

$$e_c = E_c \cos(\omega_c t + \theta). \qquad \ldots(2)$$

The carrier frequency ω_c is much greater than ω_m.

The resulting modulated wave has the form

$$e = (E_c + K_a E_m \cos \omega_m t) \cos \omega_c t. \qquad \ldots(3)$$

The amplitude factor $(E_c + K_a E_m \cos \omega_m t)$ represents the sinusoidal variation of the amplitude of the wave. K_a is proportionality factor which determines the maximum variation in amplitude for a given signal voltage E_m. Phase θ is taken zero because it plays no part in this process.

The above expression can be written as

$$e = E_c \left(1 + \frac{K_a E_m}{E_c} \cos \omega_m t\right) \cos \omega_c t$$

$$= E_c (1 + m_a \cos \omega_m t) \cos \omega_c t$$

$$= E_c \cos \omega_c t + \frac{m_a E_c}{2} \cos (\omega_c + \omega_m) t + \frac{m_a E_c}{2} \cos (\omega_c - \omega_m) t, \qquad \ldots(4)$$

where $m_a = \dfrac{K_a E_m}{E_c}$ is called modulation index. The modulation index when multiplied by 100 gives percentage of modulation.

Side Bands : In equation (4) we have three terms :

(*i*) $E_c \cos \omega_c t$ represents original carrier voltage.

(ii) $\frac{m_a E_c}{2}\cos(\omega_c+\omega_m)t$

represents upper side band.

(iii) $\frac{m_a E_c}{2}\cos(\omega_c-\omega_m)t$

represents the lower side band.

The lower side band term with angular frequency $(\omega_c-\omega_m)$ and upper side band term with angular frequency $(\omega_c+\omega_m)$ are located on either side of the carrier at a frequency interval of ω_m in their frequency spectrum as shown in fig. 1.

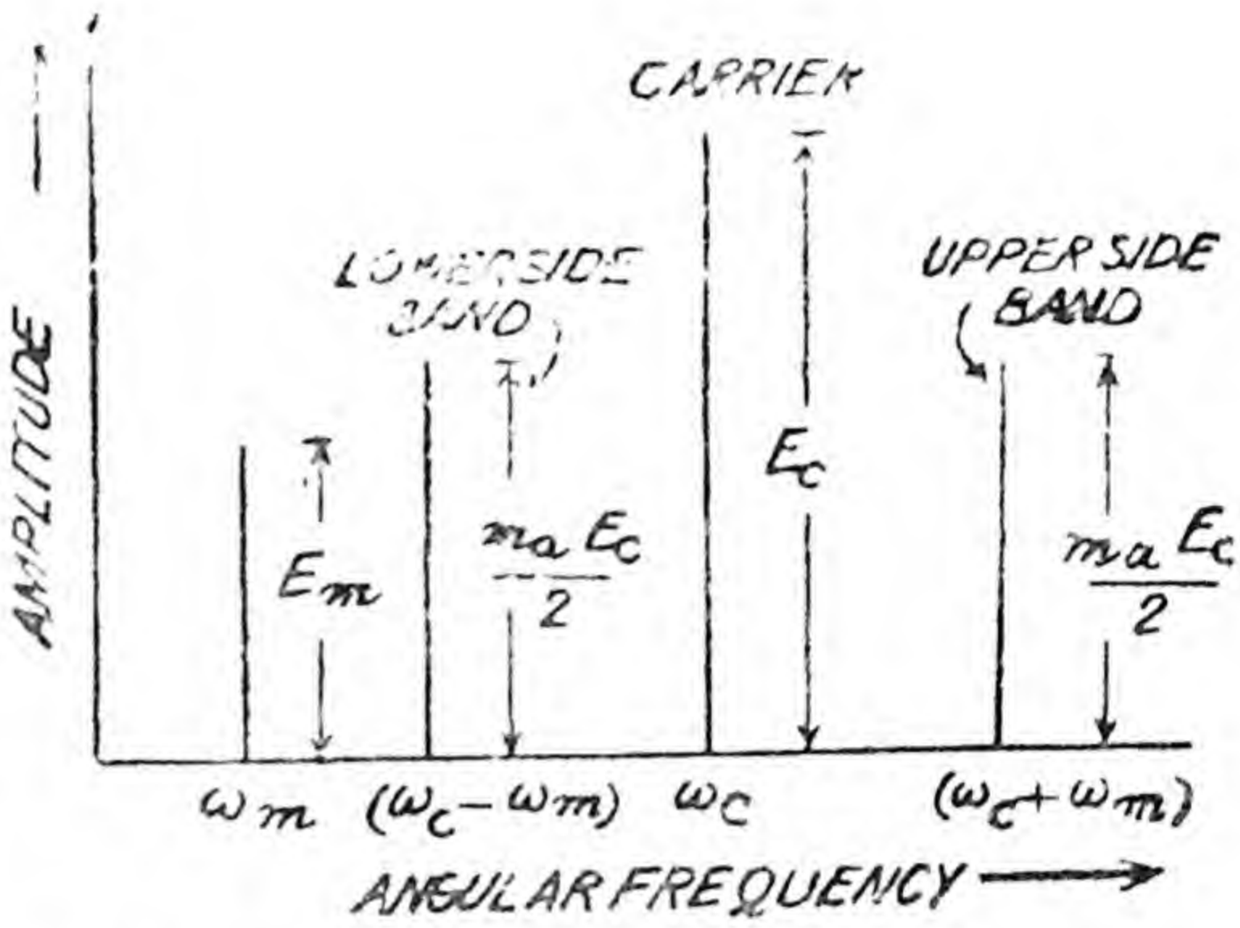


Fig. 1. Plot of frequency spec rum of amplitude modulated wave.

The magnitude of the amplitude of both the upper and lower side band is $\frac{m_a}{2}E_c$. If m_a be unity, then each side band term is half the carrier voltage in amplitude.

The carrier voltage component transmits no information, while each of the two side bands carries complete intelligence.

Waveforms : Figure (2) shows the waveforms of modulating signal and that of a modulated wave called the modulation envelope.

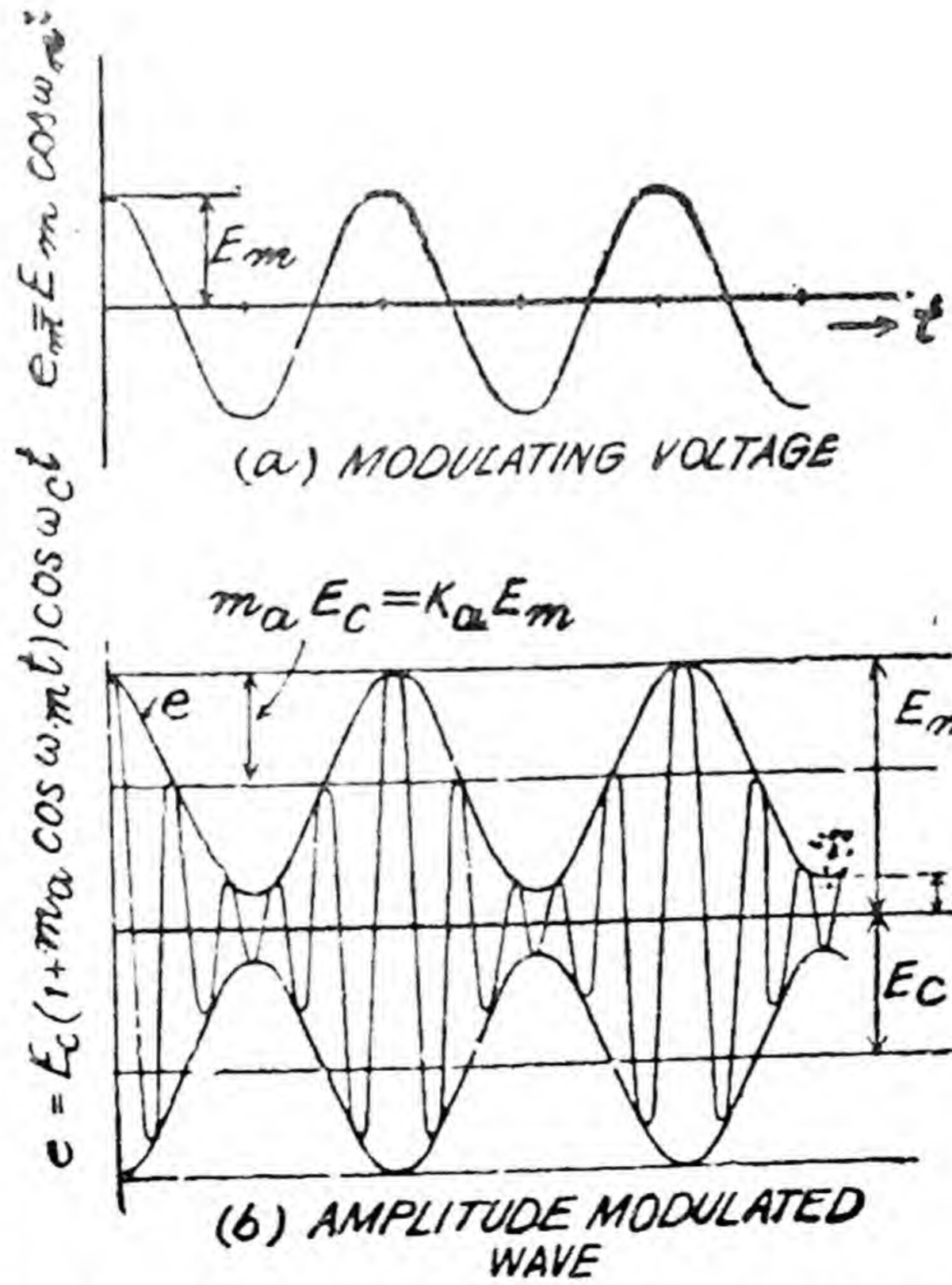


Fig. 2. Amplitude-modulated wave representation.

It is clear from fig. 2 (*b*) that $m_a E_c = E_{max} - E_c$.

$$\therefore\ m_a = \frac{E_{max} - E_c}{E_c} \qquad \ldots(5a)$$

Also $m_a E_c = E_c - E_{min.}$, $m_a = \frac{E_c - E_{min}}{E_c}$ $\ldots(5b)$

Equating both, $E_{max.} - E_c = E_c - E_{min.}$

$$\therefore\ 2E_c = E_{max.} + E_{min.}$$

Putting this value of E_c in 5 (*a*), we get

$$m_a = \frac{E_{max.} - E_{min.}}{E_{max.} + E_{min.}}. \qquad \ldots(5c)$$

And therefore percentage of modulation

$$= \frac{E_{max.} - E_{min.}}{E_{max.} + E_{min.}} \times 100.$$

Modulated waves with various degrees of modulation :

Modulated wave is expressed as

$$e = E_c\ (1 + m_a \cos \omega_m t) \cos \omega_c t,$$

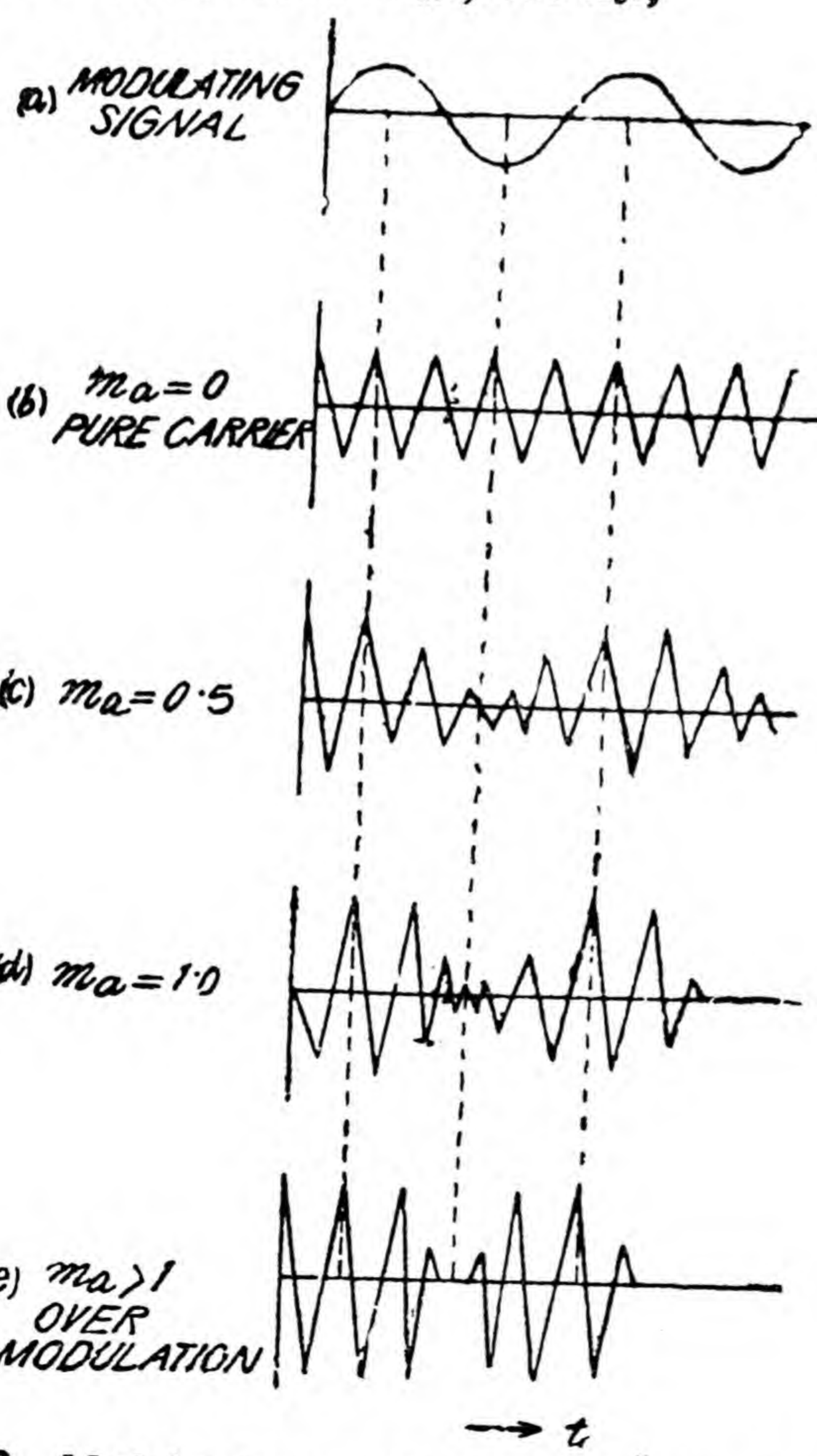


Fig. 3. Modulated wave with various degrees of modulation.

where $m_a = \frac{K_a E_m}{E_o}$ is modulation factor or modulation index, and its value ranges from zero to one. We shall now take the following different cases :

(*i*) When $m_a = 0$,

which means no amplitude modulation is present and the expression for e reduces to

$$e = E_c \cos \omega_c t,$$

representing pure carrier.

(*ii*) When $m_a = 0{\cdot}5$,

which means percentage of modulation is 50%. In other words, the extent to which carrier is being modulated is equal to half of its amplitude.

(*iii*) When $m_a = 1$,

which implies 100% modulation. In other words, the extent to which the carrier is modulated is equal to its amplitude. This means that at the negative peaks of the modulating signal, carrier amplitude is reduced to zero.

(*iv*) When $m_a > 1$,

represents the case of overmodulation. For such a case expression for e is no longer valid. At the negative peaks of modulating signal, no modulated wave exists and thus such a case results in a distortion of the intelligence that is being transmitted.

Energy distribution among the three components of amplitude modulated wave : The voltage components contained in the wave, represented by equation (4) have r.m.s. values equal to $\frac{E_c}{\sqrt{2}}$, $\frac{m_a E_o}{2\sqrt{2}}$ and $\frac{m_a E_c}{2\sqrt{2}}$ respectively. Therefore when such a modulated wave, having these three components, is impressed upon some resistance, R, then the total average power dissipated would be equal to

$$P_{av} = \frac{E_c^2}{2R} + \frac{m_a^2 E_c^2}{8R} + \frac{m^2 E_c^2}{8R}$$

$$= \frac{E_c^2}{2R}\left(1 + \frac{m_a^2}{2}\right)$$

where $\frac{E_o^2}{2R}$ represents the power dissipated by the carrier component alone. Obviously,

$$P_{av} > \frac{E_c^2}{2R},$$

$$> P_o,$$

that is, power content of modulated wave exceeds that of the corresponding unmodulated carrier. This additional power, supplied by modulating source, in modulated wave furnishes the energy contained in side bands of the modulated wave. It is the useful energy that carries intelligence and its relative magnitude is

proportional to m_a^2. The ratio of useful power contained in side bands to the total power contained in the modulated wave is

$$\frac{P_{side\ bands}}{P_{av}} = \frac{\frac{E_c^2}{2R}\cdot\frac{m_a^2}{2}}{\frac{E_c^2}{2R}\left(1+\frac{m_a^2}{2}\right)} = \frac{\frac{m_a^2}{2}}{1+\frac{m_a^2}{2}}.$$

Suppose modulation is 100% then $m_a=1$. For such a case,

$$\frac{P_{side\ bands}}{P_{av}} = \frac{\frac{1}{2}}{1+\frac{1}{2}} = \frac{1}{3},$$

that is, only one-third of the total power of modulated wave is contained in the two side bands. The rest two third power resides in carrier component which is of no use. If in the modulated wave carrier be suppressed then modulator would draw less power from supply source and the process of modulation will be quite economical

30·2. METHODS OF AMPLITUDE MODULATION :

There are two main categories of amplitude modulation methods :

(*i*) *Linear Modulation* : Relation between amplitude of modulating signal E_m and depth of modulation m_a is linear. In this type of modulation, the linear portion of the tube characteristic is used. The plate current, for such a portion, is related to grid potential by a simple relation $i_p=ae_g$.

(*ii*) *Square Law Modulation* : Relation between amplitude of modulating signal E_m and depth of modulation m_a is a square law. For this type of modulation, curved portion of the tube characteristic is used and the plate current is related to grid potential by the relation $i_p=a_1e_g+a_2e_g^2$.

30·3. LINEAR MODULATION METHODS :

The important linear modulation methods are :

(*i*) linear series plate modulation (plate modulated class *C* amplifier),

(*ii*) grid bias modulation,

(*iii*) suppressor grid modulation.

We shall discuss all these methods with necessary details.

30·3-1. Linear Series Plate Modulation : This method involves a modulated class *C* amplifier. Here a voltage corresponding to the signal is superimposed upon the d.c. plate supply voltage of an ordinary class *C* amplifier. In this way the total effective plate supply voltage of the amplifier is caused to vary inaccordance with the desired envelope of modulation.

The basic circuit of a plate modulated class *C* amplifier is shown in fig 4 (*a*).

The carrier frequency voltage is introduced into the grid circuit and the modulating voltage is connected in series with the direct plate supply voltage E_{bb}. Thus the total plate supply voltage comprises the sum of a direct and sinusoidal component. The resulting waveform in the output is illustrated in fig. 4 (*b*).

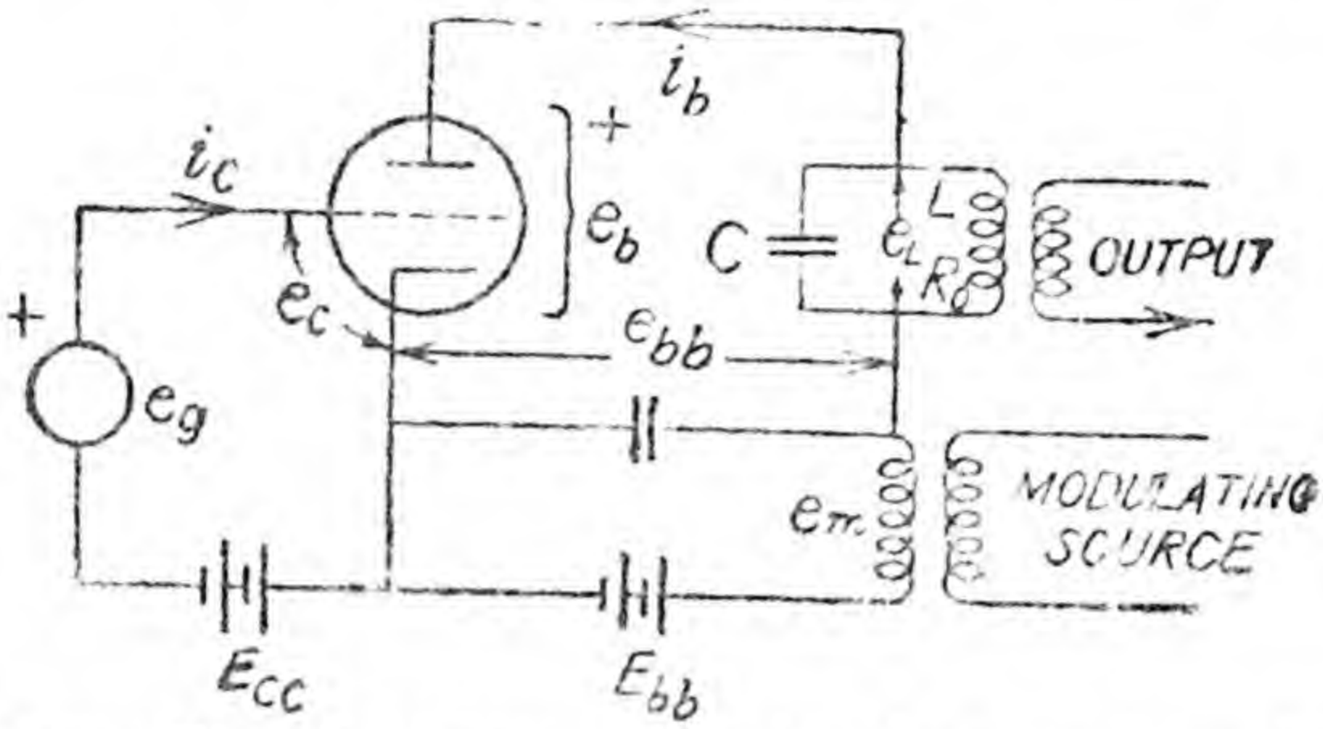


Fig. 4. (*a*) Circuit of plate modulated class C amplifier.

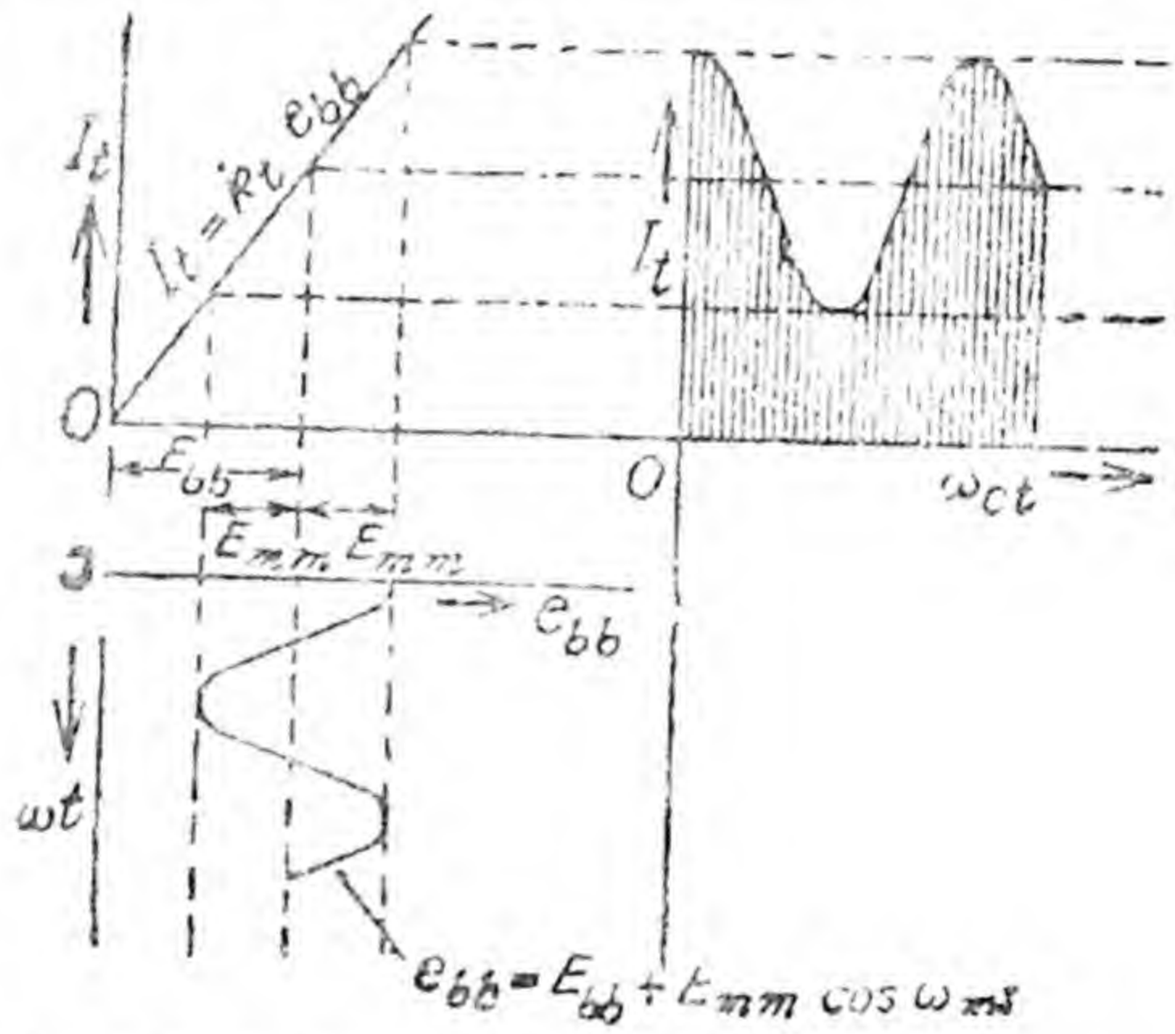


Fig 4 (*b*) Waveforms in the plate circuit of modulated class C amplifier.

Mathematical Analysis :

Let E_{bb} be the plate supply voltage,

I_t be the plate tank circuit current,

e_{bb} be the instantaneous plate supply voltage, and

E_{mm} be the amplitude of modulating signal.

It is assumed that $\omega_c \gg \omega_m$. We write for grid circuit potentials

$$e_g = E_{gc} \cos \omega_c t \text{ and } e_c = E_{gc} \cos \omega_c t + E_{cc}, \quad \ldots(1)$$

where E_{cc} is grid bias voltage.

Modulating voltage in the modulated amplifier plate circuit is given by

$$e_m = E_{mm} \cos \omega_m t. \quad \ldots(2)$$

Instantaneous plate supply voltage in modulated amplifier is given by

$$e_{bb} = E_{bb} + E_{mm} \cos \omega_m t$$

$$= E_{bb} (1 + m_a \cos \omega_m t), \text{ where } m_a = \frac{E_{mm}}{E_{bb}}. \quad \ldots(3)$$

Since r.m.s. tank current I_t and the d c. plate current I_b are related to the plate potentials by linear relation, we can express them by the following relations :

$$\left.\begin{aligned} I_t &= K_t e_{bb}, \\ I_b &= K_b e_{bb}, \\ I_{bb} &= K_b E_{bb}, \end{aligned}\right\} \qquad \ldots(4)$$

Combining eqs. (3) and (4), we get

$$I_t = K_t E_{bb} (1 + m_a \cos \omega_m t)$$

and

$$\begin{aligned} I_b &= K_b E_{bb} (1 + m_a \cos \omega_m t) \\ &= I_{bb} (1 + m_a \cos \omega_m t). \end{aligned}$$

The instantaneous tank current is given by

$$\begin{aligned} i_t &= \sqrt{2} I_t \sin \omega_c t \\ &= \sqrt{2} K_t E_{bb} (1 + m_a \cos \omega_m t) \sin \omega_c t. \end{aligned} \qquad \ldots(5)$$

R.M.S. value of output potential across the tank circuit is given by

$$E_L = jXI_t = XI_t \angle 90°,$$

where X is reactance of inductance or condenser in the tuned circuit.

Hence the instantaneous output voltage is given by

$$\begin{aligned} e_L &= X\sqrt{2}. I_t \sin (\omega_c t + 90) \\ &= \sqrt{2}. XK_t E_{bb} (1 + m_a \cos \omega_m t) \cos \omega_c t. \end{aligned} \qquad \ldots(6)$$

R.M.S. value of plate current at resonant frequency is given by

$$*I_p = \frac{E_L}{E_0} = j \frac{XI_t}{R_0} = j \frac{I_t}{Q} = \frac{I_t}{Q} \angle 90°.$$

Therefore the instantaneous plate current is given by

$$\begin{aligned} i_p &= \sqrt{2} I_p \sin \omega_c t \\ &= \sqrt{2}. \frac{I_t}{Q} \sin (\omega_c t + 90) \\ &= \frac{\sqrt{2}}{Q}. K_t E_{bb} (1 + m_a \cos \omega_m t) \cos \omega_c t. \end{aligned} \qquad \ldots(7)$$

Average power supplied by the d.c. plate power source is given by

$$P_{bb} = \frac{1}{T_m} \int_0^{T_m} E_{bb} I_b \, dt,$$

where T_m is the periodic time of the modulating cycle

Or

$$\begin{aligned} P_{bb} &= \frac{1}{T_m} \int_0^{T_m} E_{bb}. I_{bb} (1 + m_a \cos \omega_m t). dt \\ &= E_{bb} \; I_{bb} = K_b E_{bb}^2. \end{aligned} \qquad \ldots(8)$$

Average power supplied by the modulating source is

$$P_m = \frac{1}{T_m} \int_0^{T_m} e_m \, I_b \, dt$$

---

*At resonant frequency, $R_0 = \frac{L}{RC}$.

Therefore $\frac{X}{R_0} = \frac{\omega_0 L}{L/CR} = \omega_0 CR = \frac{1}{Q}$.

$$= \frac{1}{T_m} \int_0^{T_m} E_{mm} \cos \omega_m t . I_{bb} (1 + m_a \cos \omega_m t)\, dt$$

$$= E_{mm} . I_{bb} \frac{m_a}{2} = E_{bb} . I_{bb} \frac{m_a^2}{2} = P_{bb} \frac{m_a^2}{2}. \quad \ldots(9)$$

Then total average input will bc

$$P_i = P_{bb} + P_m$$

$$= P_{bb} + P_{bb} \frac{m_a^2}{2} = P_{bb} \left(1 + \frac{m_a^2}{2} \right) \quad \ldots(10)$$

The a.c. output power across the tank circuit is given by

$$P_0 = \frac{1}{T_m} \int_0^{T_m} e_L \cdot i_p \, di = \frac{1}{T_m} \int_0^{T_m} R_0 i_p^2 \, dt$$

$$= \frac{1}{T_m} \int_0^{T_m} R_0 . 2 \left(\frac{K_t E_{bb}}{Q} \right)^2 (1 + m_a \cos \omega_m t)^2 \cos^2 \omega_c t . dt$$

$$= \left(\frac{K_t E_{bb}}{Q} \right)^2 \frac{R_0}{T_m} \int_0^{T_m} (1 + m_a^2 \cos^2 \omega_m t + 2 m_a \cos \omega_m t)\, (1 + \cos 2\omega_c t)\, dt.$$

Since $\omega_c \gg \omega_m$, therefore average value of $\cos 2\omega_c t$ and $2m_a \cos \omega_m t$ will be zero.

So

$$P_0 = \frac{R_0 K_t^2}{Q^2} E_{bb}^2 . \left(1 + \frac{m_a^2}{2} \right)$$

$$= \frac{R_0 K_t^2}{Q^2 K_b} P_{bb} . \left(1 + \frac{m_a^2}{2} \right)$$

$$= \eta \, P_{bb} \left(1 + \frac{m_a^2}{2} \right),$$

where η is defined below. The unmodulated carrier power is obtained by putting $m_a = 0$ so that $P_c = \eta \, P_{bb}$.

Plate circuit efficiency $\eta = \dfrac{P_0}{P_i} = \dfrac{P_0}{P_{bb} + P_m}$

$$= \frac{R_0 \frac{K_t^2}{Q^2} \frac{1}{K_b} . P_{bb} \left(1 + \frac{m_a^2}{2} \right)}{P_{bb} \left(1 + \frac{m_a^2}{2} \right)} \quad \ldots(11)$$

$$= \frac{R_0 K_t^2}{Q^2 K_b}. \quad \ldots(12)$$

Plate dissipation $P_p = P_i - P_0 = P_i (1 - \eta)$

$$= (P_{bb} + P_m) (1 - \eta)$$

$$= P_{bb} \left(1 + \frac{m_a^2}{2} \right) (1 - \eta). \quad \ldots(13)$$

Equation (13) gives the plate dissipation of modulated amplifier tube. With 100% modulation plate dissipation will increase by a factor 3/2 from its unmodulated value ($m_a = 0$). Therefore while designing plate modulated class C amplifier, only 2/3 of the rated plate dissipation is taken under no modulation condition.

Plate dissipation of modulating amplifier tube P_{pm}, can also be related to P_m. If η_m be the plate circuit efficiency of modulating amplifier, then d.c. power input to modulating amplifier

$$=\frac{P_m}{\eta_m},$$

so that plate dissipation of modulating amplifier will be

$$P_{pm}=\text{d.c. power input—a.c. power output}$$

$$=\frac{P_m}{\eta_m}-P_m$$

$$=P_m\left(\frac{1}{\eta_m}-1\right)$$

$$=P_{bb}\left(\frac{m_a^2}{2}\right)\left(\frac{1}{\eta_m}-1\right). \qquad ...(14)$$

Advantages : (*i*) Distortion in modulated wave is small.
(*ii*) Plate circuit efficiency is high.
(*iii*) Adjustments are easy.

The main drawback of this method is that a large amount of modulating power is required.

Example 1. *In a plate modulated class C amplifier the plate dissipation of r. f. amplifier tube with modulation is 500 watts. If the amplifier efficiency is 60% and modulation index is 50%, calculate :*

(*a*) *d.c. power supplied by the supply source,*
(*b*) *power modulated by the modulation amplifier,*
(*c*) *modulated carrier power output.*

(*a*) From eq. (13), d.c. power supplied by the supply source, P_{bb}, can be calculated on putting

$$P_p=500 \text{ watts.}$$
$$m_a=0{\cdot}5,$$
$$\eta=0{\cdot}6.$$

Therefore

$$P_p=P_{bb}\left(1+\frac{m_a^2}{2}\right)(1-\eta),$$

$$500=P_{bb}\left(1+\frac{0{\cdot}25}{2}\right)(1-0{\cdot}6),$$

$$=P_{bb}\ (1{\cdot}125)\ (0{\cdot}4)$$

$$P_{bb}=\frac{500}{0{\cdot}45}=1111{\cdot}11 \text{ watts.}$$

(*b*) Power supplied by the modulating amplifier

$$P_m=P_{bb}\ \frac{m_a^2}{2}$$

$$=1111{\cdot}11\times\frac{0{\cdot}25}{2}$$

$$=1111\cdot11\times0\cdot125$$
$$=138\cdot89 \text{ watts.}$$

(*c*) Modulated carrier power output

$$P_0=\eta\ P_{bb}\left(1+\frac{m_a^2}{2}\right)$$
$$=0\ 6\times1111\cdot11\times1\cdot125$$
$$=749\cdot99 \text{ watts.}$$

Example 2. *The plate dissipation of a plate modulated class C amplifier is 500 watts with modulation index 0·6. The efficiency of r.f. amplifier is 50% and of the modulator valve (modulating amplifier) is 25%. Find*

(*a*) *carrier power and valve dissipation at 100% modulation.*

(*b*) *the modulating signal power output, plate dissipation and the rating of modulator valve to give 100% modulation.*

(*c*) *the overall efficiency at 60% modulation.*

It is given that $P_p=500W$,

$m_a=0\cdot6$,

$\eta=50\%=2\cdot5$,

$\eta_m=25\%=0\cdot25$.

(*a*) We are to find P_{bb} and P_o, we know that

$$P_{bb}=\frac{P_p}{\left(1+\frac{m_a^2}{2}\right)(1-\eta)}$$
$$=\frac{500}{1\cdot15\times0\cdot5}=847\cdot4\ W.$$

Carrier power output can be obtained from modulated power output,

$$P_o=\eta\ P_{bb}\left(1+\frac{m_a^2}{2}\right)$$

by putting $m_a=0$, so that

$$P_o=\eta\ P_{bb}$$
$$=0\cdot5\times847\cdot4$$
$$=423\cdot7 \text{ watts.}$$

Valve dissipation at 100% modulation

$$P_p=P_{bb}\left(1+\frac{m_a^2}{2}\right)(1-\eta)$$
$$=847\cdot4\times\tfrac{3}{2}\times0\cdot5=635\cdot65 \text{ watts.}$$

(*b*) From equation (14), plate dissipation of modulator valve, P_{pm}, (*i.e.* of modulating amplifier) is given by

$$P_{pm}=P_{bb}\left(\frac{m_a^2}{2}\right)\left(\frac{1}{\eta_m}-1\right).$$

For 100% modulation $m_a=1$, given that $\eta_m=0\ 25$.

$$P_{pm}=423\cdot7\ (\tfrac{1}{2})\left(\frac{1}{0\cdot2}-1\right)=423\cdot7\times\tfrac{1}{2}\times3$$

$$=635\cdot55 \text{ watts.}$$

When a.c. output modulating signal power is zero (*i.e.* under no modulation) then total d.c. input power to modulator appears as plate dissipation. This is the maximum value of plate dissipation and termed as rating of modulator valve. Thus rating of modulator valve

$(P_i)_m$=power input to modulator valve (modulating amplifier valve)

$$=\frac{P_m}{\eta_m}=\frac{P_{bb}\cdot\frac{m_a^2}{2}}{\eta_m}$$

$$=\frac{847\cdot4\times\frac{1}{2}}{0\cdot25}$$

$$=1694\cdot8 \text{ watts.}$$

(*c*) Over all efficiency

$$=\frac{\text{Total modulated power output, } P_0}{\text{d.c. power input to the modulated amplifier } (P_{bb}) \text{ and modulating amplifier } (p_i)_m.}\times100$$

We know that

$$P_0=\eta\ P_{bb}\left(1+\frac{m_a^2}{2}\right)$$

$$=0\cdot5\times847\cdot4\times\left(1+\frac{0\cdot36}{2}\right),$$

at 60% modulation.

Or $$P_0=499\cdot96 \text{ watts.}$$

Total input power to both the amplifiers will remain same irrespective of modulation index therefore

total input power$=P_{bb}+(P_i)_m$

$$=847\cdot4+1694\cdot8$$

$$=2542\cdot2 \text{ watts.}$$

Therefore over all efficiency

$$=\frac{499\cdot96}{2542\cdot2}\times100$$

$$=19\cdot6\%.$$

30·3-2. GRID BIAS MODULATION :

In grid bias modulation, modulating voltage corresponding to the signal that is to be modulated on the radio frequency wave, is superimposed on a fixed bias in grid circuit of a class *C* amplifier.

The basic circuit of grid bias modulated class *C* amplifier is shown in figure 5.

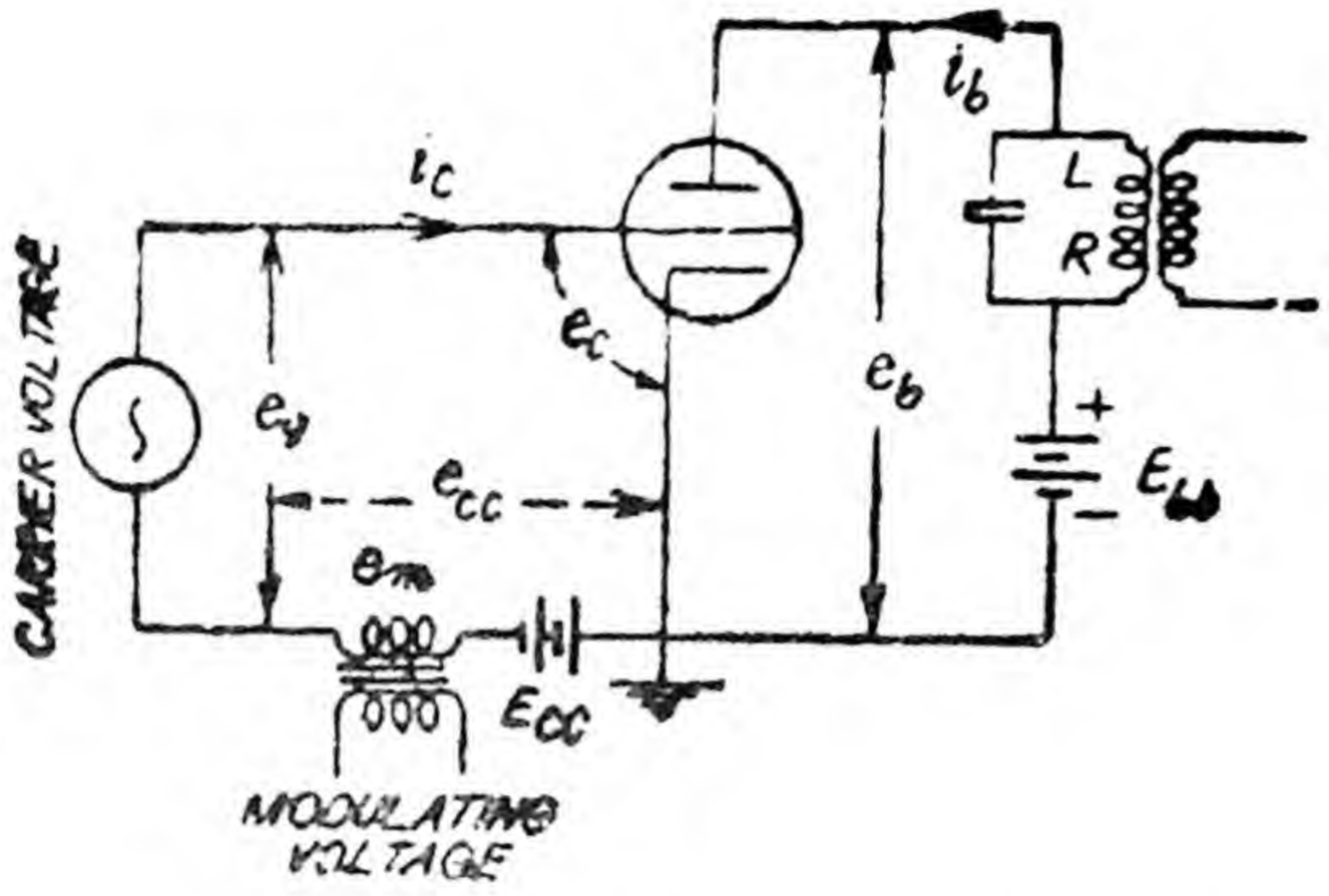


Fig. 5. Basic circuit of grid modulated class *C* amplifier.

The modulating voltage e_m will cause small periodic variation in the steady grid bias E_{cc}. The steady grid bias E_{cc} and the modulating vol age e_m, then constitute the variable grid bias e_{co}.

The plot of r.m.s tank current I_t against the grid bias e_c is a straight line. Then on including the modulating voltage e_m in series with fixed bias E_{co} similar variations of tank current I_t will take place. The amplitude of instantaneous tank current, I_t then varies at modulation frequency.

Thus it is apparent that by properly choosing the bias and superimposing upon it a modulating voltage, it is possible to obtain a completely modulated wave.

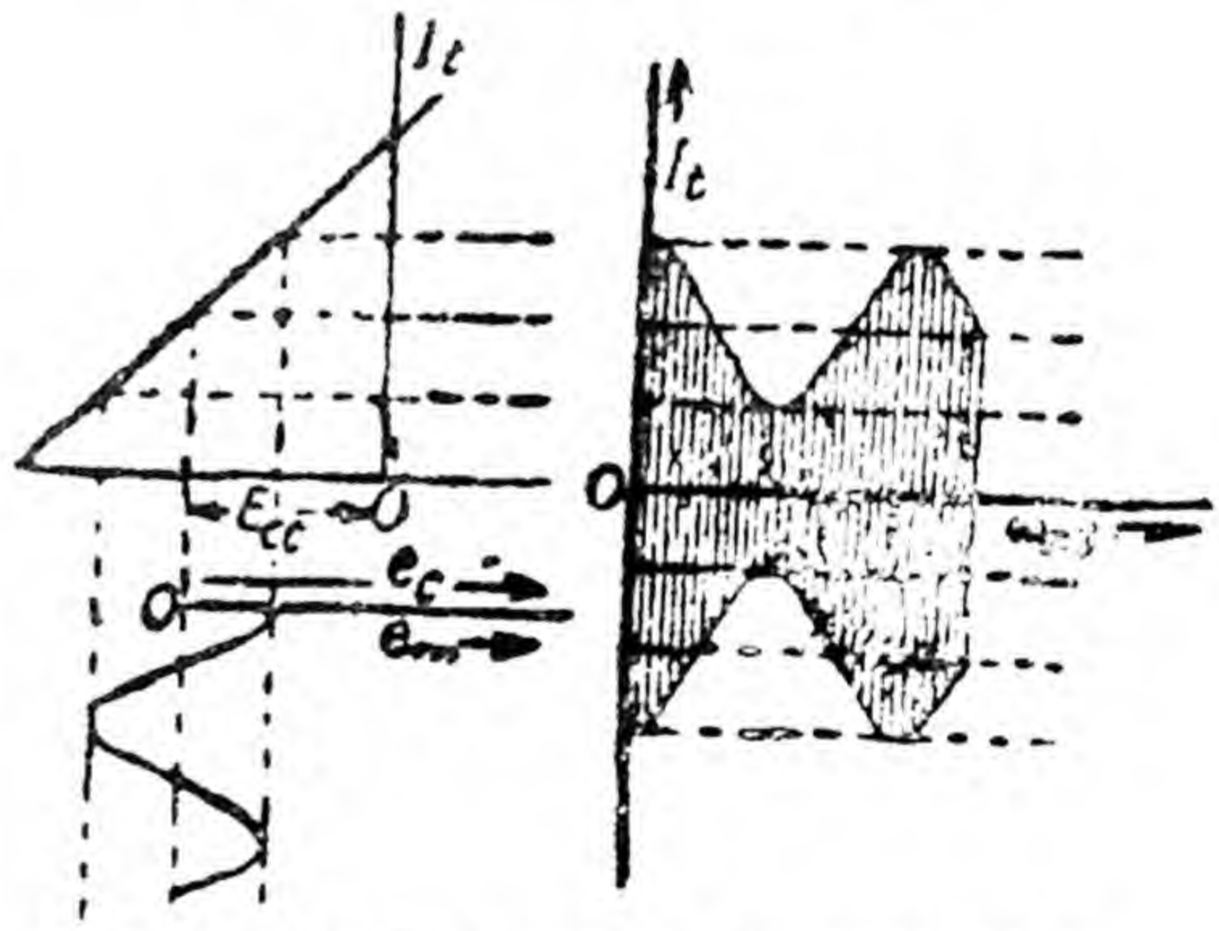


Fig. 6. Variation of tank current in grid bias modulation.

Mathematical Analysis : Similar to plate modulation analysis, we can also proceed here to calculate d.c. input power, a.c. output power, plate dissipation, etc.

Let the carrier voltage be represented by

$$e_g = E_{gm} \sin \omega_c t, \qquad \ldots(1)$$

and the modulating voltage

$$e_m = E_{mm} \cos \omega_m t. \qquad \ldots(2)$$

Thus the total varying grid bias voltage is given by

$$e_{cc} = e_m + E_{cc}$$
$$= E_{cc} + E_{mm} \cos \omega_m t. \qquad \ldots(3)$$

The total grid to cathode voltage can now be represented as

$$e_c = e_b + e_{cc}$$
$$= e_{gc} \sin \omega_c t + E_{mm} \cos \omega_m t + E_{co}. \quad ...(4)$$

If I_t represents the r.m.s. value of tank current for linear modulation characteristic, we can write

$$I_t = K_t (e_{co} - E_{co}), \quad ...(5)$$

where E_{co} is the grid to cathode voltage for zero tank current.

The instantaneous tank current which varies at frequency ω_o can be written as

$$i_t = \sqrt{2} I_t \sin \omega_o t$$
$$= \sqrt{2} K_t (e_{co} - E_{co}) \sin \omega_o t$$
$$= \sqrt{2} K_t [E_{mm} \cos \omega_m t + E_{co} - E_{co}] \sin \omega_c t \quad ...(6)$$
$$= \sqrt{2}\ K_t (E_{cc} - E_{co}) \left[1 + \frac{E_{mm} \cos \omega_m t}{E_{oo} - E_{oo}} \right] \sin \omega_o t$$
$$= \sqrt{2}.I_t [1 + m_a \cos \omega_m t] \sin \omega_o t, \quad ...(7)$$

where m_a is the modulation index and is given by

$$m_a = \frac{E_{mm}}{E_{cc} - E_{co}}. \quad ...(8)$$

Since the grid bias e_{co} consists of two components—one E_{co}, the steady component and other e_m the variable component, the d.c. plate current denoted by I_b, inspite of being constant, will vary inaccordance with the variation of variable grid bias e_{co}. I_b can be represented as

$$I_b = K_b (e_{cc} - E_{co}). \quad ...(9)$$

The steady value of d.c. plate current is then given by

$$I_{bb} = K_b (E_{co} - E_{co}). \quad ...(10)$$

In the equation (9) putting the value of e_{cc},

$$I_b = K_b (E_{co} + E_{mm} \cos \omega_m t - E_{co})$$
$$= K_b (E_{co} - E_{co}) \left[1 + \frac{E_{mm} \cos \omega_m t}{E_{cc} - E_{co}} \right]$$
$$= I_{bb} (1 + m_a \cos \omega_m t). \quad ...(11)$$

Tank circuit voltage, being the a.c. output voltage, is given by

$$E_0 = jXI_t. \quad ...(12)$$

where X represents the capacitive or inductive reactance at resonant frequency ω_0.

Putting value of I_t in equation (12), we get

$$E_0 = jXK_t (E_{co} - E_{co}) [1 + m_a \cos \omega_m t].$$

Thus the corresponding instantaneous value of output voltage is

$$e_0 = \sqrt{2} XK_t (E_{co} - E_{co}) [1 + m_a \cos \omega_m t] \sin (\omega_c t + 90°)$$
$$= \sqrt{2} XK_t (E_{cc} - E_{co}) [1 + m_a \cos \omega_m t] \cos \omega_c t. \quad ...(13)$$

D.C. Input Power : If T_m is the periodic time of modulating voltage, the average power input by the d.c. plate source is given by

$$P_{bb}=\frac{1}{T_m}\int_0^{T_m} E_{bb}\, I_b\, dt.$$

$$P_{bb}=\frac{1}{T_m}\int_0^{T_m} E_{bb}I_{bb}\,(1+m_a\cos\omega_m t)\, dt.$$

$$=E_{bb}I_{bb}. \qquad \ldots(14)$$

Thus d,c. input power from plate supply source remains constant for all values of modulation index.

A. C. Output Power :

$$P_0=\frac{1}{T_m}\int_0^{T_m} e_0\, i_p\, dt=\frac{1}{T_m}\int_0^{T_m} e_0\,\frac{e_0}{R_0}\, dt,$$

where R_0 is the impedance of tank circuit at resonance at the fundamental frequency and i_p is the a.c. component of plate current at the fundamental frequency. Thus

$$P_0=\frac{1}{T_m}\int_0^{T_m}\frac{e_0^2}{R_0}\, dt.$$

Putting the value of e_0,

$$=\frac{1}{T_m}\int_0^{T_m}\frac{1}{R_0}.\ 2X^2K_t^2\,(E_{cc}-E_{c0})^2\,(1+m_a\cos\omega_m t)^2\cos^2\omega_0 t$$

$$=\frac{1}{T_m}\int_0^{T_m}\frac{R_0K_t^2}{Q^2}\,(E_{cc}-E_{cc})^2\,(1+m_a^2\cos^2\omega_m t$$

$$+2m_a\cos\omega_m t)\,(1+\cos 2\omega_0 t),$$

where

$$\frac{X}{R_0}=\frac{1}{Q}.$$

Since $\omega_c \gg \omega_m$, the contribution of $\cos 2\omega_c t$ is zero. The average value of $2m_a\cos\omega_m t$ is zero. Hence

$$P_0=\frac{R_0K_t^2}{Q^2}\,(E_{cc}-E_{c0})^2\left(1+\frac{m_a^2}{2}\right). \qquad \ldots(15)$$

Plate Circuit Efficiency :

Plate circuit efficiency is given by

$$\eta=\frac{P_0}{P_t}=\frac{\dfrac{R_0K_t^2}{Q^2}\,(E_{cc}-E_{c0})^2\left(1+\dfrac{m_a^2}{2}\right)}{E_{bb}K_b\,(E_{cc}-E_{c0})}$$

$$=\frac{R_0K_t^2}{Q^2K_bE_{bb}}\,(E_{cc}-E_{c0})\left(1+\frac{m_a^2}{2}\right). \qquad \ldots(16)$$

Thus it is quite obvious that the plate circuit efficiency increases with modulation index. For a constant value of modulation index, it would be maximum by making ($E_{cc}=E_{c0}$) as large as possible.

Plate Dissipation :

Plate dissipation is given by

$$P_p = P_{bb} - P_0$$

$$= E_{bb}K_b(E_{cc} - E_{c0}) - \frac{R_0K_t^2}{Q^2}(E_{cc} - E_{c0})^2\left(1 + \frac{m_a^2}{2}\right). \quad ...(17)$$

Putting for E_{bb} from equation (16), we have

$$P_p = \frac{R_0K_t^2}{\eta Q^2}(E_{cc} - E_{c0})^2\left(1 + \frac{m_a^2}{2}\right) - \frac{R_0K_t^2}{Q^2}(E_{cc} - E_{c0})^2\left(1 + \frac{m_a^2}{2}\right). \quad ...(18)$$

The expression shows that plate dissipation decreases as the modulation index increases.

For unmodulated efficiency we put $m_a = 0$, so that

$$\eta_{unmod.} = \frac{R_0K_t^2}{Q^2K_bE_{bb}}(E_{cc} - E_{c0}). \quad ...(19)$$

Therefore, $$\eta = \eta_{unmod.}\left(1 + \frac{m_a^2}{2}\right). \quad ...(20)$$

Similarly, the modulated power is

$$P_0 = \frac{R_0K_t^2}{Q^2}(E_{cc} - E_{c0})^2\left(1 + \frac{m_a^2}{2}\right)$$

and for unmodulated power we put $m_a = 0$, so that

$$P_{unmod.} = \frac{R_0K_t^2}{Q^2}(E_{cc} - E_{c0}). \quad ...(21)$$

Therefore, $$P_0 = P_{unmod.}\left(1 + \frac{m_a^2}{2}\right). \quad ...(22)$$

In view of equation (21), we can write equation (18) for plate dissipation. We find that

$$P_p = \frac{P_{unmod.}}{\eta}\left(1 + \frac{m_a^2}{2}\right) - P_{unmod.}\left(1 + \frac{m_a^2}{2}\right).$$

Obviously the $P_{unmod.}$ is carrier power and can be denoted by P_c, so that

$$P_p = \frac{P_c}{\eta}\left(1 + \frac{m_a^2}{2}\right) - P_c\left(1 + \frac{m_a^2}{2}\right)$$

$$= P_c\left(1 + \frac{m_a^2}{2}\right)\left(\frac{1}{\eta} - 1\right). \quad ...(23)$$

In grid modulated class C amplifier plate dissipation is maximum when $m_a = 0$, so that

$$(P_p)_{max} = P_c\left(\frac{1}{\eta_{unmod.}} - 1\right). \quad ...(24)$$

This is also termed as the rating of the valve.

Advantages :

(i) Grid modulated class *C* amplifier has the advantage of requiring very little modulating signal power for complete modulation as compared to plate modulated class *C* amplifier. The reason being that only a small amount of energy is drawn from the modulatin g source to compensate for proper losses in the grid circuit. This greatly reduces the size and cost of modulating system.

(ii) Due to the low power requirement of modulating signal, this type of modulation is especially suitable for television transmitters.

Disadvantages :

(i) It has much lower plate circuit efficiency and it also varies throughout the modulation cycle. At zero modulation index, plate circuit efficiency is about 34% (at the positive peak of modulation cycle) which is only about one-half of the plate circuit efficiency of class *C* amplifier. At 100% modulation plate circuit efficiency is about 5 i%.

(ii) It is somewhat more typical in adjustment than plate modulated class *C* amplifier because circuit adjustments are more sensitive to plate supply voltage and load impedan ce.

Also due to wide variations in grid voltage during operation, the modulation tends to be somewhat non-linear. It has higher amplitude distortion.

30·3-3. SUPPRESSOR GRID MODULATION :

Amplitude modulation can be obtained by applying to the suppressor grid the modulating voltage superimposed upon a suitable negative bias in a pentode class *C* amplifier.

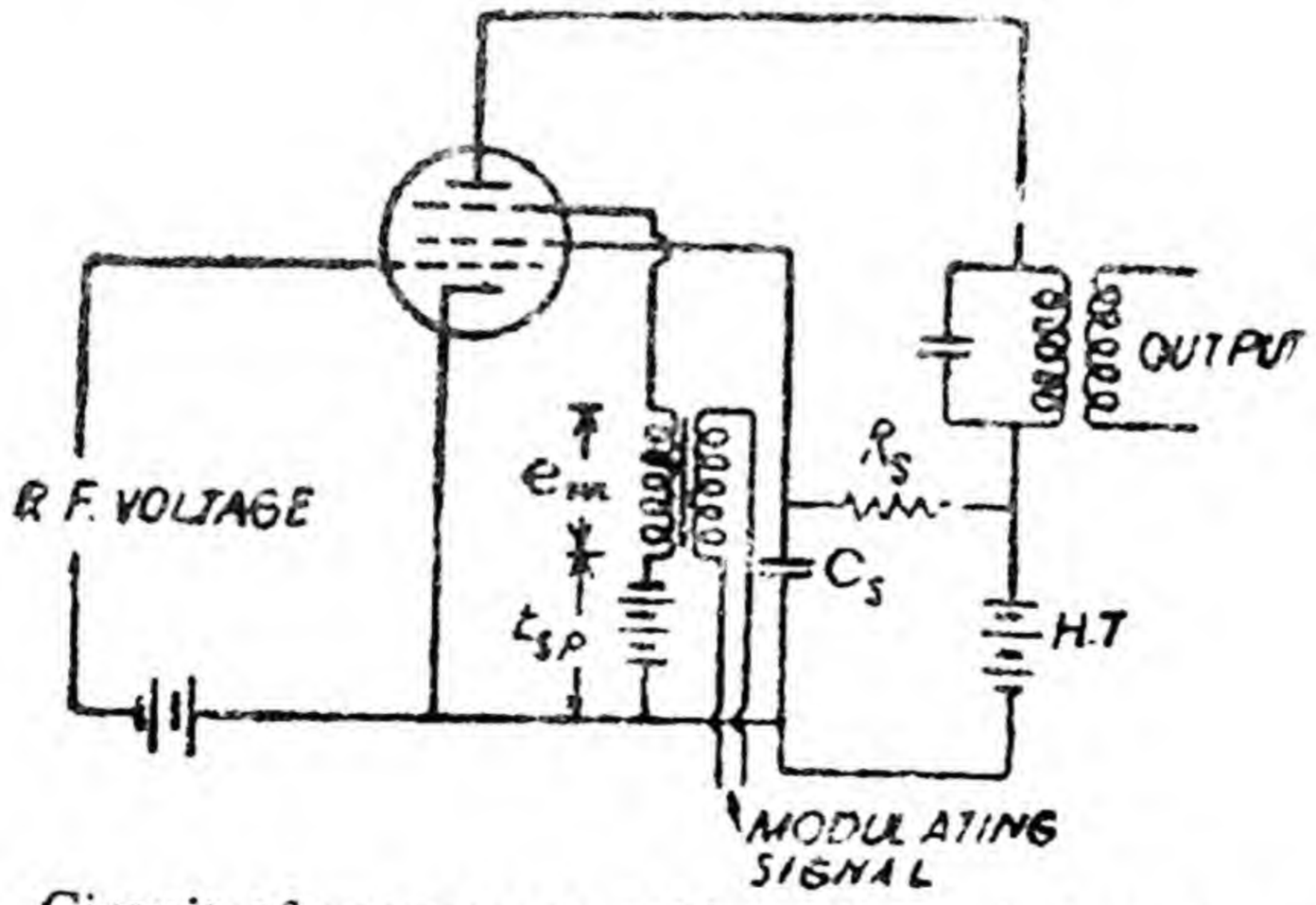


Fig. 7 (*a*) Circuit of suppressor grid modulated class *C* amplifier.

The suppressor grid is biased sufficiently negative. The modulation voltage and this bias are adjusted so that the modulation

characteristic lies entirely in the region of negative suppressor grid voltage. The power required in modulation will be almost zero because the suppressor grid current is zero.

The circuit of suppressor grid modulated class C amplifier is shown in fig. 7 (a). The resistance R_S limits the screen losses by simultaneously lowering the screen voltage and screen current. The capacitor C_S should be sufficiently large for bypassing carrier frequency.

The relation in output voltage and suppressor grid voltage is depicted in the diagram 7 (*b*).

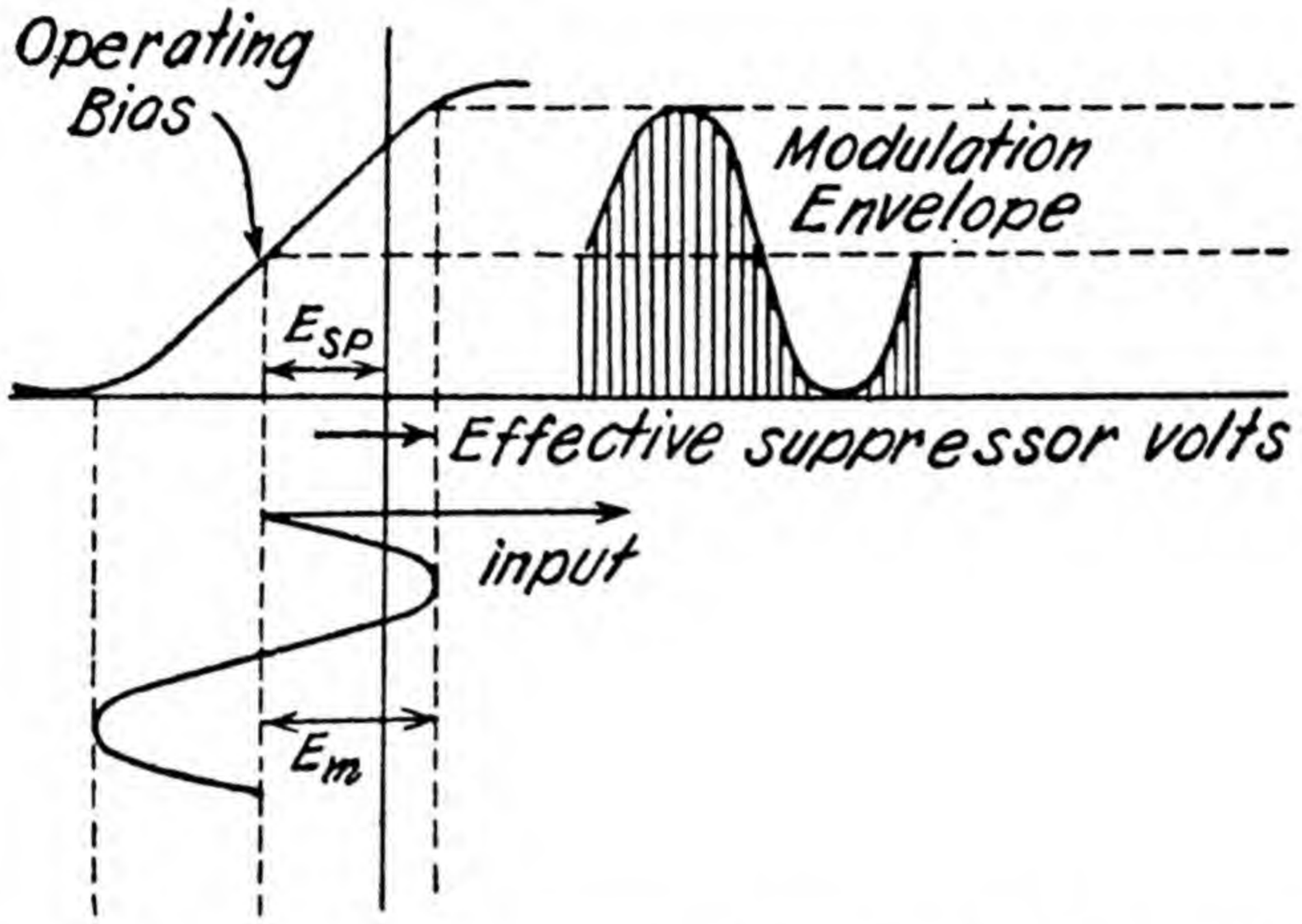


Fig. 7. (*b*) **Relation between amplifier output and suppressor grid voltage.**

Advantages :

(*i*) Low modulation power is required since the suppressor grid is always negative.
(*ii*) Modulation is sufficiently linear over the effective range in use.
(*iii*) Neutralisation is not required.
(*iv*) Adjustment of modulated amplifier is simple.

30 3-4. TRANSISTOR MODULATORS :

Transistor *RF* power amplifiers can be employed for the purpose of modulation in a way similar to vacuum tube circuits. Modulating signal can be fed to any of the three elements *viz.* collector, base and emitter. The corresponding circuits are called collector modulation, base modulation and emitter modulation circuits.

Collector Modulation : Refer to fig. 8. Modulating signal has been applied to the collector, in series with d.c. collector supply voltage. Transformer T, matches the output stage of modulating amplifier with the collector load of class C, r.f. power amplifier Q. Capacitor, C, offers low reactance to radio frequencies but a high reactance to modulating signal frequencies. Thus it grounds the carrier and serves to keep r.f. out of modulating circuit and

out of power supply. To obtain 100% modulation, maximum value of modulating voltage, E_{mm}, must equal power supply voltage, V_{CC}. Under these conditions, the r.f. output of the modula-

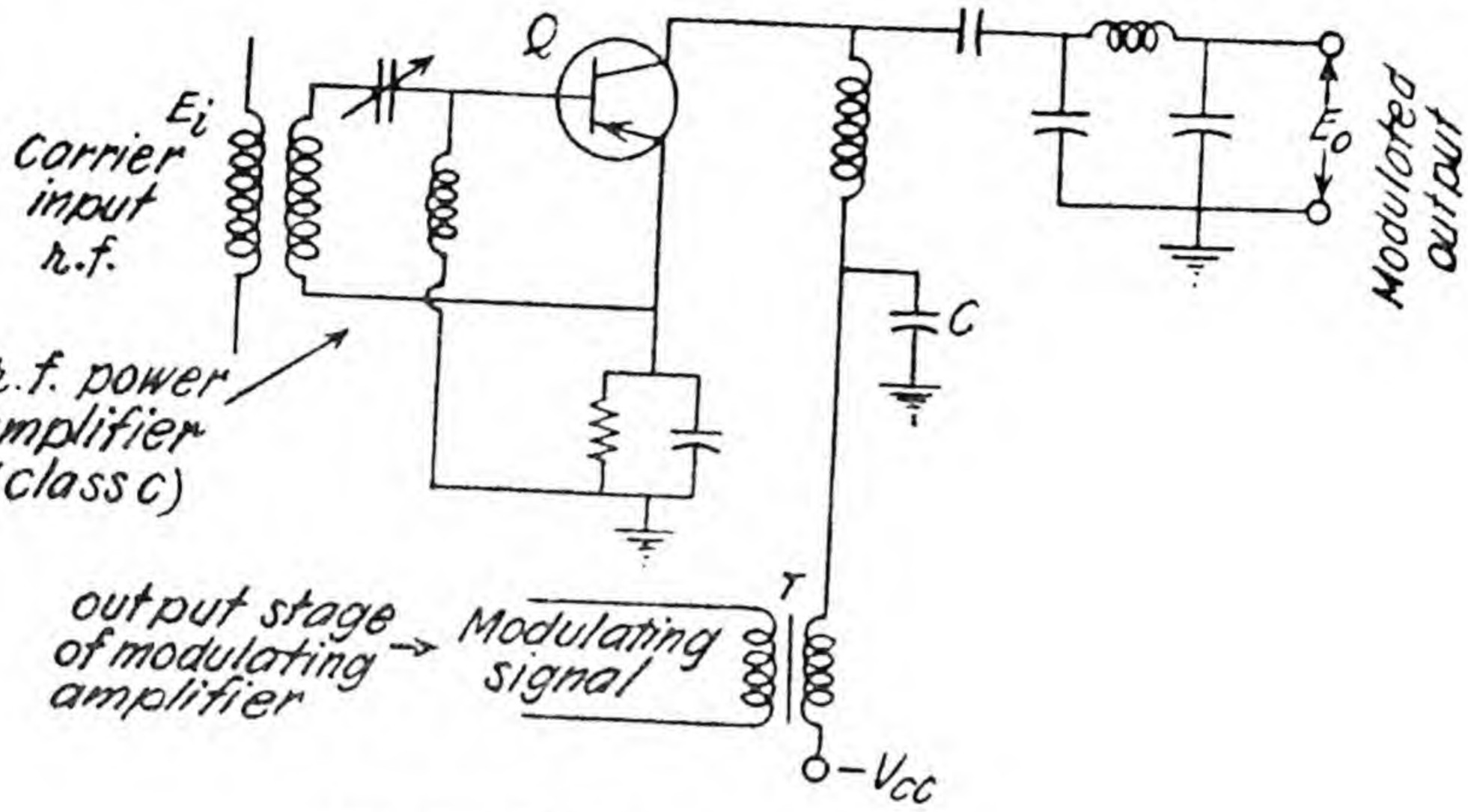


Fig. 8. Collector modulator.

ted amplifier is equal to zero at the negative peak of the modulating signal. Then we have

$$E_{mm} = V_{CC}$$

Since the average collector current of the modulated amplifier is reduced to zero from I_C during the negative half cycle of the modulated signal, the maximum value of modulating signal current is

$$I_{mm} = I_C$$

Thus modulator power output will be

$$P_0 = \frac{E_{mm} I_{mm}}{2} = \frac{V_{CC} I_C}{2}$$

$= \frac{1}{2} \times$ the power supplied by the power supply.

The effective load on the modulator is given by

$$R_L = \frac{E_{mm}}{I_{mm}} = \frac{V_{CC}}{I_C}$$

Collector modulation has the advantages over base modulation (to be described) of *better linearity, higher collector efficiency,* and *higher power output per transistor,* but also a disadvantage of requiring more modulating power.

Base Modulation : It is very similar to grid bias modulation of vacuum tube circuit. Refer to fig. 9. The modulating signal is fed into the base circuit of the transistor. Its salient features are :

(i) it has low modulating power requirement compared to collector modulation,

(ii) it has poorer linearity and is more difficult to adjust,

(iii) power output and efficiency are also low.

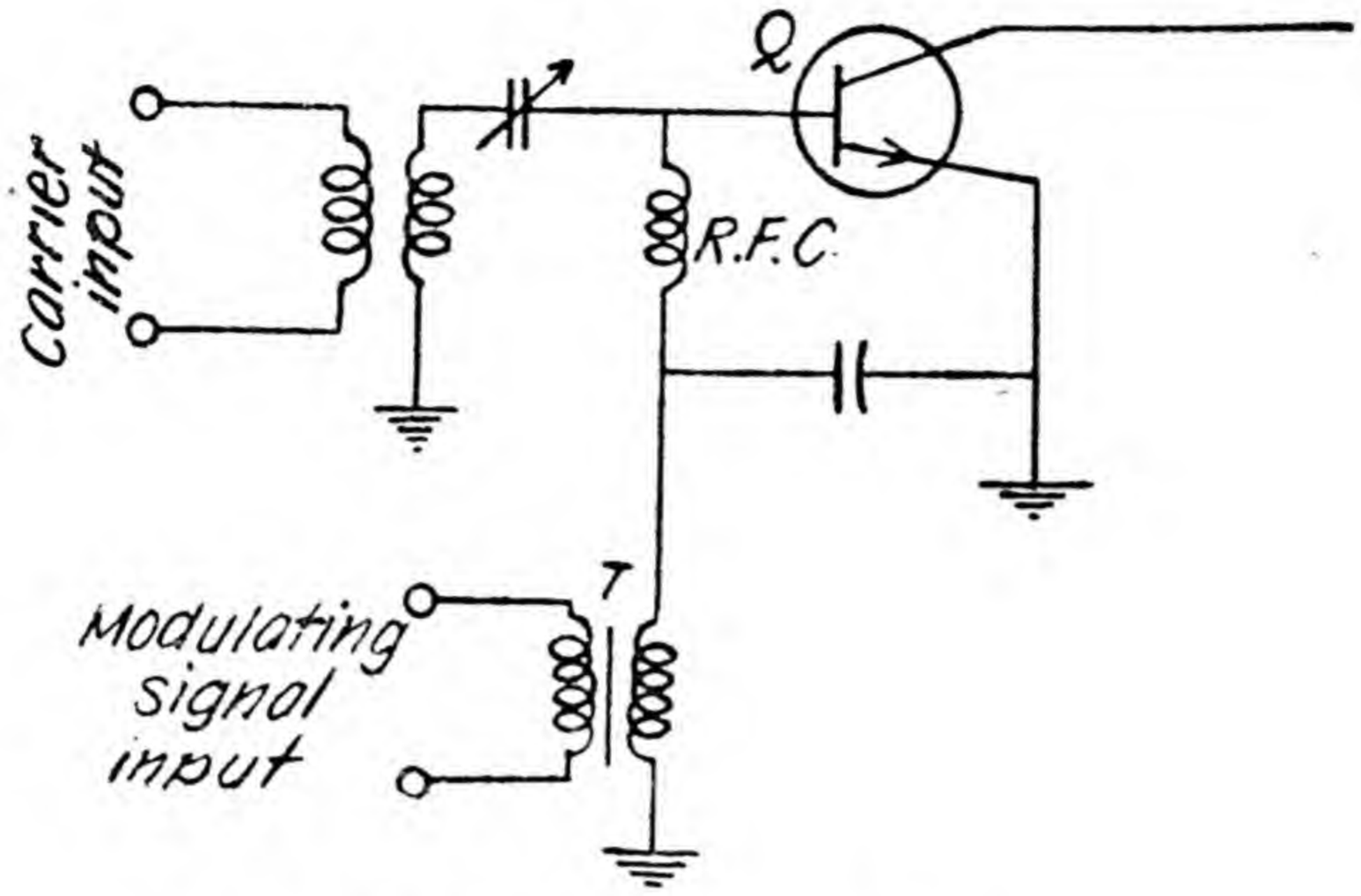


Fig. 9. Base modulation circuit.

Emitter modulation : Its characteristics lie in between base and collector modulation. Refer to fig. 10. The modulating signal is applied to the emitter circuit. Resistor *R*, in combination with

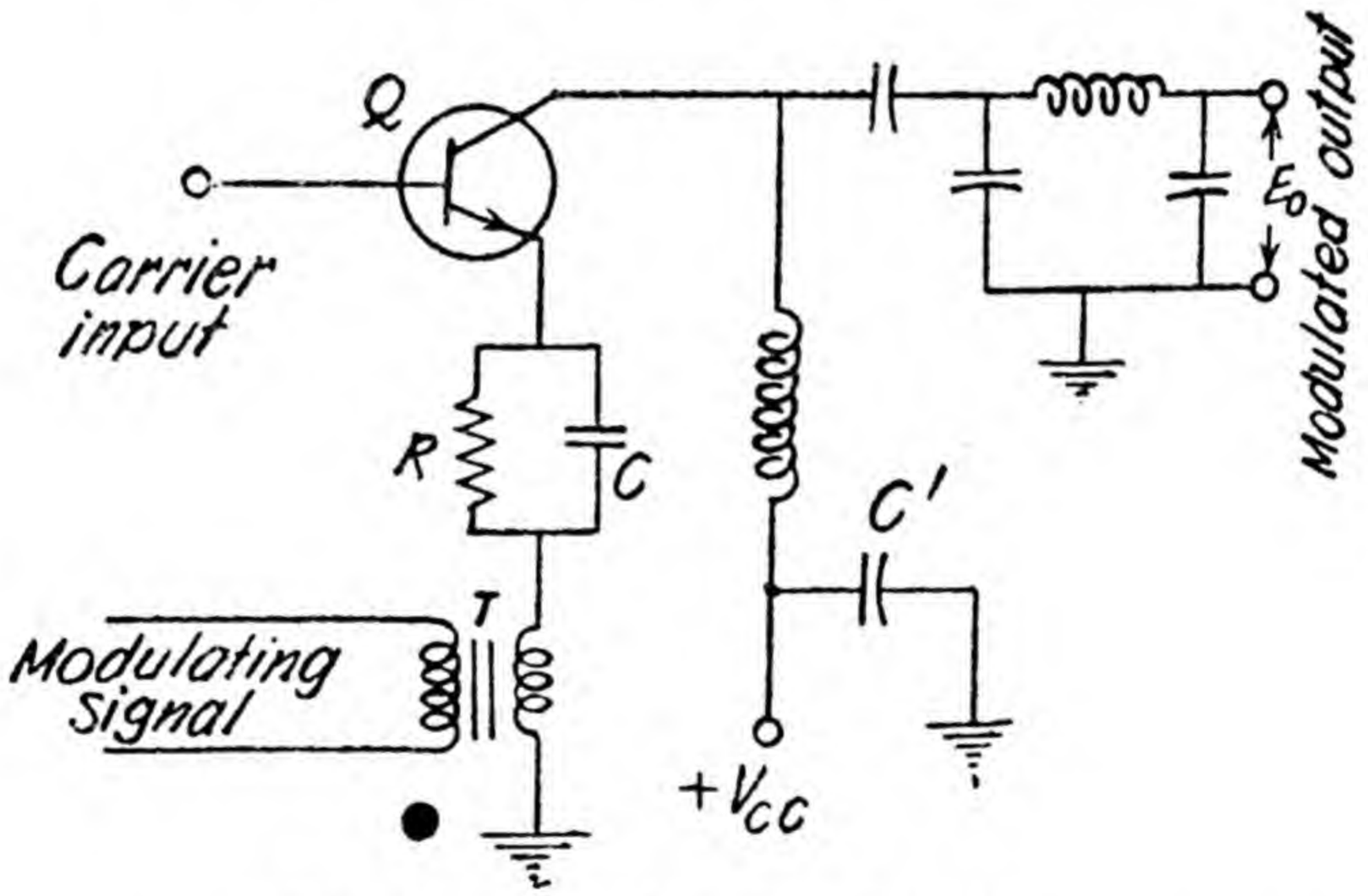


Fig. 10. Emitter modulator.

C, develops the proper bias. Transformer, *T*, provides better matching of the impedance. *C'* serves to keep r.f. out of power supply.

30·4. SQUARE LAW MODULATION :

Important methods in square law modulation are :

(*i*) grid modulation or Vander Bijl modulator or square law modulation,

(*ii*) balanced modulator,

(*iii*) square law diode modulation.

We shall discuss only two of these methods.

30·4-1. GRID MODULATION :

The circuit of Vander Bijl modulator is shown in fig. 11.

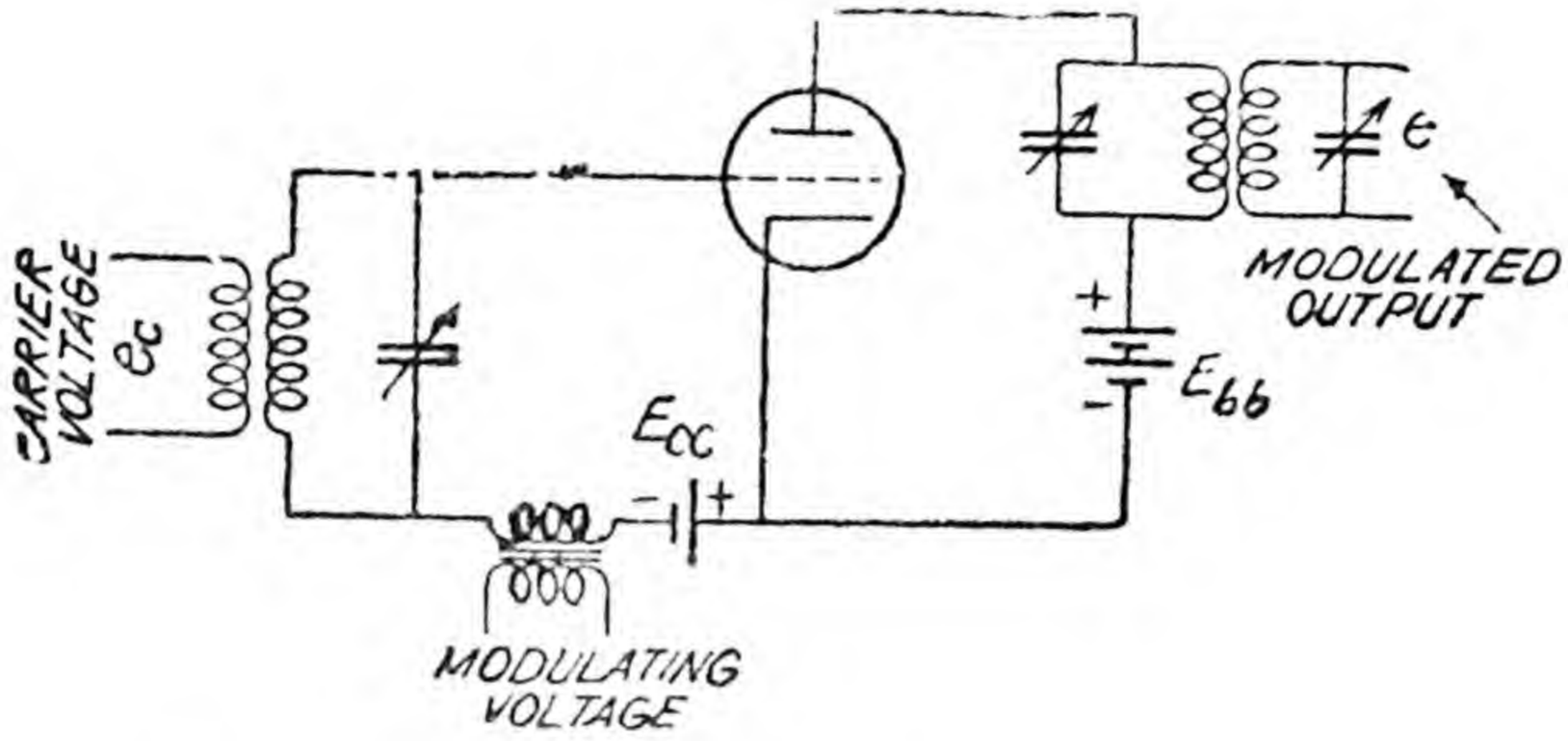


Fig. 11. Circuit of Vander Bijl modulator.

The carrier and modulation frequency potentials are both applied in the grid circuit of a triode or tetrode etc. The modulated output is obtained from plate tank circuit which is tuned to the carrier frequency.

Analysis : In fig 12, dynamic transfer characteristic of triode, used in the circuit, is shown alongwith the plate current waveform. It is suppo sed that the transfer curve is parabolic over the range of operation and therefore a.c. plate current i_p may be expressed in terms of the a.c grid voltage e_g by the first two terms of Taylor series expansion, *i.e.*,

$$i_p = a_1 e_g + a_2 e_g^2, \qquad \ldots(1)$$

where a_1 and a_2 are constants.

The excitation potential e_g is of the form

$$e_g = e_c + e_m$$
$$= E_c \cos \omega_c t + E_m \cos \omega_m t, \qquad \ldots(2)$$

and thus from eqs. (1) and (2), expression for plate current becomes

$$i_p = a_1 [E_m \cos \omega_m t + E_c \cos \omega_c t] + a_2 [E_m \cos \omega_m t + E_c \cos \omega_c t]^2,$$
$$i_p = a_1 E_m \cos \omega_m t + a_1 E_c \cos \omega_c t + a_2 E_m^2 \cos^2 \omega_m t + a_2 E_c^2 \cos^2 \omega_c t + 2a_2 E_c E_m \cos \omega_m t \cos \omega_c t$$
$$= a_1 E_c \cos \omega_c t + a_2 E_m E_c \cos (\omega_c + \omega_m) t + a_2 E_c E_m \cos (\omega_c - \omega_m) t + a_1 E_m \cos \omega_m t + \frac{a_2 E_m^2}{2} + \frac{a_2 E_m^2}{2} \cos 2\omega_m t + \frac{a_2 E_c^2}{2} + \frac{a_2 E_c^2}{2} \cos 2\omega_c t. \qquad \ldots(3)$$

Assuming $\omega_c \gg \omega_m$, so that the circuit, tuned to ω_c, responds only to the terms of frequency ω_c, $(\omega_c + \omega_m)$ and $(\omega_c - \omega_m)$. Rest terms of above equation do not produce appreciable voltage across the tuned circuit and thus the plate current is given by

$$i_p = a_1 E_c \cos \omega_c t + a_2 E_m E_c \cos (\omega_c + \omega_m) t + a_2 E_m E_c \cos (\omega_c - \omega_m) t$$
$$= a_1 E_c \left(1 + \frac{2a_2 E_m}{a_1} \cos \omega_m t \right) \cos \omega_c t$$

$$=a_1E_o\,(1+m_a\cos\omega_m t)\cos\omega_c t,$$

where modulation index $m_a=\dfrac{2a_2E_m}{a_1}$.

Therefore form of modulated output is

$$e=E_o\,(1+m_a\cos\omega_m t)\cos\omega_c t. \qquad ...(4)$$

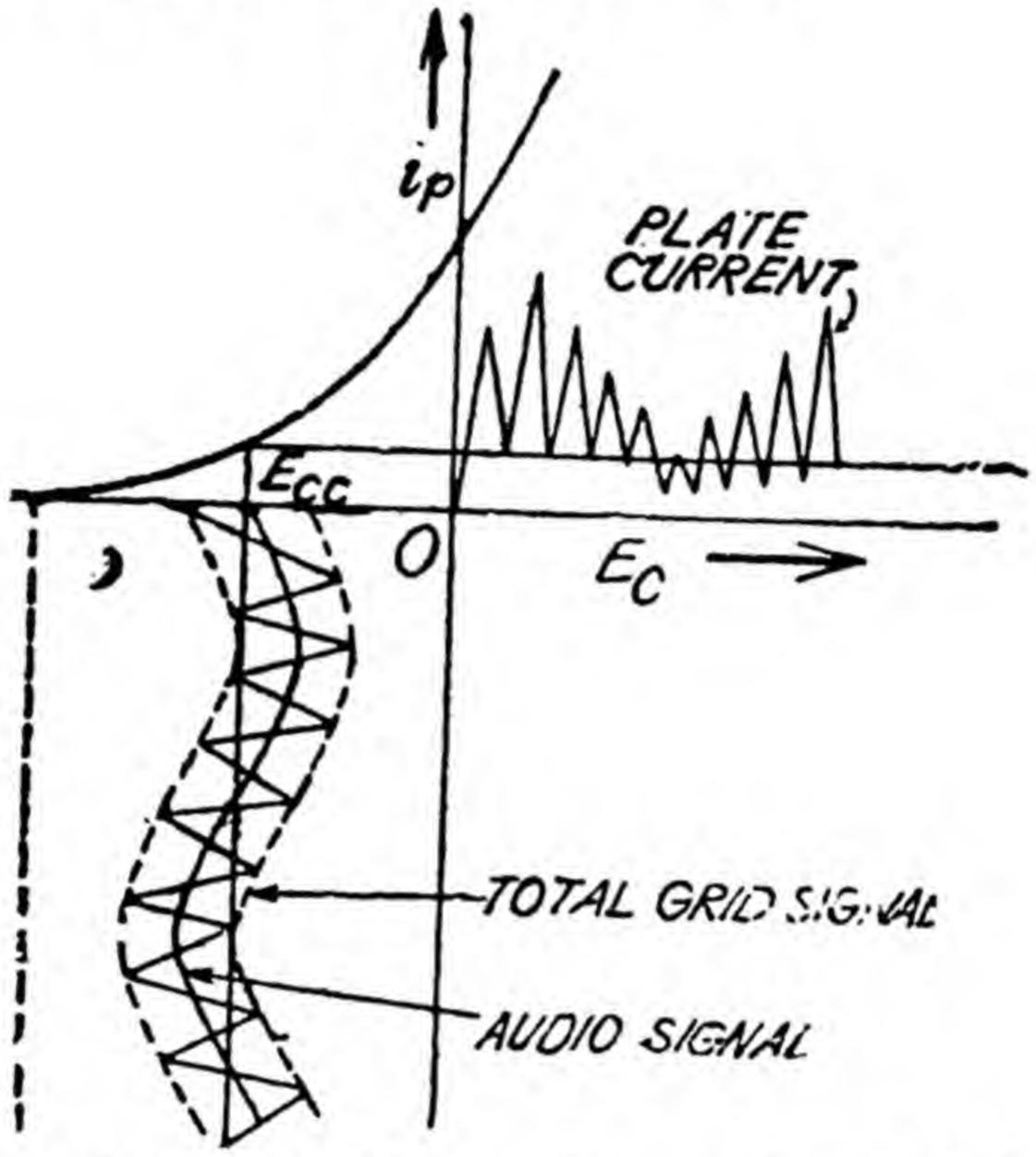


Fig. 12. Waveshapes of plate current and total grid signal.

Advantages :

(*a*) In this method grid current is very small due to which little power is required from the source of carrier voltage and modulation voltage.

(*b*) Distortion in the output is not too large.

Disadvantages :

(*a*) The plate circuit efficiency is low.

(*b*) In order to suppress the amplitudes of undesired side frequencies, the modulation index should be low : therefore the method is suitable for small percentage modulation.

30·4-2. BALANCED MODULATOR :

This consists of a circuit similar to push pull amplifier arrangement.

In this method we shall discuss the carrier suppression by means of balanced modulator and the single side band generation by a set of balanced modulators.

(A) Modulator for Carrier Suppression · The circuit for the carrier suppression is shown in fig. 13.

It is assumed that (*i*) the tubes T_1 and T_2 have identical characteristics, (*ii*) the circuit is symmetrical with respect to the centre tap of the transformer windings.

The carrier voltage is applied to the grids of the two tubes in the same phase, while the modulating signal is applied in the opposite phase to the two tubes by means of centre tapped transformer. The outputs of the two tubes are combined through a transformer with a centre-tapped primary.

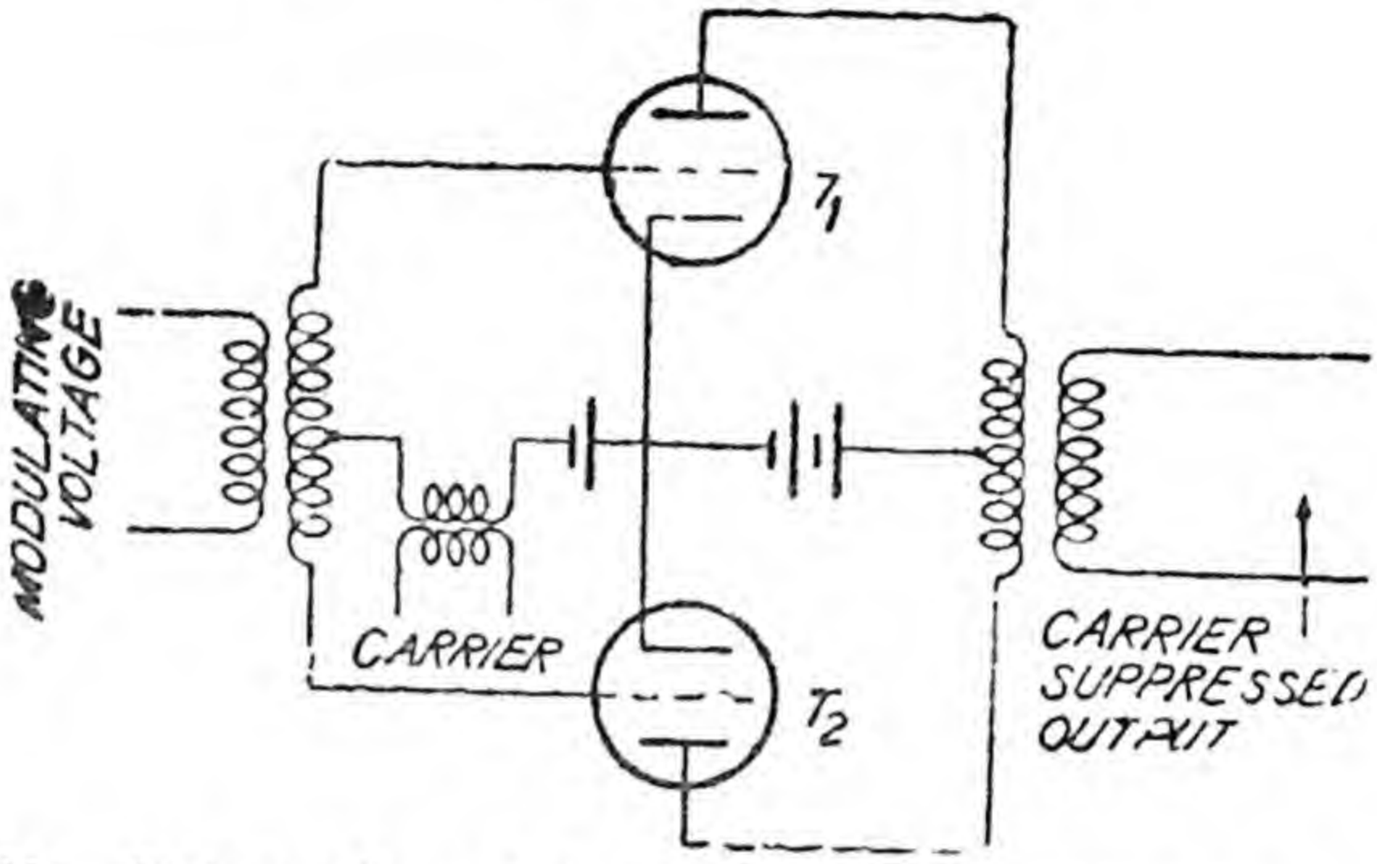


Fig. 13. Balanced modulator circuit for carrier suppression.

The grid bias is so adjusted that the tubes operate as class *A* amplifier and operation takes place over the curved portion of the dynamic characteristics of the two identical tubes. The a.c. component of plate current can, therefore. be represented as

$$i_p = a_1 e_g + a_2 e_g^2 + \ldots$$

Since modulating signal is applied in opposite phase to the two tubes, we can write

$$e_{g1} = E_m \cos \omega_m t + E_c \cos \omega_c t$$

and

$$e_{g2} = -E_m \cos \omega_m t + E_c \cos \omega_c t$$

So that for tube T_1,

$$\begin{aligned} i_{p1} &= a_1 (E_m \cos \omega_m t + E_c \cos \omega_c t) \\ &\quad + a_2 (E_m \cos \omega_m t + E_c \cos \omega_c t)^2 + \ldots \\ &= a_1 E_m \cos \omega_m t + a_1 E_c \cos \omega_c t + a_2 E_m^2 \cos^2 \omega_m t \\ &\quad + 2a_2 E_m E_c \cos \omega_m t \cos \omega_c t + a_2 E_c^2 \cos^2 \omega_m t \end{aligned}$$

and for tube T_2,

$$\begin{aligned} i_{p2} &= a_1 (-E_m \cos \omega_m t + E_c \cos \omega_c t) \\ &\quad + a_2 (-E_m \cos \omega_m t + E_c \cos \omega_c t)^2 + \ldots \\ &= -a_1 E_m \cos \omega_m t + a_1 E_c \cos \omega_c t + a_2 E_m^2 \cos^2 \omega_m t \\ &\quad - 2a_2 E_m E_c \cos \omega_m t \cos \omega_c t + a_2 E_c^2 \cos^2 \omega_c t. \end{aligned}$$

As these currents flow in opposite directions from the centre tap of the primary, the output voltage across the secondary will be proportional to the difference of i_{p1} and i_{p2}. Therefore

$$\begin{aligned} e_0 &= K (i_{p1} - i_{p2}) \\ &= 2K (a_1 E_m \cos \omega_m t + 2a_2 E_m E_c \cos \omega_m t \cos \omega_c t) \\ &= 2K [a_1 E_m \cos \omega_m t + a_2 E_m E_c \{\cos (\omega_c + \omega_m) t \\ &\quad + \cos (\omega_c - \omega_m) t\}] \end{aligned}$$

As the load is tuned to carrier frequency ω_c and ω_c is much greater than ω_m, the first term in the above expression will not develop any significant output voltage. Therefore output voltage is

$$e_0 = 2K\, a_2\, E_m\, E_c\, [\cos(\omega_c + \omega_m)\, t + \cos(\omega_c - \omega_m)\, t]$$

which consists of two side bands only.

Thus it is clear from above analysis that the carrier will be cancelled while the side bands appear in the output. This suppression of carrier results in the economy of power.

Waveform when carrier has been suppressed is shown in fig. 14.

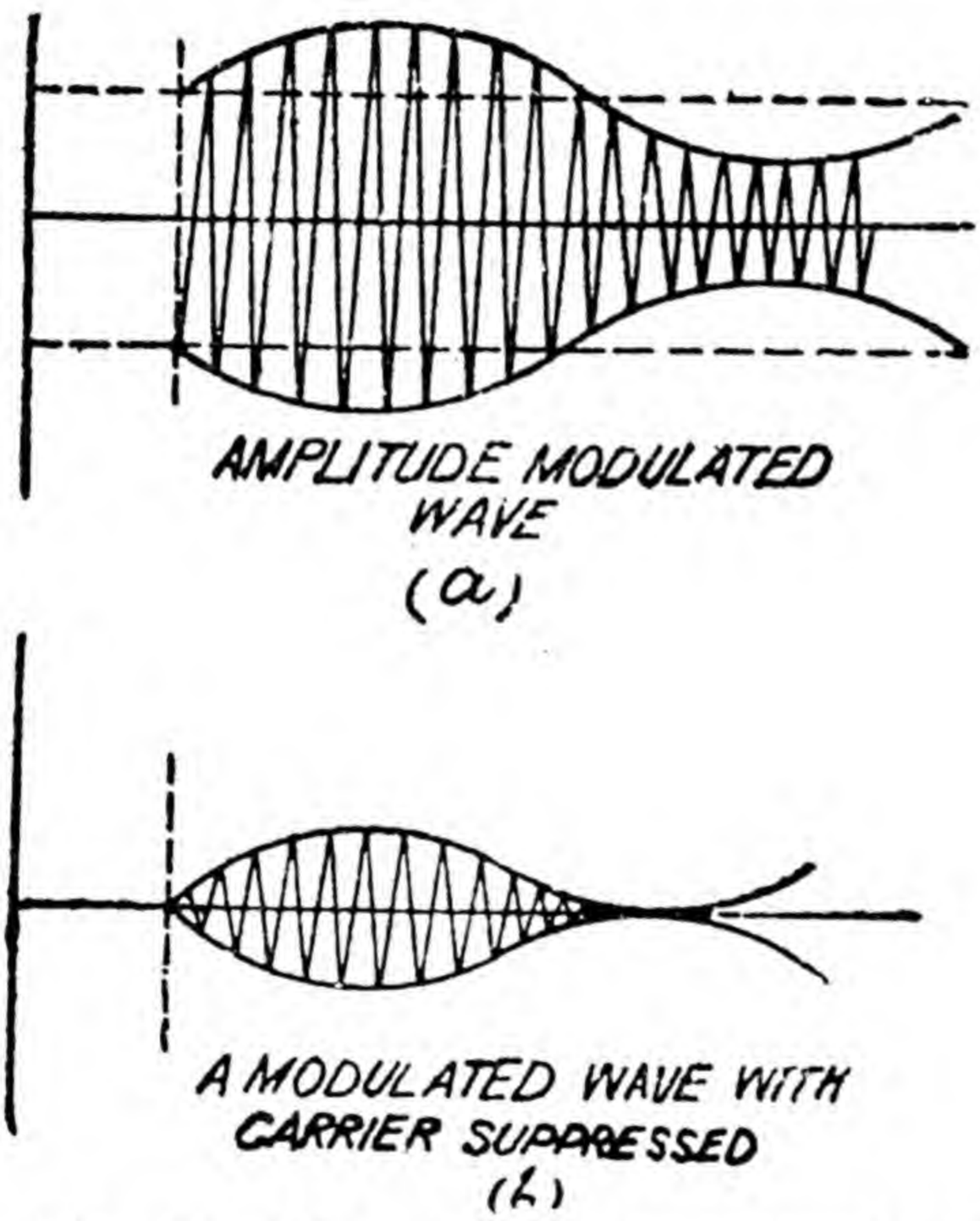


Fig. 14. Modulated wave representation.

Transistor Balanced Modulator :

Refer to fig. 15. Here the voltages have been denoted by v and V, instead of letters e and E.

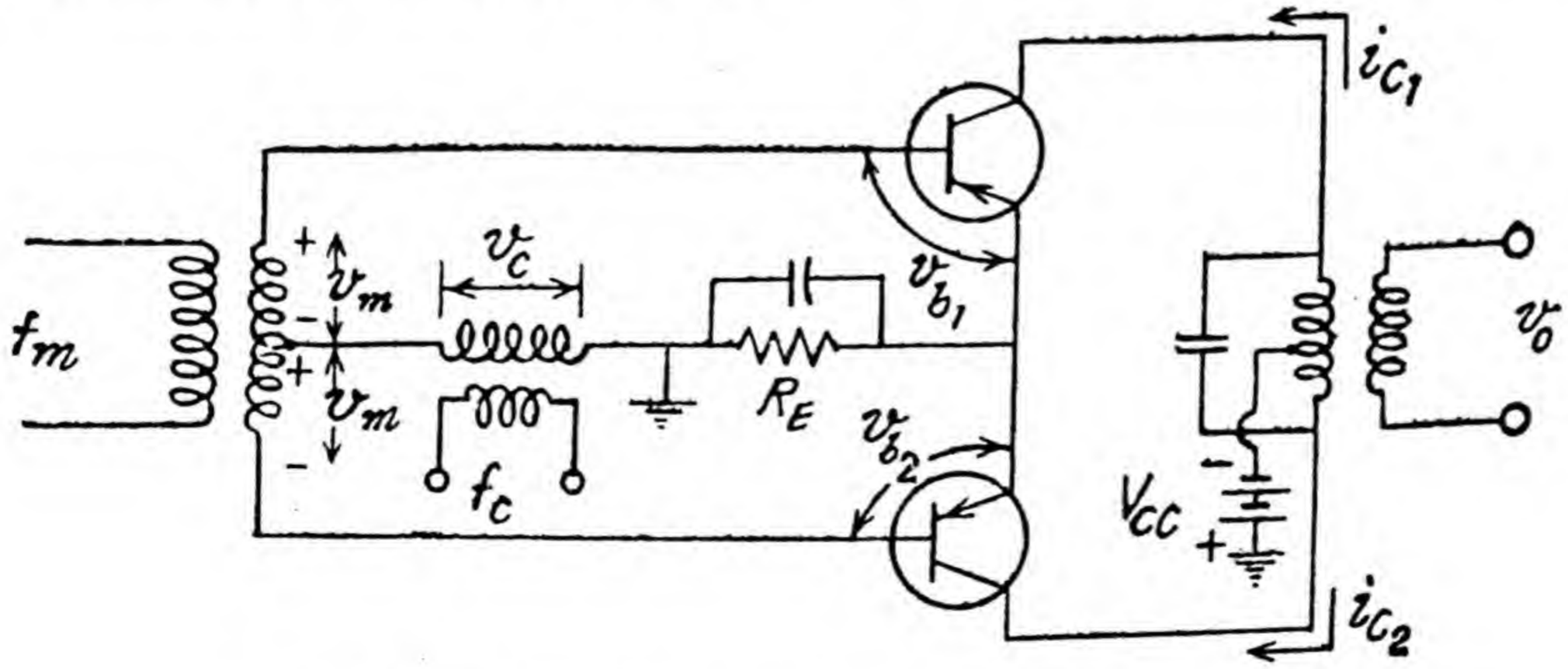


Fig 15. Transistor balanced modulator.

In this circuit,

$$v_{b1}=v_{m1}+v_o$$
$$=E_m \cos \omega_m t+E_c \cos \omega_c t$$
$$v_{b2}=v_{m2}+v_c$$
$$=-v_{m1}+v_c=-E_m \cos \omega_m t+E_o \cos \omega_c t$$

Then $i_{c1}=a_1 v_{b1}+a_2 v_{b1}^2$

$$i_{c2}=a_1 v_{b2}+a_2 v_{b2}^2$$

or $$i_{c1}=a_1 (E_m \cos \omega_m t+E_c \cos \omega_c t)+a_2 (E_m \cos \omega_m t+E_o \cos \omega_c t)^2$$

$$i_{c2}=a_1 (-E_m \cos \omega_m t+E_o \cos \omega_c t)+a_2 (-E_m \cos \omega_m t+E_o \cos \omega_c t)^2$$

so that output voltage is

$$v_0=K (i_{c1}-i_{c2})$$
$$=K [2a_1 E_m \cos \omega_m t+4a_2 E_m E_o \cos \omega_m t \cos \omega_c t]$$

The output circuit is tuned to carrier frequency. It will eliminate, ω_m, so that

$$v_0=K.4a_2 E_m E_c \cos \omega_m \cos \omega_c t$$
$$=2K a_2 E_c E_m [\cos (\omega_c+\omega_m) t+\cos (\omega_c-\omega_m) t]$$

output circuit will accept the side band frequencies $(\omega_c-\omega_m)$ and $(\omega_c+\omega_m)$.

Some features of this method are :

(*i*) efficiency is large because 2/3 power is saved due to suppression of carrier,

(*ii*) it affords a means of secrecy system, and

(*iii*) one side band can also be eliminated.

(B) SINGLE SIDE BAND GENERATION : An amplitude modulated wave consists of carrier and two side bands. The amplitude, frequency and phase of the carrier component remain unaffected by the presence or absence of modulation or, in other words, the carrier has none of the intelligence to be transmitted. Therefore as discussed in (*a*) there is no need to transmit this component and is suppressed for the economy of power.

In the remaining two side bands, each contains all of the information present in the modulated wave. Therefore, it is possible to convey all the information represented by a modulating signal by transmitting only a single side band while suppressing both the carrier and the other side band The frequency band required by such single side band system is half of that occupied by modulated wave having two side bands and the carrier. Two thirds of power is saved in single side band system only due to the suppression of carrier.

Phase Cancellation Method :

The method of obtaining a single side band consists of two balanced modulators. The radio frequency input of the two balanced modulators differs in phase by 90°. The modulating inputs are identical except that the corresponding frequency components applied to the two modulators also differ by 90° in phase. The procedure is explained in the block diagram (fig. 16).

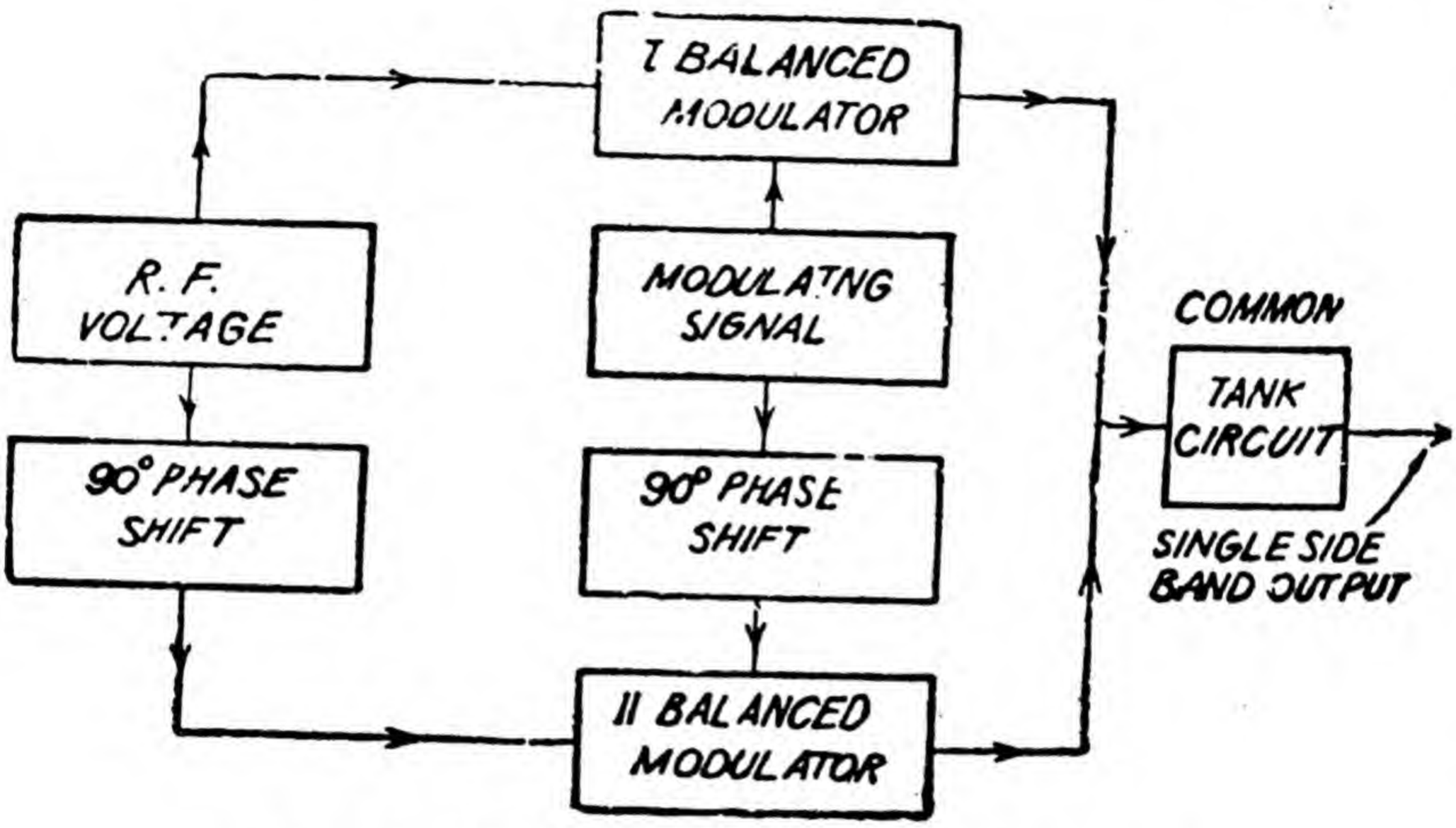


Fig. 16. Block diagram of side band system.

It is clear form the following analysis that the output of the system will have only one side band.

For the first balanced modulator, the equation of wave with carrier removed can be written as

$$e_1 = m_a E_o \sin \omega_m t \sin \omega_o t. \qquad ...(1)$$

Since for second balanced modulator, the modulating frequency ω_m and carrier frequency ω_o are shifted 90° in phase, therefore equation of wave will be

$$e_2 = m_a E_c \sin (\omega_m + 90°)t . \sin (\omega_o + 90°)t$$
$$= m_a E_c \cos \omega_m t \cos \omega_o t. \qquad ...(2)$$

Since e_1 and e_2 are combined in the common tank circuit, the resultant output will be given by

$$E_0 = e_1 + e_2 = m_a E_c \sin \omega_m t \sin \omega_c t + m_a E_o \cos \omega_m t \cos \omega_o t$$
$$= \frac{m_a E_c}{2}\left[\cos (\omega_o - \omega_m)\, t - \cos (\omega_c + \omega_m)\, t\right]$$
$$+ \frac{m_a E_c}{2}\left[\cos (\omega_o - \omega_m)\, t + \cos (\omega_c + \omega_m)\, t\right]$$
$$= m_a E_o \cos (\omega_o - \omega_m)\, t, \qquad ...(3)$$

which respresents the lower side band.

The schematic circuit diagram can be drawn for single side system as shown in fig. 17.

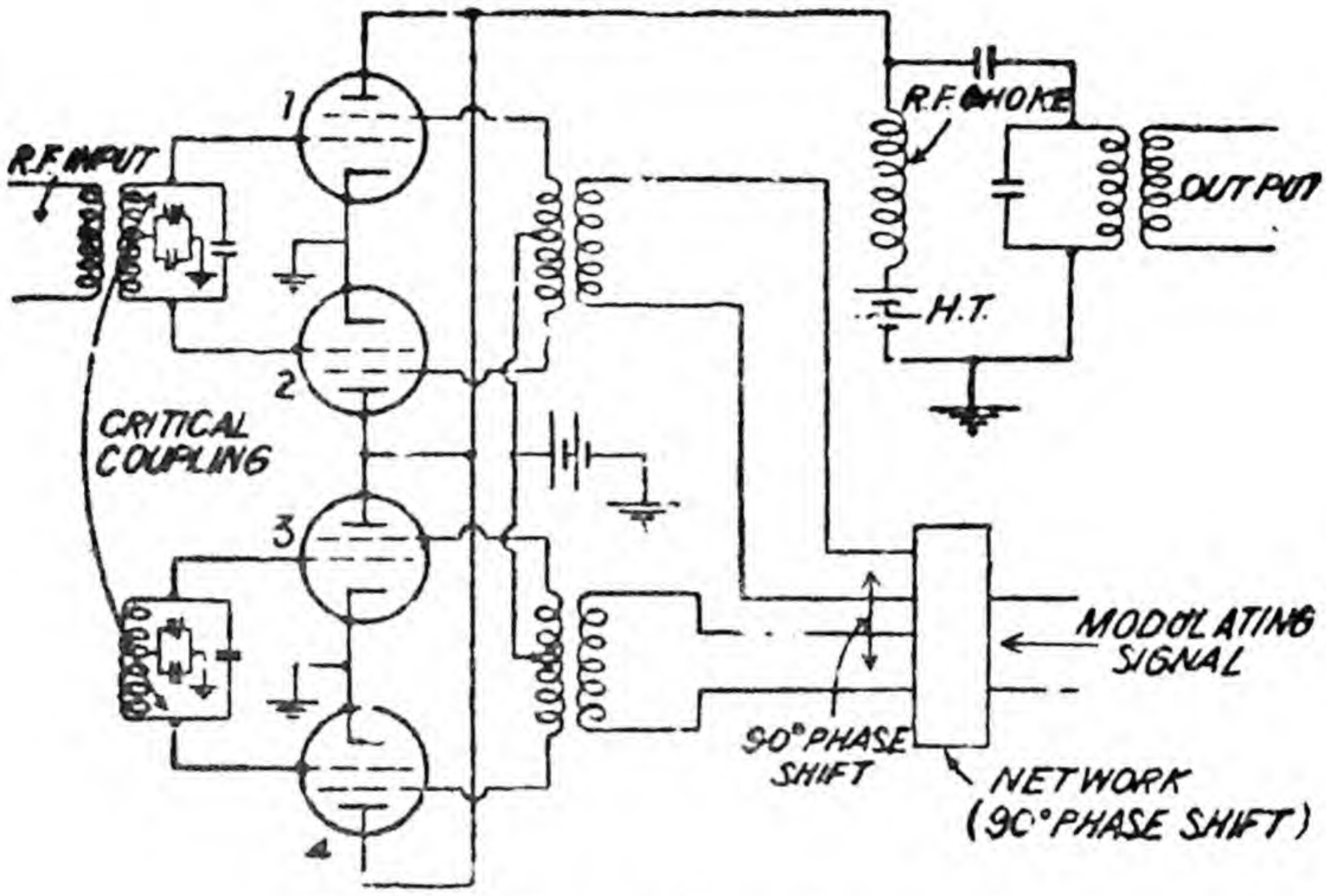


Fig. 17. Single sideband system circuit.

The radio frequency voltages are applied to the control grids. The magnitudes of r.f. voltage are kept *same* by obtaining them from a pair of critically coupled resonant circuits tuned to the same frequency. The modulating voltages are applied to the screen grids having equal magnitude but 90° phase difference.

Advantages of single Side Band (SSB) Transmission :

(*i*) Bandwidth in *SSB* transmission is half than in double side band system *e.g.* bandwidth occupied by one radio telephone channel is reduced to half *i.e.* only 3*kc/s* instead of 6*kc/s*. Thus twice the number of channels can be accommodated in a given frequency spectrum through the use of *SSB* system. In other-words, double the number of stations can be accommodated on the same band.

(*ii*) In *SSB* system, no carrier is transmitted and therefore possibility of interference with other channels is avoided.

(*iii*) Since for demodulation of *SSB* singal, carrier of same frequency, phase and of requisite strength is to be inserted at the receiver, one can not get output audio signal without the knowledge of the carrier. Since the carrier is not transmitted in *SSB* system, some secrecy is automatically achieved in such a transmission.

(*iv*) The improvement in signal to noise ratio is from 10 to 12 db at the receiver output over that in double sideband system.

(*v*) *SSB* system eliminates the possibility of distortion due to selective fading. Fading occurs due to multipath propagation of electromagnetic waves. Thus radio waves at the same frequency may travel by two paths which may be of different lengths so that

signals received by these paths may be of unequal strengths and phases, resulting in fading. In selective fading, fading is selective over the received band. When transmitted signals consists of two sidebands and carrier (double sideband system) the following three types of selective fading occurs. In one type one sideband fades completely leaving other sideband and carrier unaffected, in other type of fading carrier alone fades out, and in third type of fading the amplitude and phase of one sideband component varies with respect to the other sideband and carrier. This latter type of fading gives rise to severe distortion. In *SSB* system since only one sideband is present, possibility of such type of fading is eliminated.

(*vi*) *SSB* system provides an improvement in signal to noise ratio of atleast 9 dbs. Thus in order to get the same signal to noise ratio at the receiver output, the transmitter average power output may be reduced by 9 dbs. *i.e.* to about one eighth as compared with *DSB* transmitter. Therefore *SSB* transmitter will require less number of output amplifying power stages ; consequently weight, volume and operating cost is reduced.

Disadvantages of SSB Transmission :

The major disadvantages of *SSB* system is that the transmitter and receiver become more complex and performance required is of high standard as briefed below :

For demodulation of *SSB* signal, carrier is inserted at the receiver. The frequency of this reinserted carrier must be within 15 c/s of the suppressed carrier frequency in case of speech and 4c/s in case of music. Such a requirement complicates the demodulating process ; because to meet it, this is necessary to transmit a pilot signal or the carrier voltage itself at a very low level for synchronising the receiver oscillator frequency. This signal has to be filtered at the receiver with the use of highly selective filters, amplified and then either reinserted to provide the carrier or is used to control the carrier frequency produced by the local oscillator. Design of these highly selective filters is thus involved in *SSB* receiver. This complexity contributes to an addition in cost.

Applications of SSB System :

Because of complexity and cost of *SSB* receiver, this system is not used for commercial broadcasting. It finds use in other fields such as

(*i*) Police wireless communication.
(*ii*) *SSB* telegraph system.
(*iii*) Point to point radio telephone communication.
(*iv*) In V.H.F. and U.H.F. communications.

30·5. FREQUENCY MODULATION :

As described earlier frequency modulation is produced when the instantaneous frequency of carrier is varied inaccordance with the modulating signal, while the amplitude of the carrier remains constant.

Characteristics of F.M. Wave : The amount of deviation in frequency is proportional to the amplitude of the signal, while the rate is determined by the frequency of modulation source. For example, suppose 100 mc. carrier is being modulated by a sinusoidal signal and for given modulation system there is 2 kc., frequency shift per volt of signal magnitude. Then, if the modulating signal has an amplitude of 20 volt and frequency of 2 kc., the instantaneous frequency of modulated wave will vary between 99·96 mc., and 100·04 mc, at a rate of 2000 times per second.

Figure 18 (*a*) illustrates the form of a frequency modulated wave when a sinusoidal modulating voltage is employed.

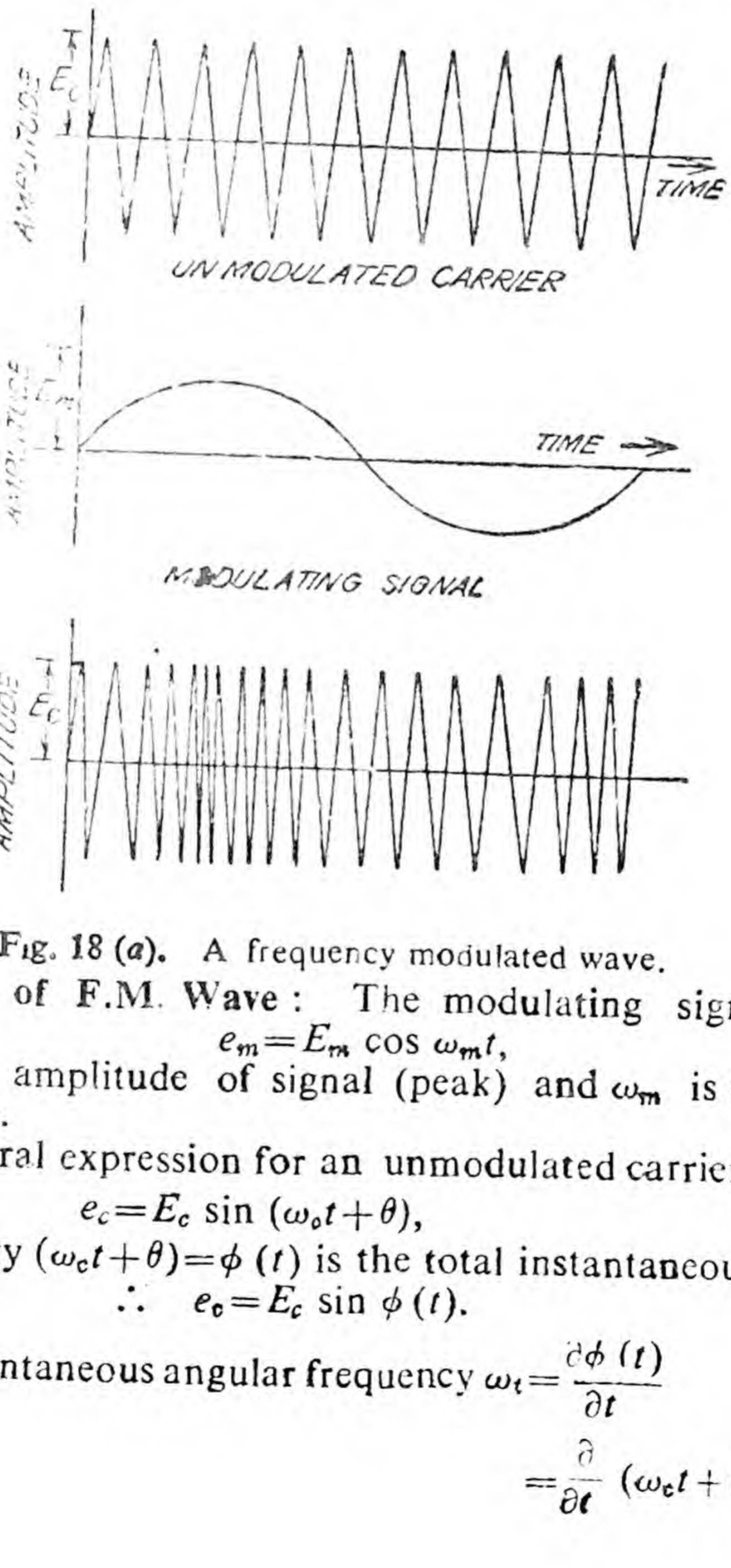


Fig. 18 (*a*). A frequency modulated wave.

Analysis of F.M. Wave : The modulating signal can be written as

$$e_m = E_m \cos \omega_m t, \qquad \ldots(1)$$

where E_m is amplitude of signal (peak) and ω_m is frequency of signal voltage.

The general expression for an unmodulated carrier is given by

$$e_c = E_c \sin (\omega_c t + \theta), \qquad \ldots(2)$$

where quantity $(\omega_c t + \theta) = \phi(t)$ is the total instantaneous 'phase' of function,

$$\therefore \quad e_c = E_c \sin \phi(t). \qquad \ldots(3)$$

The instantaneous angular frequency $\omega_i = \dfrac{\partial \phi(t)}{\partial t}$

$$= \frac{\partial}{\partial t} (\omega_c t + \theta) = \omega_c.$$

By definition, the instantaneous angular frequency of frequency modulated wave can be written as

$$\omega_i = \omega_c + K_f E_m \cos \omega_m t, \qquad \ldots(4)$$

where proportionality factor K_f determines the maximum variation in frequency for a given signal strength E_m.

$$\omega_i = \frac{\partial \phi\,(t)}{\partial t}.$$

So
$$\phi(t) = \int_0^t \omega_i \, dt$$

$$= \int_0^t (\omega_c + K_f E_m \cos \omega_m t)\, dt$$

$$= \omega_c t + K_f \frac{E_m}{\omega_m} \sin \omega_m t + \theta_0. \qquad \ldots(5)$$

The initial phase θ_0 is neglected as it plays no part in modulation process.

Thus f.m. wave can be written as

$$e = E_0 \sin \left(\omega_0 t + K_f \frac{E_m}{\omega_m} \sin \omega_m t \right) \qquad \ldots(6)$$

The instantaneous frequency of f.m. wave is

$$f = \frac{\omega}{2\pi} = f_0 + K_f \frac{E_m}{2\pi} \cos \omega_m t.$$

Therefore
$$f_{max.} = f_c + K_f \frac{E_m}{2\pi},$$

and
$$f_{min.} = f_0 - K_f \frac{E_m}{2\pi}.$$

$\therefore$ Frequency deviation $\Delta f = f_{max.} - f_0 = f_c - f_{min.}$

$$= \frac{K_f E_m}{2\pi}. \qquad \ldots(7)$$

The instantaneous frequency can then be expressed as

$$f = f_c + \Delta f \cos \omega_m t. \qquad \ldots(7a)$$

$$\text{The modulation index} = m_f = \frac{\text{frequency deviation}}{\text{modulating frequency}}$$

$$= \frac{\Delta f}{f_m} = \frac{K_f . E_m}{\omega_m}. \qquad \ldots(8)$$

And therefore the expression for f.m. wave becomes

$$e = E_0 \sin (\omega_c t + m_f \sin \omega_m t). \qquad \ldots(9)$$

In f.m. broadcasting system, Δf has a maximum permissible value of 75 kc/s and f_m may vary from about 33 c/s to 15 kc/s. Consequently, m_f can have values both greater than unity or less than unity whereas in amplitude modulation m_a should not exceed unity for distortionless transmission. Variation in transmitted frequency is shown in fig. 18 (*b*).

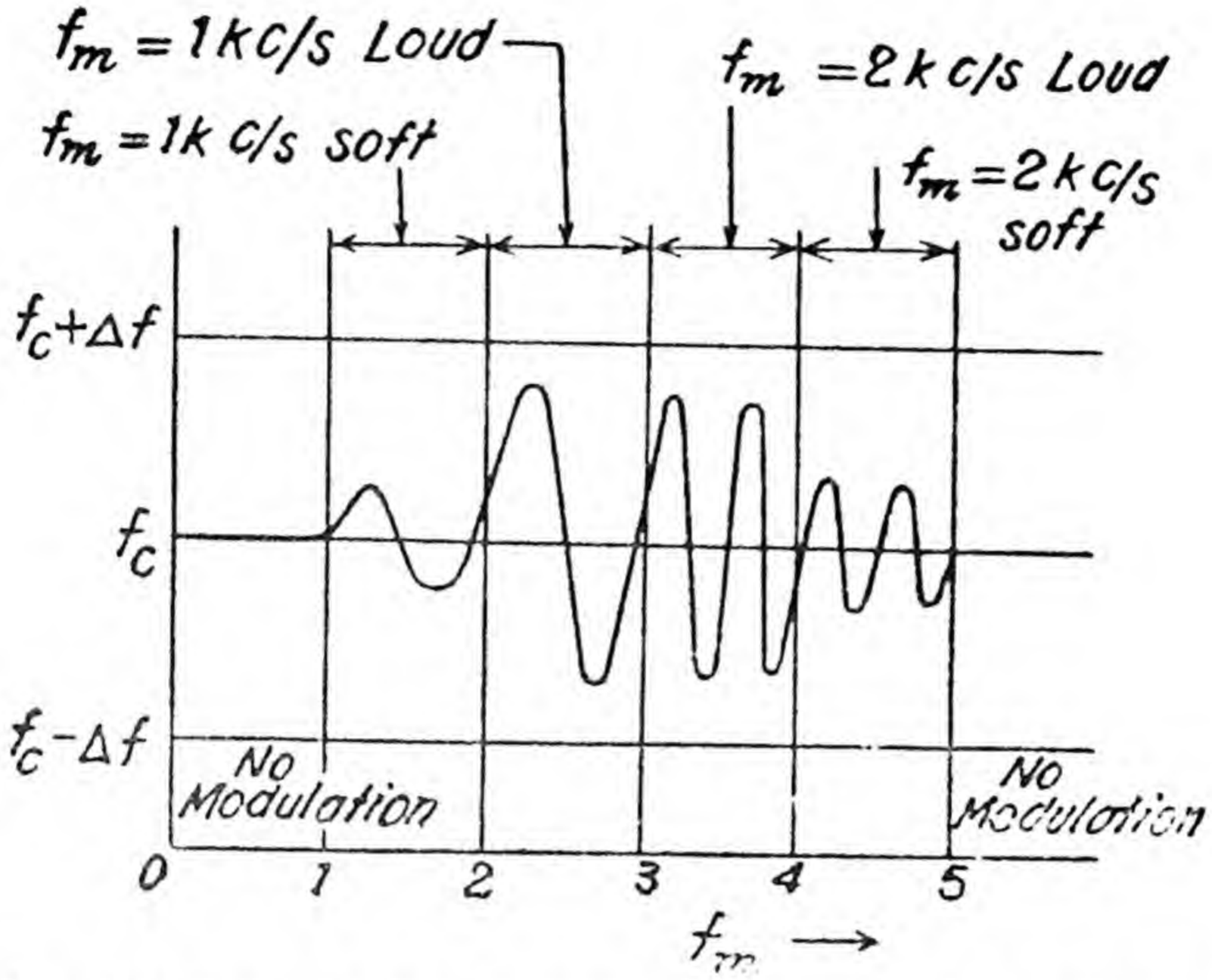


Fig. 18 (*b*). Variation of the transmitted frequency for an F. M. wave.

Power Relation in F.M. Wave ; In amplitude modulated wave, intelligence is contained in side frequencies. The amplitude and power level of the side frequencies increase with percent modulation and consequently, an increase in modulation level also increases the power level of modulated wave. On the other hand in frequency modulated wave, as the amplitude is not affected by modulation, the *total* power in the wave does not change, regardless of the degree of modulation. It is also obvious from the operation of reactance tube, where it has been shown that the effect of modulating signal is to change the capacitance in oscillator tank circuit so as to cause a shift Δf in frequency of the oscillator, the modulating signal does not change the power output of the oscillator.

Frequency Spectrum of F.M. Wave : The expression of f.m. wave can be written as

$$e = E_0 [\sin \omega_c t . \cos (m_f \sin \omega_m t) + \cos \omega_c t \sin (m_f \sin \omega_m t)]. \quad ...(10)$$

Evaluating above expression with the help of Bessel's expression,

$$\begin{aligned} e &= E_c \sin \omega_c t \, [J_0 (m_f) + 2J_2 (m_f) \cos 2\omega_m t + 2J_4 (m_f) \cos 4 \omega_m t + \ldots] \\ &\quad + E_c \cos \omega_0 t \, [2J_1 (m_f) \sin \omega_m t + 2J_3 (m_f) \sin 3\omega_m t + \ldots] \\ &= J_0 (m_f) \, E_c \sin \omega_c t \\ &\quad + J_1 (m_f) \, E_c [\sin (\omega_0 + \omega_m) t - \sin (\omega_c - \omega_m) t] \\ &\quad + J_2 (m_f) \, E_c [\sin (\omega_0 + 2\omega_m) t - \sin (\omega_c - 2\omega_m) t] \\ &\quad + J_3 (m_f) \, E_0 [\sin (\omega_0 + 3\omega_m) t - \sin (\omega_0 - 3\omega_m) t]. \quad ...(11) \end{aligned}$$

It is clear from above expression (11) that the spectrum of f.m. wave consists of carrier and an infinite number of side bands. The side frequency components extend above and below the carrier by an amount ω_m, $2\omega_m$, $3\omega_m$ etc. The side frequency pair differing

from carrier frequency by ω_m is known as the first order side frequencies and that differing by $2\omega_m$ as second order side frequencies etc.

The amplitudes of the different frequency components of f.m. wave determined by the values of various order Bessel functions depend upon the value of m_f. A graph of several of these functions is drawn in fig. 19.

It is to be noted from the graph that $J_0(m_f)$ at $m_f=2{\cdot}4$ has a root. This means that carrier will vanish when frequency deviation $\Delta f=2{\cdot}4$. Bessel functions of the first kind for m_f values upto 5 are given in the table.

TABLE : Bessel Functions of First kind :

x	n or order								
m_f	J_0	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8
0·00	1·00								
0·5	0·94	0·24	0·03						
1·0	0·77	0·44	0·11	0·02					
1·5	0·51	0·56	0·23	0·06	0·01				
2·0	0·22	0·58	0·35	0·13	0·03				
2·5	−0·05	0·50	0·45	0·22	0·07	0·02			
3	−0·26	0·34	0·49	0·31	0·13	0·04	0·01		
4	−0·40	−0·07	0·36	0·43	0·28	0·13	0·05	0·02	
5	−0·18	−0·33	0·05	0·36	0·39	0·26	0·13	0·05	0·02

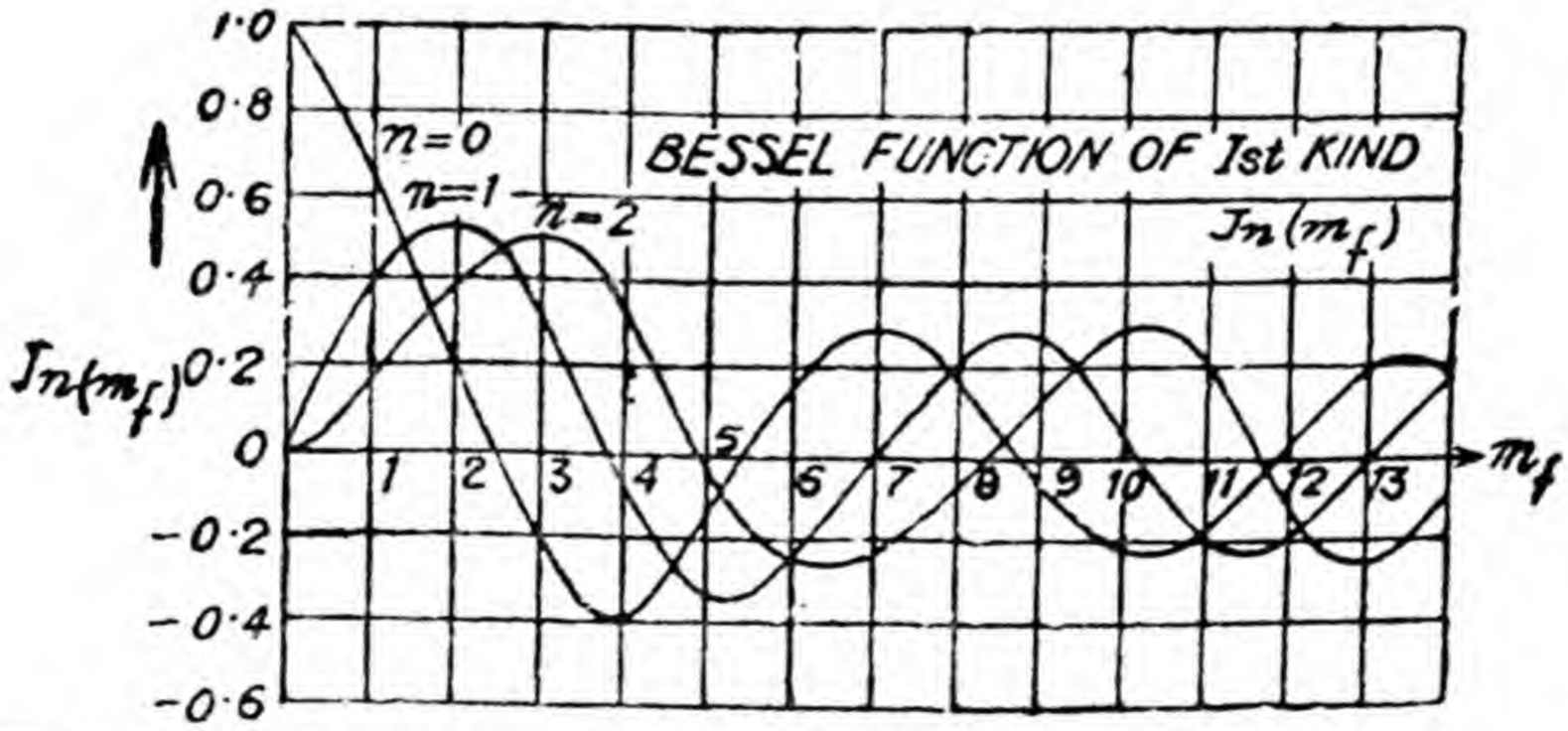


Fig. 19. Variation of amplitude in f. m. wave with m_f.

Bandwidth of F. M. Waves : It is defined as the width of frequency spectrum that contains all components having an amplitude equal to or greater than one percent of the amplitude of unmodulated wave (carrier). Suppose a given wave has n significant side frequency pairs then bandwidth is equal to

$$*BW = 2nf_m. \qquad \ldots(12)$$

The relation follows from the fact that each side frequency pair extends above and below the carrier by an *equal* amount and the spacing in the spectrum between components of successive order is equal to modulating frequency f_m. Further, from equation (8),

$$f_m = \frac{\Delta f}{m_f}.$$

Therefore,

$$BW = 2n\frac{\Delta f}{m_f}$$

or
$$\frac{BW}{2\Delta f} = \frac{n}{m_f}, \qquad \ldots(13)$$

which predicts that ratio of bandwidth to the total frequency deviation depends upon the number of significant side-frequency pairs (n) obtained with a given modulation index (m_f). For values $m_f > 10$ it is seen that the ratio $\frac{n}{m_f}$ tends to approach unity and the bandwidth is almost equal to or slightly larger than $2\Delta f$. For narrow band frequency modulation. modulation index is kept at a low value by employing small frequency deviations. Moreover, the number of significant side frequency pairs, n, increases as the modulation index, m_f, of the wave is increased. Therefore, m_f is kept low for narrow band transmission.

Suppose frequency of modulating signal f_m is doubled, then, though Δf will remain unaffected, but m_f will be halved (eq. 8).

Reduction in the value of m_f will, in its turn, reduce the value of n, the number of significant side frequency pairs. If the reduction in the value of n and m_f is by like amounts, bandwidth will remain constant even after doubling the modulating frequency.

Example 1. Find the bandwidth needed for a frequency modulated signal if the frequency deviation is 60 kc/s and the modulatirg frequency is 5 kc/s.

Frequency deviation $\Delta f = 60\ kc/s.$

Modulating frequency $f_m = 5\ kc/s.$

Then modulation index will be

$$m_f = \frac{\Delta f}{f_m} = \frac{60}{5}$$

$$= 12.$$

---

*For nth side frequency pair, one component of the pair will have frequency $(\omega_c + n\omega_m)$ and another $(\omega_c - n\omega_m)$ so that total width on frequency spectrum for this pair will be $= (\omega_c + n\omega_m) - (\omega_c - n\omega_m) = 2n\omega_m$.

From Bessel table corresponding to $m_f=12$, it can be found that there are sixteen significant side frequency pairs (*i.e.* the highest *J* coefficient for $m_f=12$ is J_{16}). That is, $n=16$, so that

$$BW=2nf_m$$
$$=2\times16\times5\ kc/s.$$
$$=160\ kc/s.$$

Example 2. A 5 kc/s. modulating frequency provides $m_f=2$ in an *FM* wave. What bandwidth is required for passing this wave ? Keeping deviation, Δf, constant if frequency of modulating signal is doubled, what will be the effect on modulation index ?

Given :

Modulation index, $m_f=2$

Modulating frequency, $f_m=5\ kc/s.$

From Bessel table, the highest *J* coefficient included for this value of m_f is J_4. Thus $n=4$, *i.e.* there are four significant side frequency pairs. Thus required bandwidth will be

$$BW=2n\ f_m$$
$$=2\times4\times5\ kc/s.$$
$$=40\ kc/s.$$

If f_m is doubled *i. e.* $2\times5=10\ k/c.$, then as

$$f_m=\frac{\Delta f}{m_f}$$

for same Δf, the modulation index, m_f, should be halved.

30·6. REACTANCE TUBE METHOD FOR PRODUCING F.M. WAVE :

This method is quite simple and is being widely adopted these days. The method is to control the frequency of an oscillator by the modulating voltage with the aid of reactance tube.

(A) Capacitive Reactance Tube :

We know that the frequency of tank circuit depends upon the reactance of the components, *i.e.*, capacitor and inductor. So if we couple some other reactance which varies with modulating voltage with the tank circuit, then the frequency generated will vary with coupled reactance and hence with modulating voltage. The output wave will be then frequency modulated.

The reactance tube is merely a pentode tube associated with a phase splitting network *CR* arranged in such a way that when a.c. voltage is applied between plate and cathode, the resulting a.c. plate current is in phase quadrature with this applied a.c. voltage. Consequently, if the plate current leads the applied a.c. voltage by

90° then the tube will offer a capacitive reactance between its plate cathode terminals, or if the plate current legs behind the a.c. voltage by 90° then the tube will offer an inductive reactance between its plate-cathode terminals X and Y.

The basic circuit arrangement and a. c. equivalent circuit is shown in figs. 20 (a) and 20 (b).

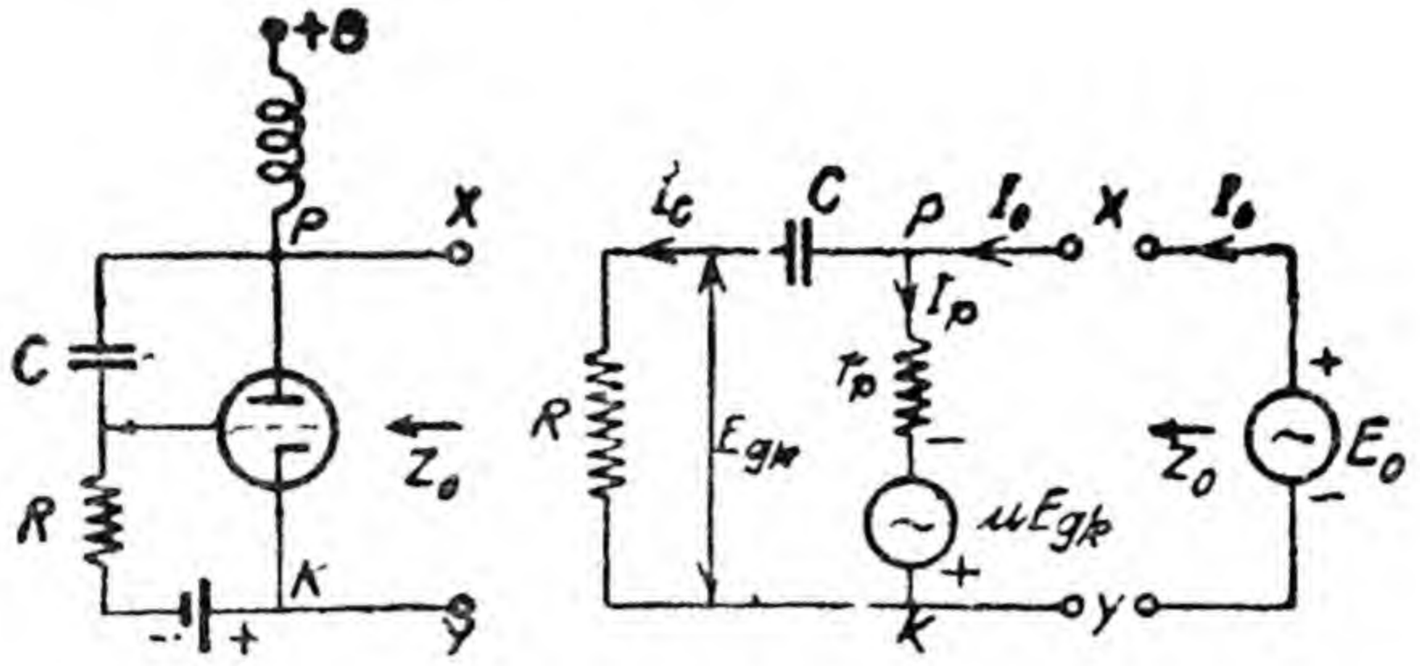


Fig. 20 (a). Basic circuit arrangement.

(b) a c. equivalent circuit.

When $R \ll \frac{1}{\omega C}$, then the current I_0 flowing through R and C leads the applied voltage by 90°. The grid to cathode voltage E_{gk} ($=I_cR$) will also lead the applied a c. voltage by 90°. The plate current I_p, being in phase with E_{gk} will also lead the applied a.c. voltage by 90°. Thus plate current is leading the plate voltage by 90° which happens in the case of a capacitive reactance or in otherwoids, tube offers a capacitive reactance between plate-cathode terminals X and Y. We shall now calculate the magnitude of this reactance.

Let us denote the impedance by Z_0 which the tube offers to a.c. voltage E_0 applied between its plate and cathode. Grid current is assumed to be zero. Then I_c, the current flowing through capacitor C will be

$$\mathbf{I}_c = \frac{\mathbf{E}_0}{R + \frac{1}{j\omega C}} = \frac{\mathbf{E}_0}{R - jX_c}.$$

The grid to cathode voltage will be

$$\mathbf{E}_{gk} = R\mathbf{I}_c = \frac{R\mathbf{E}_0}{R - jX_c}.$$

The a.c. plate current is given by

$$\mathbf{I}_p = \frac{\mathbf{E}_0 + \mu\mathbf{E}_{gk}}{r_p} = \frac{\mathbf{E}_0}{r_p} + \frac{\mu}{r_p}\,\frac{R\mathbf{E}_0}{R - jX_c}$$

Let I_0 be the current drawn by the reactance tube circuit from the applied voltage. Then

$$\mathbf{I}_0=\mathbf{I}_p+\mathbf{I}_c$$

$$=\frac{\mathbf{E}_0}{r_p}+g_m\frac{R\mathbf{E}_0}{R-jX_c}+\frac{\mathbf{E}_0}{R-jX_c}. \qquad ...(1)$$

Therefore output terminal impedance is given by

$$Y_0=\frac{1}{Z_0}=\frac{I_0}{E_0}$$

$$=\frac{1}{r_p}+\frac{g_m R}{R-jX_c}+\frac{1}{R-jX_c}$$

$$=\frac{1}{r_p}+\frac{1}{R-j/\omega C}+\frac{1}{(1/g_m)-j/g_m\omega\ CR}. \qquad ...(2)$$

If, however, r_p and $(R-j/\omega C)$ are large compared to $(1/g_m)-(j/g_m\omega CR)$ then the first two terms in above equation can be neglected. Then

$$Y_0=\frac{1}{(1/g_m)-j/g_m\ CR\omega}. \qquad ...(3)$$

Further if $\frac{1}{\omega CR}\gg 1$, that is $\frac{1}{g_m\ \omega CR}\gg\frac{1}{g_m}$ then we can write

$$Y_0=-\frac{g_m\ \omega CR}{j}$$

or

$$Z_0=-\frac{j}{\omega CRg_m}=-\frac{j}{\omega C_{eq}},$$

which is purely reactive. Thus the tube offers an equivalent capacitance CRg_m. The modulator circuit is shown in fig. 20 (*c*).

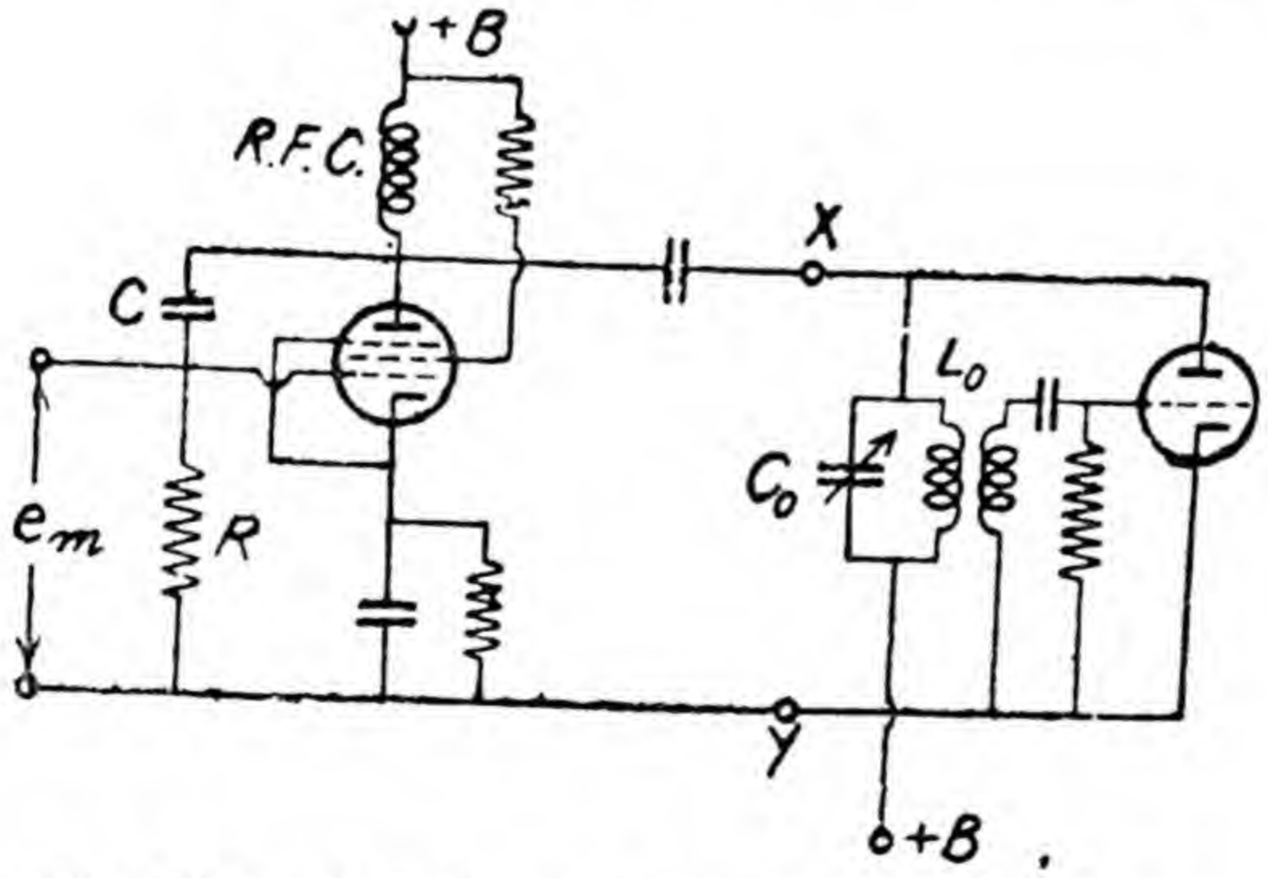


Fig. 20. (*c*) Reactance tube modulator.

Between *X-Y* terminals an oscillator circuit is connected. Thus reactance tube shunts the tank circuit of r. f. oscillator. When a modulating voltage is applied to the control grid of the tube, this varies g_m which in turn varies the equivalent capacitance C_{eq} which

is associated with C_0 and hence frequency of the oscillator is modulated.

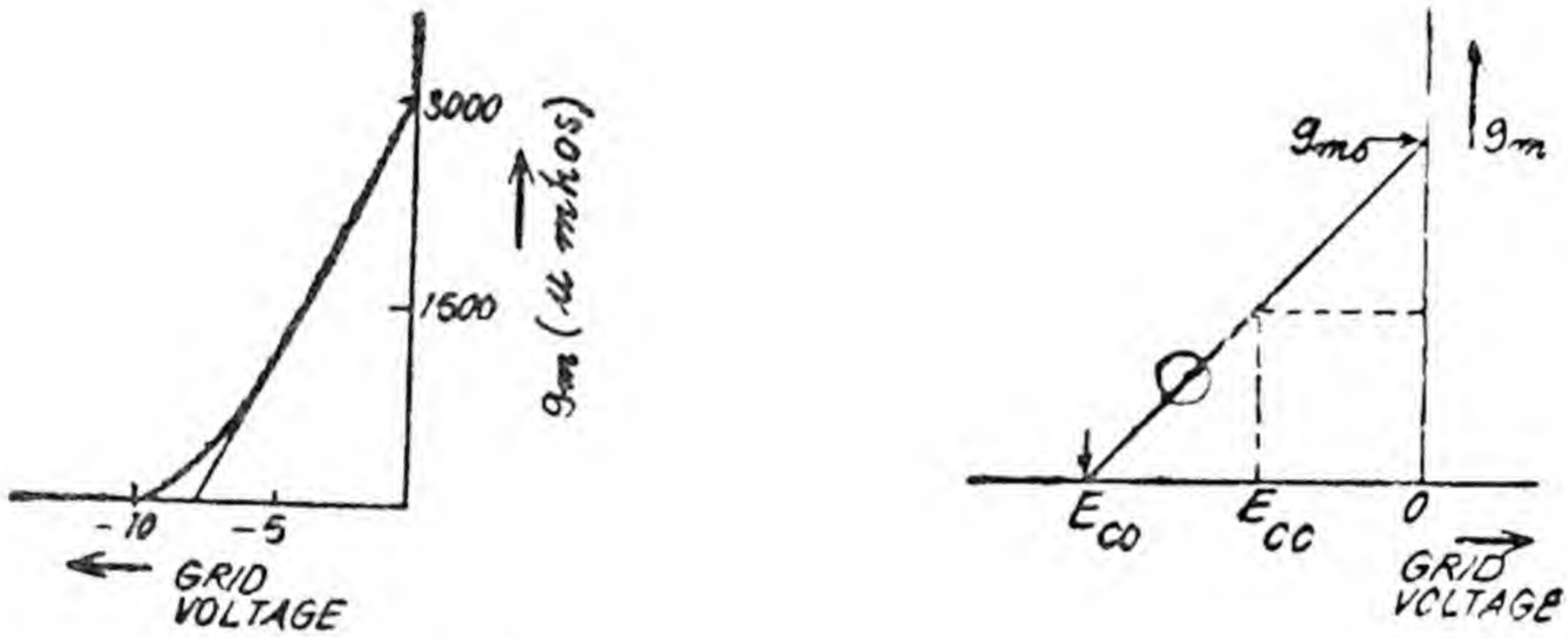
Now we shall calculate the frequency of oscillations and the modulation index. The variation of transconductance of the tube with control grid potential is shown in fig. 21 (*a*) and redrawn in fig. 21 (*b*). Suppose the bias applied is E_{co} then instantaneous control grid potential will be

$$e_c = E_{cc} + e_m = E_{cc} + E_m \cos \omega_m t, \qquad \ldots(4)$$

where e_m is the modulating signal applied at the grid. The straight line equation for the fig. 21 (*b*) can be written as

$$g_m = \text{slope of control grid voltage} + \text{intercept}$$

$$= - \left(\frac{g_{m_1}}{E_{c0}}\right)(E_{cc} + E_m \cos \omega_m t) + g_{m0}. \qquad \ldots(5)$$



(*a*) (*b*)

Fig. 21. Plot of g_m as control grid voltage.

Therefore equivalent capacitance of the reactance tube circuit will become

$$C_{eq} = CRg_m$$

$$= CRg_{m0}\left(1 - \frac{E_{cc}}{E_{c0}} - \frac{E_m \cos \omega_m t}{E_{c0}}\right). \qquad \ldots(6)$$

Since this effective capacitance is in parallel with the capacitance C_0 of the tank circuit, the frequency of oscillation will be

$$f = \frac{1}{2\pi \sqrt{\{L_0 (C_0 + C_{eq})}}$$

$$= \frac{1}{2\pi \left[L_0 C_0 + L_0 CRg_{m0} \{1 - (E_{cc}/E_{c0}) - (E_m/E_{c0}) \cos \omega_m t\}\right]^{1/2}}$$

When modulating frequency is zero, then the frequency of oscillations will be purely of carrier wave, that is,

$$f_o = \frac{1}{2\pi \left[L_0 C_0 + L_0 CRg_{m0} (1 - E_{cc}/E_{c0})\right]^{1/2}}$$

Therefore, we can write

$$\frac{f}{f_o}=\frac{1}{\left[1-\frac{L_0CRg_{m0}E_m\cos\omega_m t}{E_{c0}\{L_0C_0+L_0CRg_{m0}\ (1-E_{cc}/E_{c0})\}}\right]^{1/2}}$$

$$=1+\frac{CRg_{m0}E_m\cos\omega_m t}{2E_{c0}\{C_0+CRg_{m0}\ (1-E_{cc}/E_{c0})\}} \qquad ...(7)$$

or $$f=f_c\ (1-m_f\cos\omega_m t) \qquad ...(8)$$

where modulation index m_f is given by

$$m_f=\frac{CRg_{m0}\ E_m}{2E_{c0}\ [C_0+CRg_{m0}\ (1-E_{cc}/E_{c0})]}$$

$$=\frac{E_m}{2\left(\frac{E_{c0}\ C_0}{CRg_{m0}}+E_{c0}-E_{cc}\right)^{\frac{1}{2}}} \qquad ...(9)$$

The frequency variation effect of the reactance tube should be kept small in order to maintain the required proportionality between instantaneous modulating voltage and the frequency deviation. Therefore, in frequencey modulated transmitters, submultiple of the assigned carrier frequency is first generated by crystal oscillator and modulation is performed. After modulation, the wave is then passed through frequency multiplier circuits inorder to achieve the assigned carrier frequency of the station.

(B) Inductive Reactance Tube : The basic circuit of inductive reactance tube is shown in fig. 22 (*a*) and equivalent circuit is shown in fig. 22 (*b*).

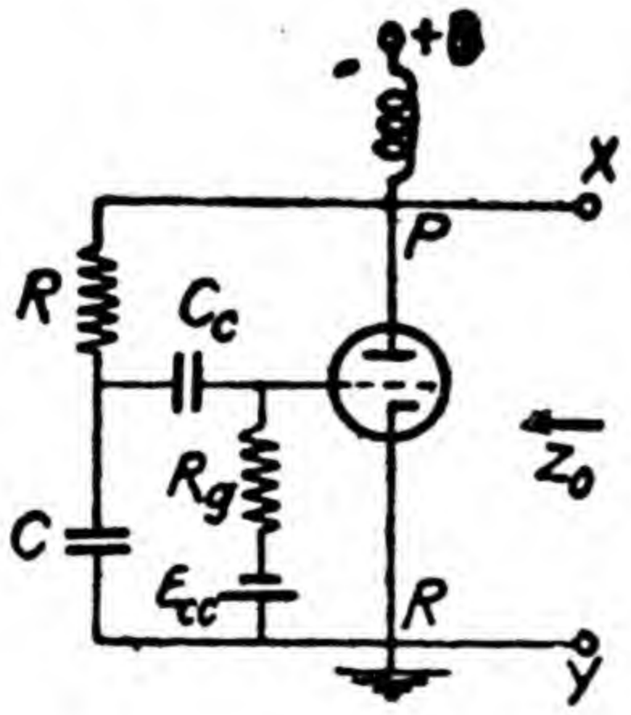


Fig. 22. (*a*) Basic circuit of inductive reactance tube.

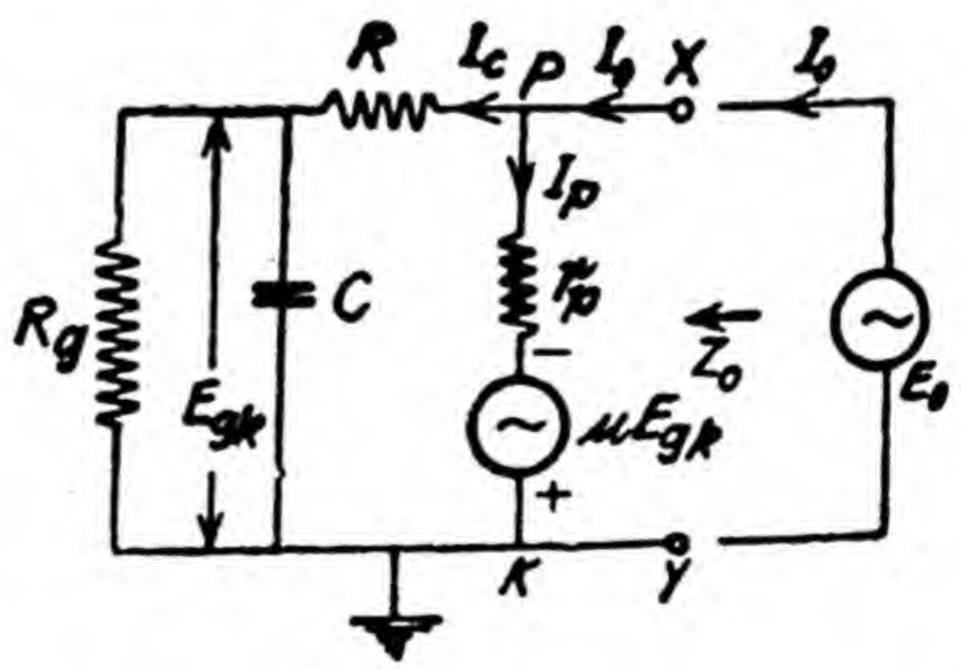


Fig. 22. (*b*) a.c. equivalent circuit.

In inductive reactance, $R \geqslant \frac{1}{\omega C}$ so that the current I_c that flows through R and C is in phase with the applied voltage. Since $E_{gk}=\frac{I_c}{j\omega C}$, it will lag behind the current I_c and the applied a.c. voltage by 90°. The plate current I_p being in phase with E_{gk} also lags behind the applied voltage by 90° which happens in the case

of a circuit having inductive reactance. The reactance tube thus behaves as an inductor.

We shall now proceed to calculate L_{eq}. offered by this reactance tube. Suppose I_0 is the current drawn by reactance tube circuit from ideal voltage source of voltage E_0 applied between plate and cathode and the impedance offered by the tube to E_0 is, say, Z_0. We find from the equivalent circuit that

$$\mathbf{I}_c = \frac{\mathbf{E}_0}{R + \dfrac{R_g}{1 + j\omega C R_g}} = \frac{\mathbf{E}_0 (1 + j\omega C R_g)}{R_g + R (1 + j\omega C R_g)}.$$

Then

$$\mathbf{E}_{gk} = \mathbf{I}_c . \frac{R_g}{1 + j\omega C R_g}$$

$$= \frac{R_g}{1 + j\omega C R_g} \times \frac{\mathbf{E}_0 (1 + j\omega C R_g)}{R_g + R (1 + j\omega C R_g)}$$

$$= \frac{R_g \mathbf{E}_0}{R_g + R(1 + j\omega C R_g)}.$$

Also

$$\mathbf{I}_p = \frac{\mathbf{E}_0 + \mu \mathbf{E}_{gk}}{r_p}$$

$$= \frac{\mathbf{E}_0}{r_p} + g_m . \frac{R_g \mathbf{E}_0}{R_g + R (1 + j\omega C R_g)}$$

Then

$$\mathbf{I}_0 = \mathbf{I}_p + \mathbf{I}_c$$

$$= \frac{\mathbf{E}_0 (1 + j\omega C R_g)}{R_g + R (1 + j\omega C R_g)} + \frac{\mathbf{E}_0}{r_p} + g_m . \frac{R_g \mathbf{E}_0}{R_g + R (1 + j\omega C R_g)}$$

Output admittance between XY is given by

$$\mathbf{Y}_0 = \frac{\mathbf{I}_0}{\mathbf{E}_0} = \frac{1 + j\omega C R_g}{R_g + R (1 + j\omega C R_g)} + \frac{1}{r_p} + \frac{g_m R_g}{R_g + R (1 + j\omega C R_g)}$$

$$= \frac{1}{R + \dfrac{R_g / j\omega C}{R_g + \dfrac{1}{j\omega C}}} + \frac{1}{r_p} + \frac{1}{\dfrac{1}{g_m} + \dfrac{R}{g_m R_g} + \dfrac{j\omega C R}{g_m}}$$

The circuit may be chosen such that first two terms in above equation are smaller compared to third one. Then

$$\mathbf{Y}_0 = \frac{1}{\dfrac{1}{g_m} + \dfrac{R}{g_m R_g} + \dfrac{j\omega C R}{g_m}}$$

Further impedance

$$\left(\frac{1}{g_m} + \frac{R}{g_m R_g} \right)$$

can be made much smaller compared to inductive reactance $j\omega CR/g_m$, that

$$\mathbf{Y}_0=\frac{1}{j\omega CR/g_m}=\frac{1}{j\omega L_{eq}},$$

where

$$L_{eq}=\frac{CR}{g_m}$$

which was desired with

$$g_m=g_{m0}\left(1-\frac{E_{cc}}{E_{c0}}-\frac{E_m}{E_{c0}}\cos\omega_m t\right)$$

which we have shown in the case of capacitive reactance tube.

The frequency of the inductive reactance tube modulated oscillator can be calculated by putting the value of L_{eq} in

$$f=\frac{1}{2\pi}\sqrt{\left(\frac{1}{C_0L'}\right)},$$

where

$$\frac{1}{L'}=\frac{1}{L_{eq}}+\frac{1}{L_0}$$

Thus

$$f=\frac{1}{2\pi}\sqrt{\left\{\frac{1}{C_0}\left(\frac{1}{L_0}+\frac{1}{L_{eq}}\right)\right\}}$$

$$=\frac{1}{2\pi}\sqrt{\left\{\frac{1}{C_0}\left(\frac{1}{L_0}+\frac{g_m}{CR}\right)\right\}}$$

$$=\frac{1}{2\pi}\sqrt{\left\{\frac{1}{C_0}\left[\frac{1}{L_0}+\frac{g_{m0}}{CR}\left(1-\frac{E_{cc}}{E_{c0}}-\frac{E_m}{E_{c0}}\cos\omega_m t\right)\right]\right\}}$$

The carrier frequency, in the absence of modulating voltage ($E_m=0$) is

$$f_o=\frac{1}{2\pi}\sqrt{\left\{\frac{1}{C_0}\left[\frac{1}{C_0}+\frac{g_{m0}}{CR}\left(1-\frac{E_{cc}}{E_{c0}}\right)\right]\right\}}$$

Then

$$\frac{f}{f_c}=\left[\frac{\dfrac{1}{L_0}+\dfrac{g_{m0}}{CR}\left(1-\dfrac{E_{cc}}{E_{c0}}-\dfrac{E_m}{E_{c0}}\cos\omega_m t\right)}{\dfrac{1}{L_0}+\dfrac{g_{m0}}{CR}\left(1-\dfrac{E_{cc}}{E_{c0}}\right)}\right]^{1/2}$$

$$=\left[1-\frac{\dfrac{g_{m0}}{CRE_{c0}}E_m\cos\omega_m t}{\dfrac{1}{L_0}+\dfrac{g_{m0}}{CR}\left(1-\dfrac{E_{cc}}{E_{c0}}\right)}\right]^{1/2}$$

$$=\left[1-\frac{E_m\cos\omega_m t}{CRE_{c0}/L_0g_{m0}+(E_{c0}-E_{cc})}\right]^{1/2}$$

$$=1-\frac{E_m\cos\omega_m t}{2\left[\dfrac{CRE_{c0}}{L_0g_{m0}}+(E_{c0}-E_{cc})\right]}$$

or $$f=f_o\ (1+m_f \cos \omega_m t)$$

where

$$m_f=-\tfrac{1}{2}\ \frac{E_m}{\frac{CRE_{c0}}{L_0 g_{m0}}+(E_{c0}-E_{cc})}$$

The above expression is for frequency modulated by the signal.

30·7. COMPARISON OF FREQUENCY AND AMPLITUDE MODULATION :

Frequency modulation	Amplitude modulation
1. There is a substantial reduction in interference effects That is, noise can be easily minimised in f.m. systems. The presence of the noise at the receiver input in addition to the desired signal tends to alter the input voltage both in instantaneous amplitude and instantaneous phase. In frequency modulation system, amplitude variations can be eliminated by slicing them with the help of limiter, so that always constant amplitude is applied to the discriminator.	In this type of modulation, alteration of the amplitude of the desired signal by the noise at the receiver input severely affects the response and amounts to marked distortion.
2. No restriction is placed on the modulation index. The instantaneous frequency deviation is proportional to the instantaneous magnitude of the modulating signal.	In amplitude modulation, use of an excessively large modulating signal may result in distortion because of over modulation. Modulation index can have a maximum value unity for distortionless modulation.
3. The average power in frequency modulated wave is the same as that contained in the unmodulated wave (carrier).	The average power in modulated wave is greater than that contained in unmodulated wave (carrier). This added power is provided by the modulating source.

30·8. PHASE MODULATION :

Definition. In the process of phase modulation the instantaneous phase of the carrier is varied by an amount that is proportional to the instantaneous amplitude of the signal and at a rate that is proportional to modulating frequency. The amplitude of carrier remains unaltered.

Analysis : The modulating signal has the form

$$e_m=E_m \sin \omega_m t \qquad \ldots(1)$$

The instantaneous phase of the wave is given by the expression

$$\theta=\theta_0+K_p E_m \sin \omega_m t, \qquad \ldots(2)$$

where proportionality factor K_p determines the maximum variation in phase for a given signal strength.

By definition the instantaneous phase for phase modulated wave is given by

$$\phi(t)=\omega_o t+\theta_0+K_pE_m \sin \omega_m t \quad ...(3)$$

Therefore the expression for phase modulated wave can be written as

$$e=E_c \sin (\omega_o t+K_pE_m \sin \omega_m t), \quad ...(4)$$

θ_0 representing the constant phase, which plays no part in the modulation process, is taken zero. The above expression can be written as

$$e=E_c \sin (\omega_c t+\Delta\phi \sin \omega_m t), \quad ...(5)$$

where $$\Delta\phi=K_pE_m, \quad ...(6)$$

and is called phase deviation If the instantaneous phase of a phase modulated wave is given by equation (3), then the instantaneous frequency of the phase modulated wave will be

$$\omega=\frac{d\phi}{dt}=\frac{d}{dt}(\omega_o t+\theta_0+\Delta\phi \sin \omega_m t)$$

$$=\omega_c+\Delta\phi\ \omega_m \cos \omega_m t$$

or $$f=f_o+\Delta f_m \cos \omega_m t$$

$$=f_o+\Delta f \cos \omega_m t, \quad ...(7)$$

where $$\Delta f=\Delta\phi . f_m, \quad ...(8)$$

and is called maximum frequency shift for a phase modulated wave. Obviously variation of the rate of change ϕ with time will produce a change in Δf and consequently in the instantaneous frequency 'f' of the phase modulated wave. This means phase modulation, like frequency modulation, is also a frequency variation process. From equation (8), we can also conclude that if we double either the amplitude E_m or the frequency f_m of the signal, then maximum frequency shift Δf will be doubled. This is in constrast to frequency modulation process in which

$$\Delta f=K_f \frac{E_m}{2\pi} \text{ (eq. 7 art. 30·5)} \quad ...(9)$$

and thus depends only on signal amplitude and not on signal frequency.

We have shown above that the expression for phase modulated wave is

$$e=E_c \sin (\omega_c t+\Delta\phi \sin \omega_m t)$$

$$=E_cJ_0(\Delta\phi) \sin \omega_c t+E_cJ_1(\Delta\phi)[\sin (\omega_c+\omega_m)\ t-\sin (\omega_o-\omega_m]t]$$

$$+E_cJ_2(\Delta\phi)[\sin (\omega_o+2\omega_m)\ t-\sin (\omega_c-2\omega_m)\ t]+\cdots$$

For the particular case $\Delta\phi=0{\cdot}5$, the values of $J_n(\Delta\phi)$ can be determined from the curves of fig. 15. These are

$$J_0\ (0{\cdot}5)=0{\cdot}938,$$
$$J_1\ (0{\cdot}5)=0{\cdot}242,$$
$$J_2\ (0{\cdot}5)=0{\cdot}030$$

and $$J_n\ (0{\cdot}5)\approx 0 \text{ for } n > 2.$$

The equation of the phase modulated wave then becomes

$$e=0{\cdot}938\ E_o \sin \omega_c t+0{\cdot}242\ E_o\ [\sin\ (\omega_o+\omega_m)\ t-\sin\ (\omega_c-\omega_m)\ t]$$
$$+0{\cdot}030\ E_o\ [\sin\ (\omega_c+2\omega_m)\ t-\sin\ (\omega_c-2\omega_m)\ t]$$
$$\approx 0{\cdot}938\ E_c \sin \omega_c t+0{\cdot}242\ E_o\ [\sin\ (\omega_c+\omega_m)\ t-\sin\ (\omega_c-\omega_m)\ t],$$

where the second side band components which have amplitude of 0 03 E_c have been neglected. Approximation is more valid when $\Delta\phi < 0{\cdot}5$. Equation can be further put in the form

$$e=0{\cdot}938\ E_c\left[\sin \omega_c t+\frac{\Delta\phi}{2}\sin\ (\omega_o+\omega_m)\ t-\frac{\Delta\phi}{2}\sin\ (\omega_c-\omega_m)\ t\right].$$

The corresponding equation for a.m. wave is

$$e=E_o\left[\sin \omega_c t+\frac{m_a}{2}\cos\ (\omega_c+\omega_m)\ t+\frac{m_a}{2}\cos\ (\omega_c-\omega_m)\ t\right].$$

It is obvious, on comparing the two equations, that both are identical if $\Delta\phi = m_a$ with only essential difference of the relative phase of the carrier and the sidebands. This difference can be eliminated if the sidebands of amplitude modulated wave are shifted in phase by 90° with respect to the carrier.

Example. A 7·5 kc/s modulating signal has sufficient amplitude to cause a phase shift of 50°. Find the amount of indirect *FM* produced.

$$\Delta f=\Delta\phi.\ f_m$$
$$=\frac{50\pi}{150}\times 7{\cdot}5$$
$$=6{\cdot}55 \text{ kc/s.}$$

Comparison with frequency modulated wave :

If we compare equation (9) art. 30·5, and equation (5), then it becomes obvious that $\Delta\phi$ determines the characteristics of phase modulated wave in the same manner as the factor m_f determines characteristics of frequency modulated wave. If $\Delta\phi$ be made numerically equal to the factor m_f, then the two resulting waves, *i.e.*, one phase modulated wave due to phase modulation process and another frequency modulated wave due to frequency modulation process will be identical.

Since $\Delta\phi$ plays the role of m_f in the case of phase modulated wave, the expression for bandwidth can be written as

$$\frac{BW}{2\Delta f}=\frac{n}{\Delta\phi}. \qquad \ldots(10)$$

If the frequency of the signal f_m is doubled, it will not affect $\Delta\phi$ (see eq. 6) and therefore, unlike to frequency modulation case, there will be no affect on n, the number of significant side frequency pairs. But the value of Δf is doubled so that

$$2\frac{BW}{(2\Delta f)}=\frac{n}{\Delta\phi}$$

or

$$\frac{BW}{2\Delta f}=2.\frac{n}{\Delta\phi} \qquad ...(11)$$

or in other words bandwith covered by more widely spaced components is doubled. This is in constrast to the case of frequency modulation where bandwidth remains constant even after doubling the signal frequency f_m.

Production of f.m. wave by phase modulation :

The undesirable variation of bandwidth with modulating frequency, as stated above, can be corrected if E_m is made to vary inversely as the modulating signal frequency f_m For this purpose, an integrating circuit can be used. If the input to this circuit is $e_i=E_m \cos \omega_m t$, then output will be

$$e_0=\frac{1}{RC}\int E_m \cos \omega_m t \, dt$$

$$=\frac{1}{RC}.\frac{E_m \sin \omega_m t}{\omega_m}$$

or

$$|\, e_0 \,|=\frac{E_m}{RC\omega_m},$$

so that for constant E_m, the amplitude of output wave, $|\, e_0 \,|$, varies inversely as ω_m. If this output $|\, e_0 \,|$ is applied to the modulator, then frequency deviation Δf will not vary with f_m as shown below. For this, output voltage, E_m, should be replaced by E_m/RCf_m. That is

$$\Delta\phi=K_p\frac{E_m}{RCf_m},$$

so that

$$\Delta f=\Delta\phi\, f_m$$

$$=K_p\frac{E_m}{RCf_m}.\, f_m$$

$$=K_p\frac{E_m}{RC} \qquad ...(12)$$

This means by doubling f_m, Δf as well as $\Delta\phi$ will remain unaffected. Consequently, bandwidth remains constant which happens in the case of frequency modulated wave. In other words, we can state that in this process in which, prior to applying the modulating voltage to modulator, an integrating circuit is employed to vary E_m inversely as f_m, a frequency modulated wave is produced by the process of the phase modulation.

30·9. PULSE MODULATION SYSTEMS :

We have discussed already the modulation of carrier frequency by sinusoidal signals. A signal waveform can be produced by taking samples of this wave at regular intervals. The sampling of information signal is done by the use of pulse on varying one of their basic characteristics such as amplitude, width (duration), or

position inaccordance with this modulating signal (Fig. 23). These sampled pulses, bearing the characteristics of modulating signal, are then modulated upon the radio frequency carrier and thus short bursts of the modulated carrier are transmitted. We shall discuss pulse modulation methods in brief.

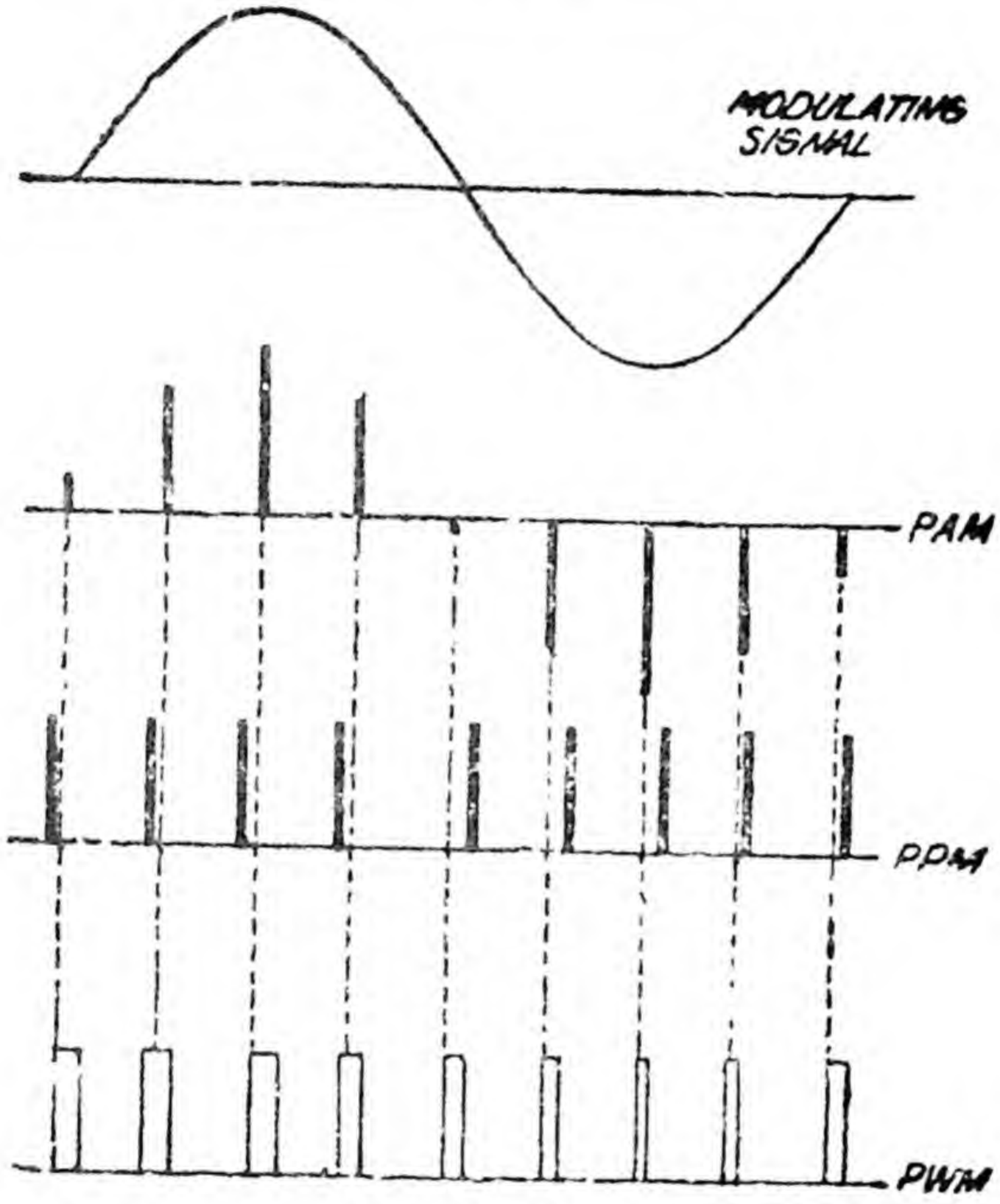


Fig. 23. Various pulse modulation methods.

30·9-1. PULSE AMPLITUDE MODULATION (PAM) :

In this process, relative amplitudes of successive pulses serve as a measure of signal amplitudes at the corresponding sampling instants. Width and position of the pulse remain constant (fix). Modulating signal is sampled at the basic rate, usually $\frac{1}{2f_m}$ where f_m is the maximum value of frequency present in the modulating signal.

As shown in figure 24 these amplitude modulated pulses modulate the radio frequency carrier. The amplitude modulated radio frequency output is then transmitted in the form of short bursts. The *duration of each burst is equal to the duration of pulses* in figure 24 (*a*). *Instantaneous amplitude of each radio frequency burst is proportional to the audio frequency signal* shown in fig. 24 (*c*).

The main advantages of pulse amplitude modulation are :

(*a*) greater peak power, compared to previous amplitude modulation, is obtained, and

(*b*) signal to noise ratio of the system is improved.

30·9-2. PULSE DURATION OR WIDTH MODULATION (I·WM) :

The width of the pulse is varied (around fixed value) inaccordance with the modulating signal whereas amplitude and position remain fixed.

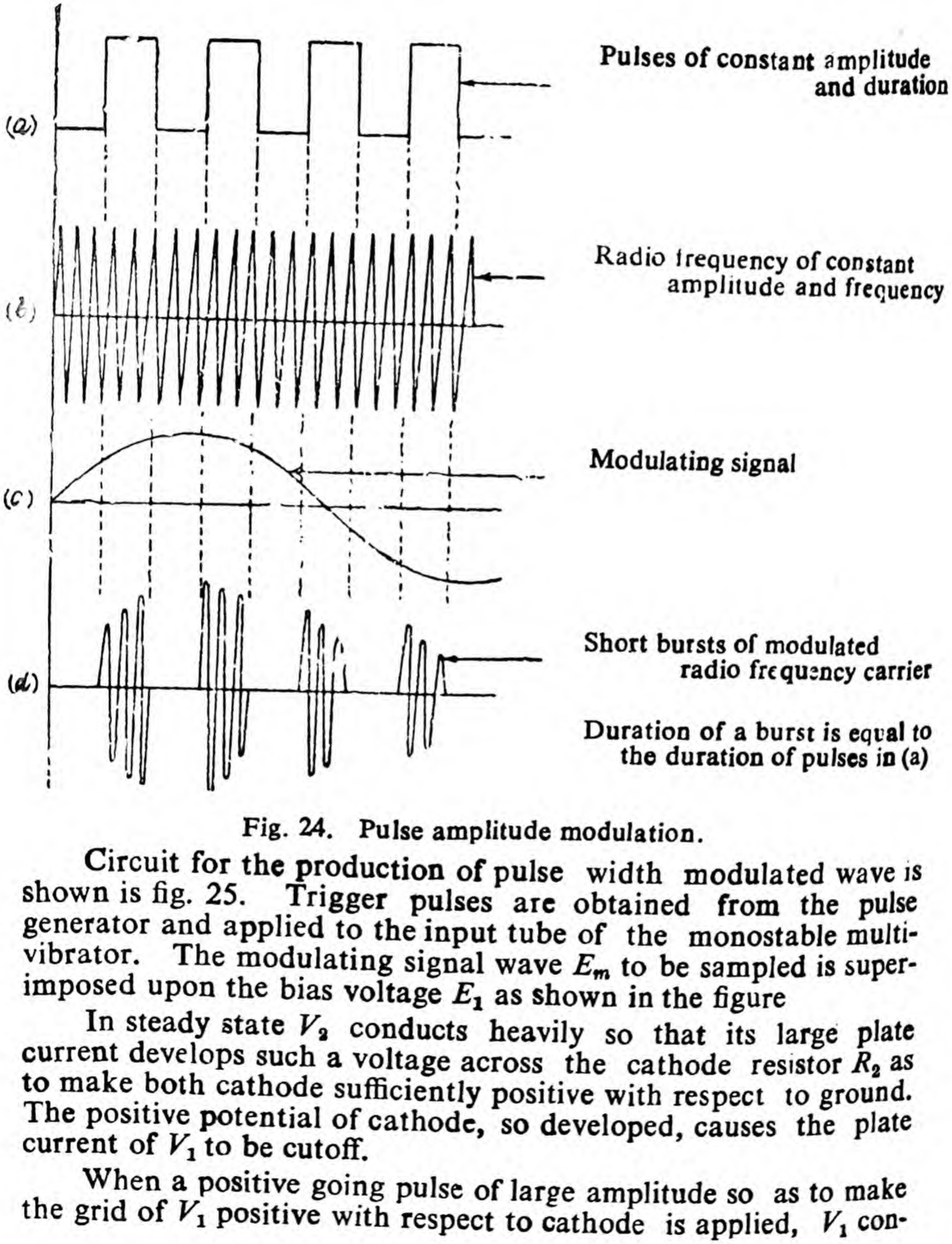


Fig. 24. Pulse amplitude modulation.

Circuit for the production of pulse width modulated wave is shown is fig. 25. Trigger pulses are obtained from the pulse generator and applied to the input tube of the monostable multivibrator. The modulating signal wave E_m to be sampled is superimposed upon the bias voltage E_1 as shown in the figure

In steady state V_2 conducts heavily so that its large plate current develops such a voltage across the cathode resistor R_2 as to make both cathode sufficiently positive with respect to ground. The positive potential of cathode, so developed, causes the plate current of V_1 to be cutoff.

When a positive going pulse of large amplitude so as to make the grid of V_1 positive with respect to cathode is applied, V_1 con-

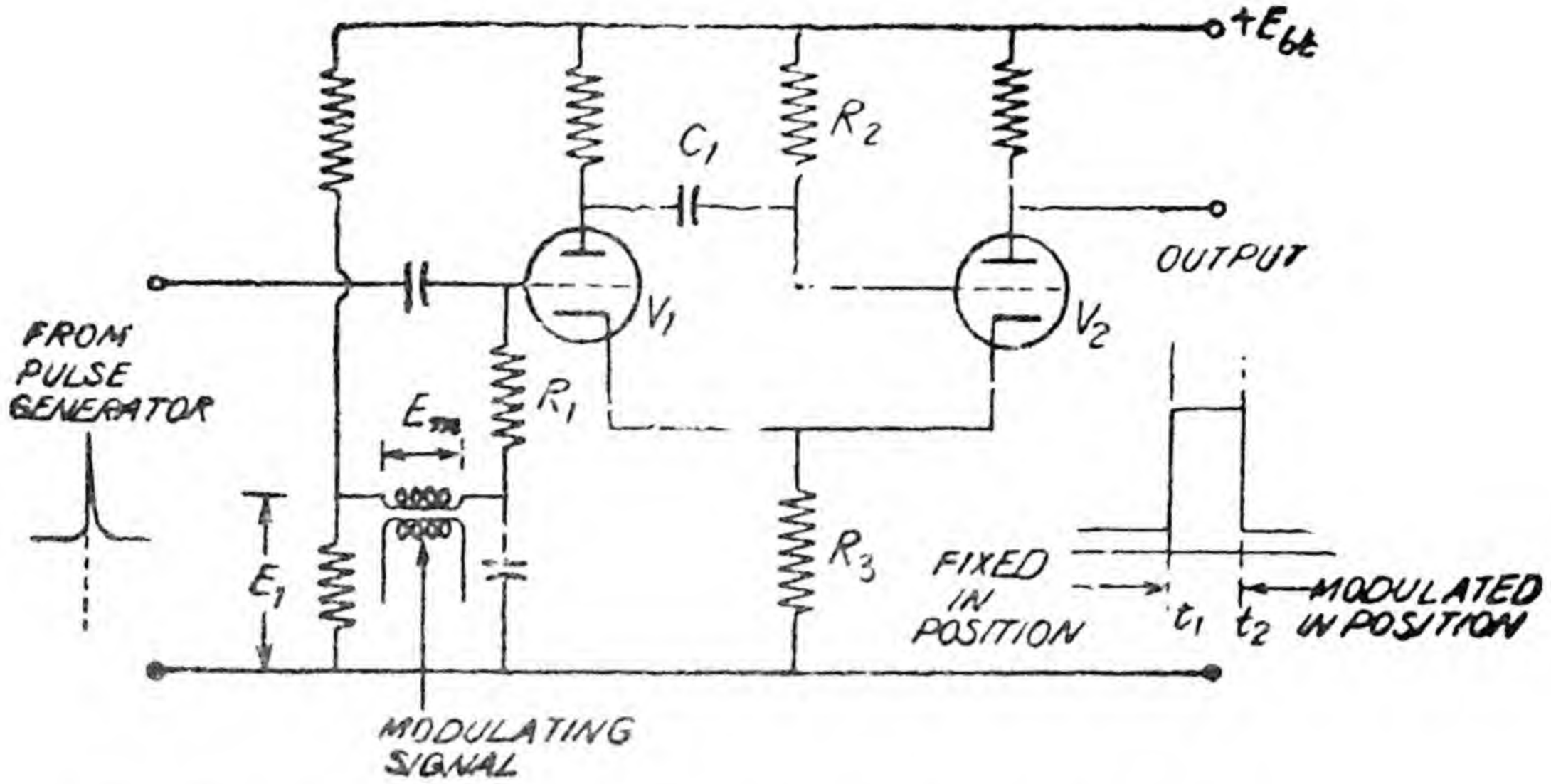


Fig 25. *PWM* by cathode coupled one shot multivibrator.

ducts. Plate potential of V_1 then falls and this decrease in the plate potential is directly applied (as the charge on condenser C_1 cannot change instantaneously) to the grid of V_2 making it negative. This reduces the plate current of V_2, making the cathodes less positive with respect to ground. Then V_1 conducts more with the result that plate potintial of V_1 decreases further and grid of V_2 becomes more negative. Action continues and finally V_2 is cut-off while V_1 remain in conduction. Now capacitor C_1 discharges through R_2. When C_1 is discharged enough to bring the grid of V_2 to a potential level where V_2 starts conducting, steady states are reverted *i.e.* V_1 is cutoff and V_2 continues to conduct.

Obviously, more negative is the grid of V_2, greater will be the time for which V_1 conducts or V_2 will take more time to start conducting. The duration (t_2-t_1) for which V_1 corducts depends upon :

(*i*) the time constant R_2C_1,

(ii) the grid bias voltage of V_1.

The grid bias of V_1 governs its plate current. The plate current in its turn, sets the cathode potential during this interval of conduction (t_2-t_1) and consequently determines the level to which the grid potential of V_2 should rise before V_2 can conduct. *i.e.*, determines the instant t_2. Since instant t_1 is fixed, interval (t_2-t_1) varies according to the grid bias voltage. Therefore in the circuit (fig. 25), we note that the modulating signal is superimposed upon grid bias and hence grid bias varies inaccordance with modulating signal which will inturn determine the interval (t_2-t_1) and hence the pulse width. Thus the output is pulse width modulated.

Transistor Circuit : An emitter coupled multivibrator generating pulse-width modulation is shown in fig. 26. It may be noted

that

(*i*) the *starting time* of pulses from a monostable multivibrator is controlled by applying trigger pulses, and

(ii) the signal, to be sampled, controls the *duration* of these pulses from the monostable multivibrator.

(iii) the gatewidth of this multivibrator depends on the voltage to which the capacitor, *C*, is charged and therefore if this voltage is varied inaccordance with signal voltage, rectangular pulses of varying width will be generated at output.

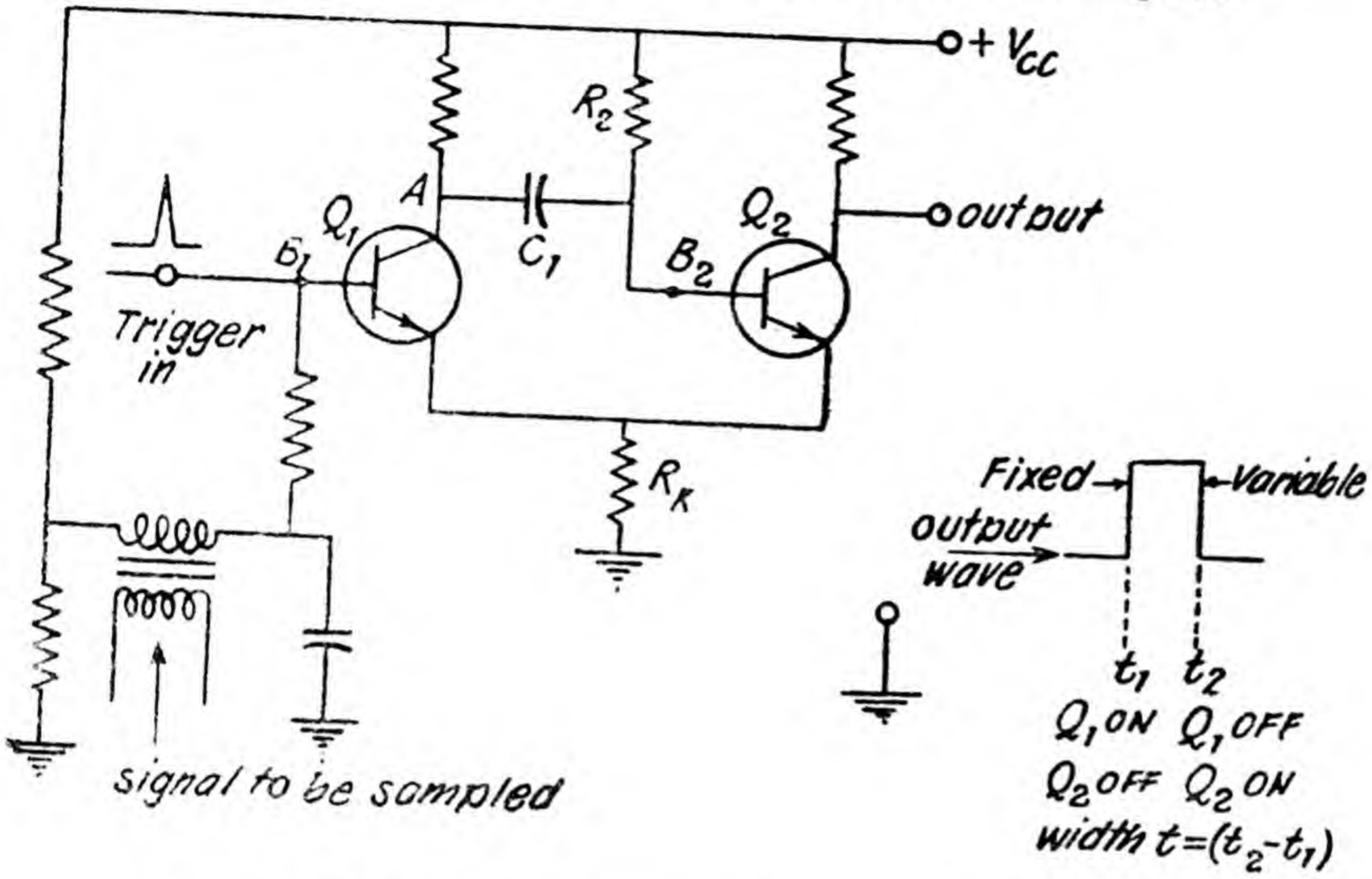


Fig. 26. Emitter coupled monostable multivibrator.

Operation :

For this type of multivibrator stable state is with Q_1 OFF and Q_2 ON. When trigger pulse (a positive going pluse of large amplitude) is applied at the base of Q_1, then it starts conducting. As it draws the collector current, voltage at *A* falls. Since any charge on condenser C_1 cannot change instantaneously, the falling voltage at A is directly applied to the base of Q_2. This reduces the plate current of Q_2, making the emitters less positive with respect to ground. Its result is that Q_1 conducts more, making now B_2 more negative. Action continues so that Q_2 is finally cutoff while Q_1 remain in conduction. This is called regenerative action.

With Q_2 OFF, capacitor C_1 begins to charge towards $+V_{cc}$ through R_2. *The rate of charging depends upon the supply voltage and time constant* R_2C_1. Thus when voltage across C_1 builds up to a level so as to make B_2 sufficiently positive, Q_2 starts conducting and Q_1 is cutoff by regenerative action. The time when Q_2 starts conducting depends upon :

(i) the time constant, R_2C_1

(ii) the voltage across the common-emitter resistor, R_k because Q_2 conducts only when the voltage at base B_2 becomes slightly more positive than voltage across R_k.

The voltage across R_k depends upon the current flowing through, which at the time is collector current of Q_1 (which is then ON). The collector current of Q_1, in turn, depends on the base bias voltage, at B_1. This base bias voltage in turn, is governed by the instantaneous changes in applied modulating signal. Thus finally we arrive at the *instant* t_2, at which Q_2 conducts, and, *is determined by the signal to be sampled.* As instant t_1 is fixed, width $(t_2 - t_1)$ of the outpul pulse varies inaccordance with the modulating signal.

30·9-3. PULSE POSITION MODULATION (PPM) :

In this process, the spacing along the time axis between the pulse do not remain uniform. but is varied inaccordance with the modulating signal (position of the pulses is shifted) whereas the amplitude and width remain fixed.

Pulse position modulation can be achieved by differentiating width modulated pulses and then rectifying the differentiated output. After rectification only those pulses which have differentiated trailing edge of width modulated pulses will be obtained as shown in figure 27.

Unmodulated position of trailing edge

(a) Width modulated pulse

(b) Width modulated pulse after differentiation

(c) Position modulated pulse after rectifying the pulse (b).

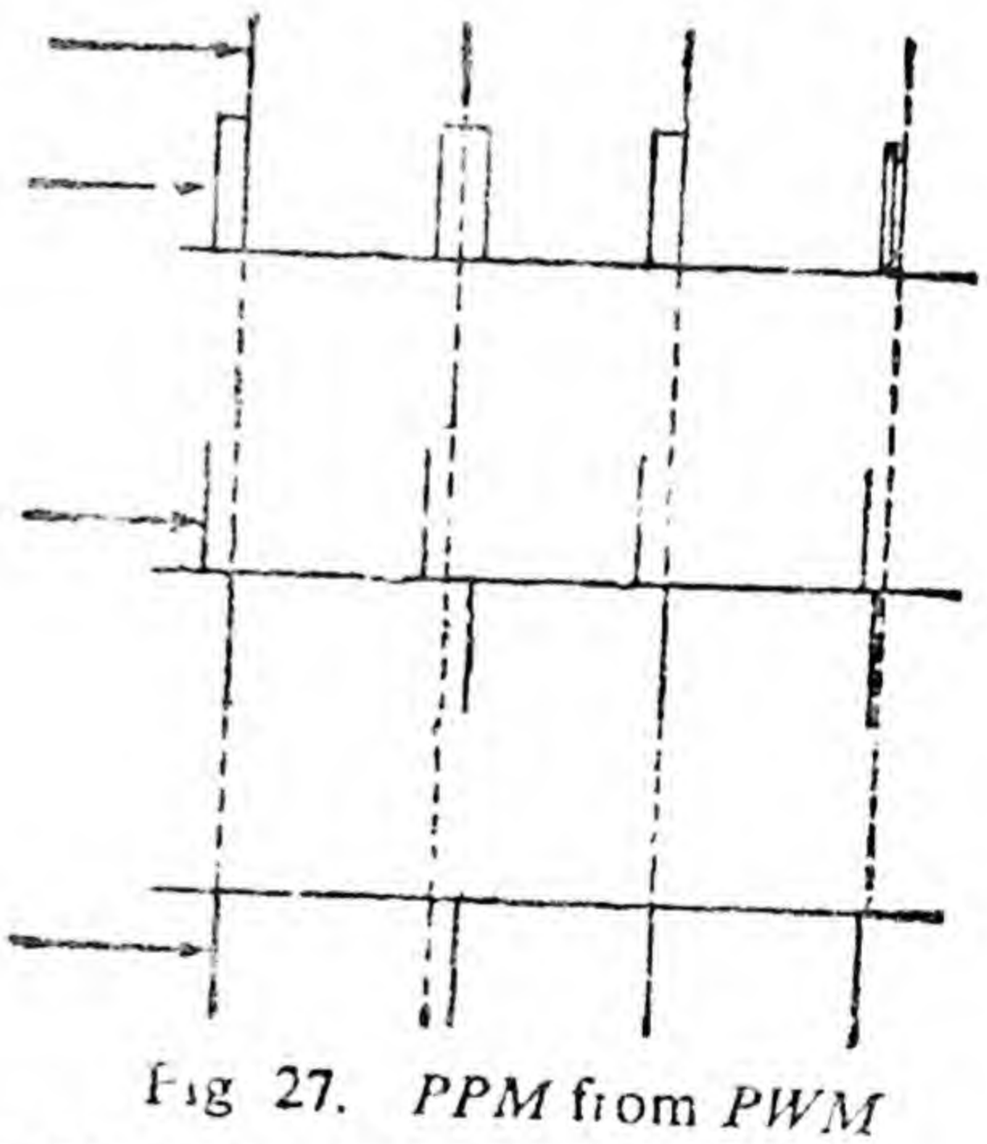


Fig 27. *PPM* from *PWM*

30·9-4. PULSE CODE MODULATION (PCM) :

In pulse code modulation, signal is first sampled, then sampled amplitude is selected interms of nearest standard amplitude. This selected standard amplitude is then indicated by a code of pulses. Thus finally signal is transmitted as code of pulses.

As shown in fig. 28, the signal is sampled at an interval t. This interval is kept small so that a sufficient number of samples can be taken. Though the signal is sampled at regular intervals, but due to sufficient number of samples taken at frequent intervals, nothing useful of the signal to be transmitted is lost.

The actual amplitude of the sampled signal is then not chosen as such for transmission, but it (sampled amplitude) is quantised

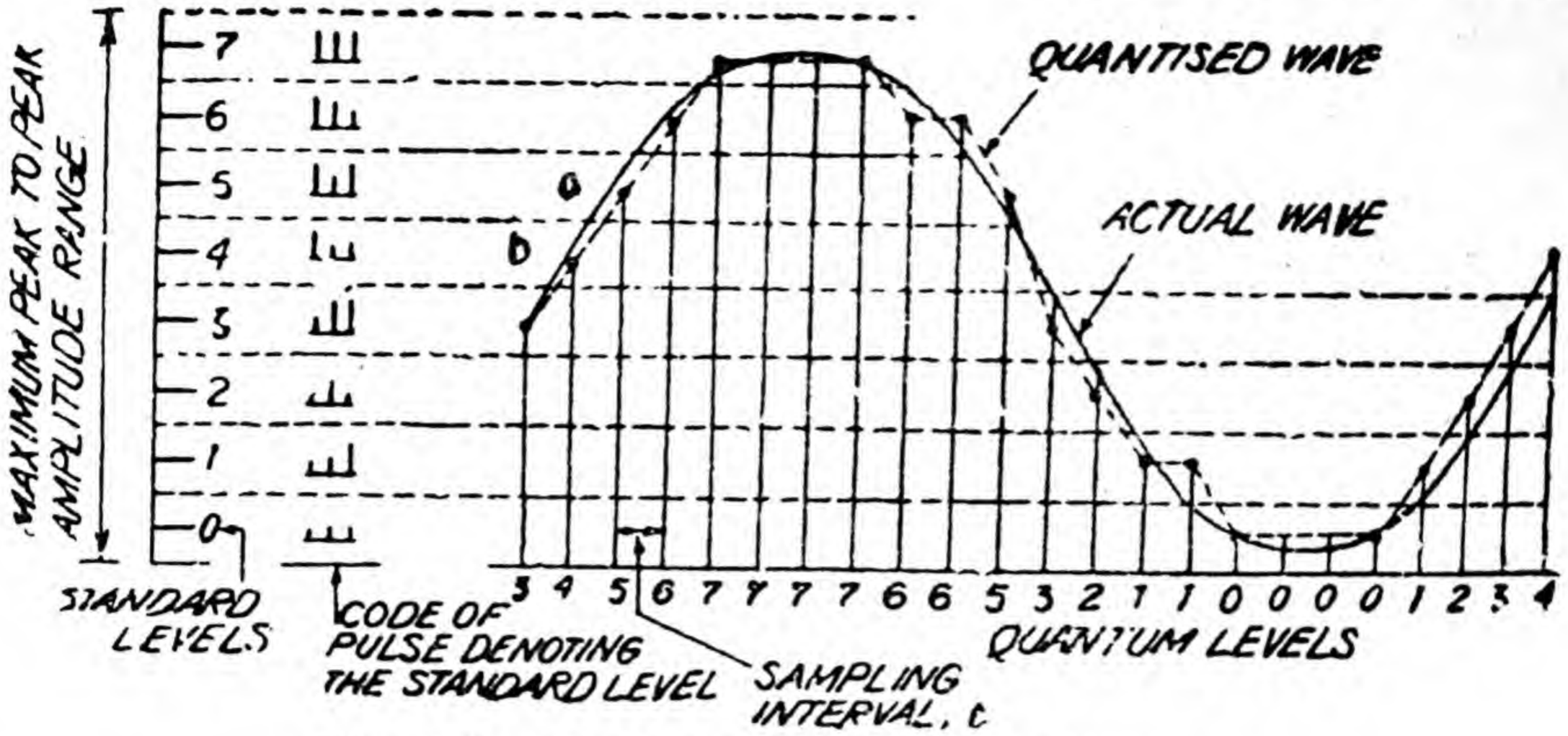


Fig. 28. *PCM* sampling and quantising.

into a nearest standard amplitude. For example, actual amplitude *b* is chosen as a nearest standard amplitude 4; actual amplitude *a* as a nearest standard amplitude 5. This is done to avoid the use of a prohibitively large number of code pulses to indicate the sampled amplitudes of the signal. In this process of quantising the information, some distortion (as is obvious from dotted and solid curves) is introduced, but this error is minimised by increasing the number of standard amplitudes.

These selected standard amplitudes are not transmitted as such but are indicated by a code of pulses. As shown in the figure, each of the standard amplitudes from 0 to 7, can be specified by a different combination of pulses in a three pulse code in which a pulse may be either present or absent, or in other words code mainly concerns with the presence or absence of the pulse and such a code corresponds to binary digital code in which the presence of a pulse is indicated by 1 and absence by 0†. Thus the code for the eight standard amplitudes in three pulse code is

	2^2	2^1	2^0
7 =	1	1	1
6 =	1	1	0
5 =	1	0	1
4 =	1	0	0
3 =	0	1	1
2 =	0	1	0
1 =	0	0	1
0 =	0	0	0

Obviously, seven digit code is capable of transmitting $2^7 = 128$ standard amplitudes. A *PCM* system block diagram is shown in figure 29.

---

†Remember *zero* state of the binary is that state in which output tube conducts while input tube is cut off. This is initial state and no trigger pulse is applied to the first tube. When any pulse of appropriate polarity and amplitude is applied to the tube second, a reversal of state occurs and output tube is cut off while input tube conducts and this is the *one* state of the binary.

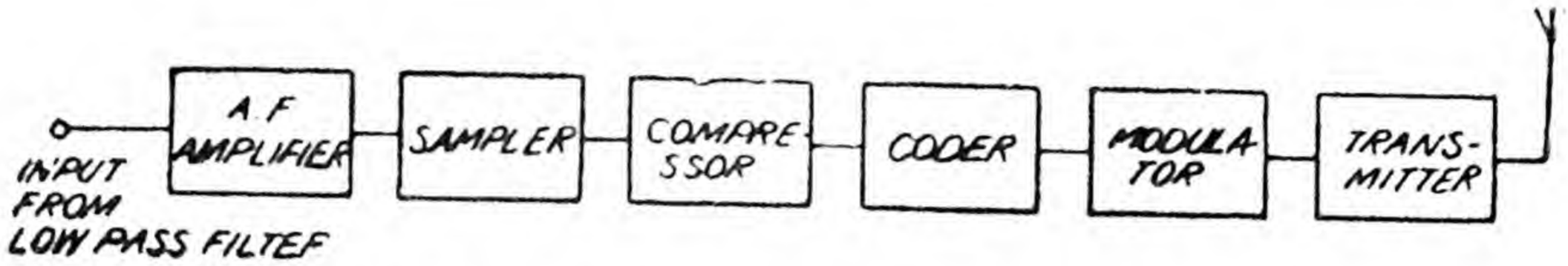


Fig. 29. Block diagram of a PCM system.

30·9-5. ADVANTAGES OF PULSE MODULATION :

(*i*) In such systems, amp itude limiting techniques may be employed to minimise noise effects.

(*ii*) The transmission can be made independent of fading and other difficulties.

(*iii*) Multiple signals may be transmitted on a signal carrier.

(*iv*) Energy is radiated only in short bursts and therefore the consumption of power is low.

EXPERIMENT

To measure the modulation of a carrier wave and to show its relationship with the power in sidebands.

Procedure : An arrangement, shown in figure 30, is used for the measurement of modulation. For the supply of various modulating signal a low frequency oscillator is used and the carrier wave, over which the modulation is performed, is generated by a high frequency oscillator. The modulated output of h. f. o. is connected to the vertical plates of C.R O. The horizontal plates are connected either to linear base or to the terminals of modulating oscillator.

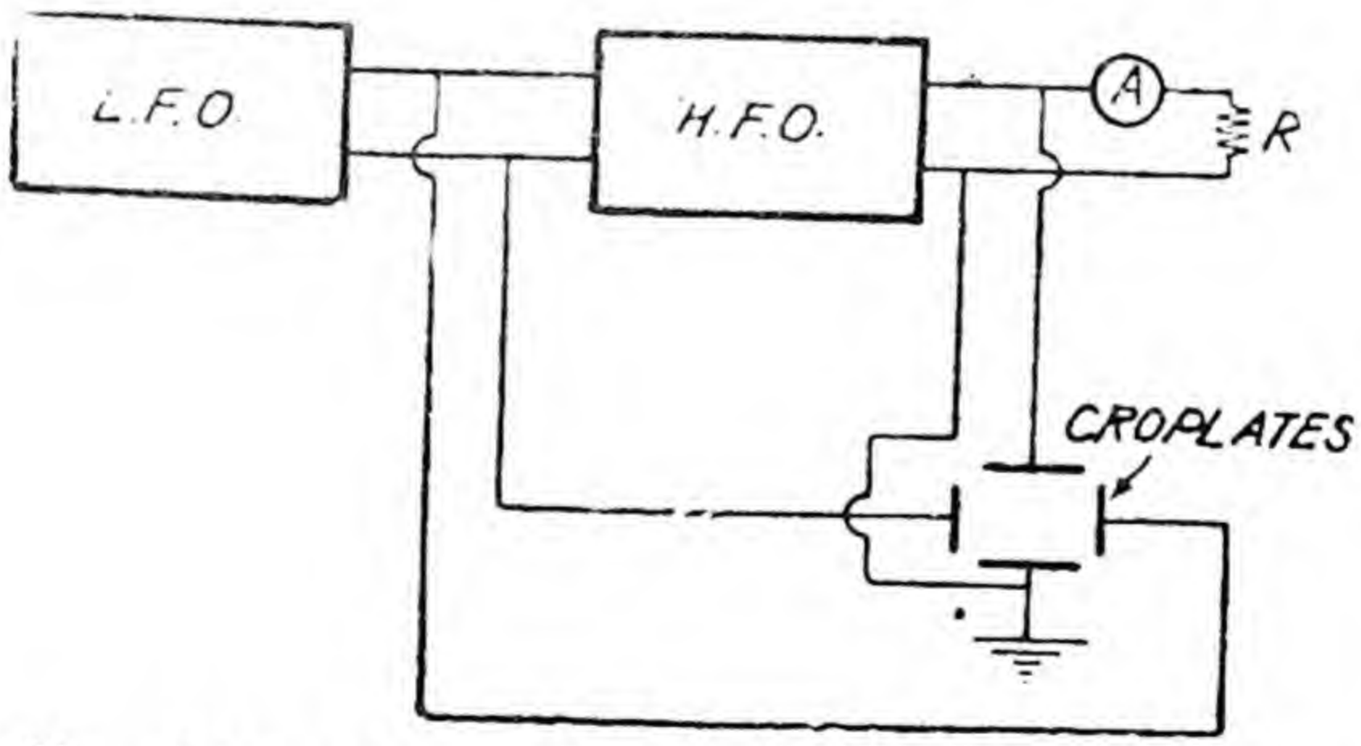


Fig. 30. Arrangement for the measurement of modulation.

For the power measurement, output terminal of h.f.o. are also connected to a known load *R* through a milliameter *A*. We proceed as follows :

(*i*) First modulating oscillator output is kept at zero and output of modulated generator is adjusted to give a suitable reading. The ammeter *A* then records the carrier current *I*. It is noted.

(*ii*) Now the output of modulating oscillator is raised to give 90% modulation approximately.

The ammeter reading is noted. The stationary pattern on the screen of C.R.O. will have the form as shown in fig. 31. The exact value of mod ulation can be calculated from

$$m=\frac{E_{max.}-E_{min.}}{E_{max.}+E_{min.}},$$

where $E_{max.}$ and E_{min}, are the maximum and minimum amplitudes of the modulated wave as shown in fig. 31.

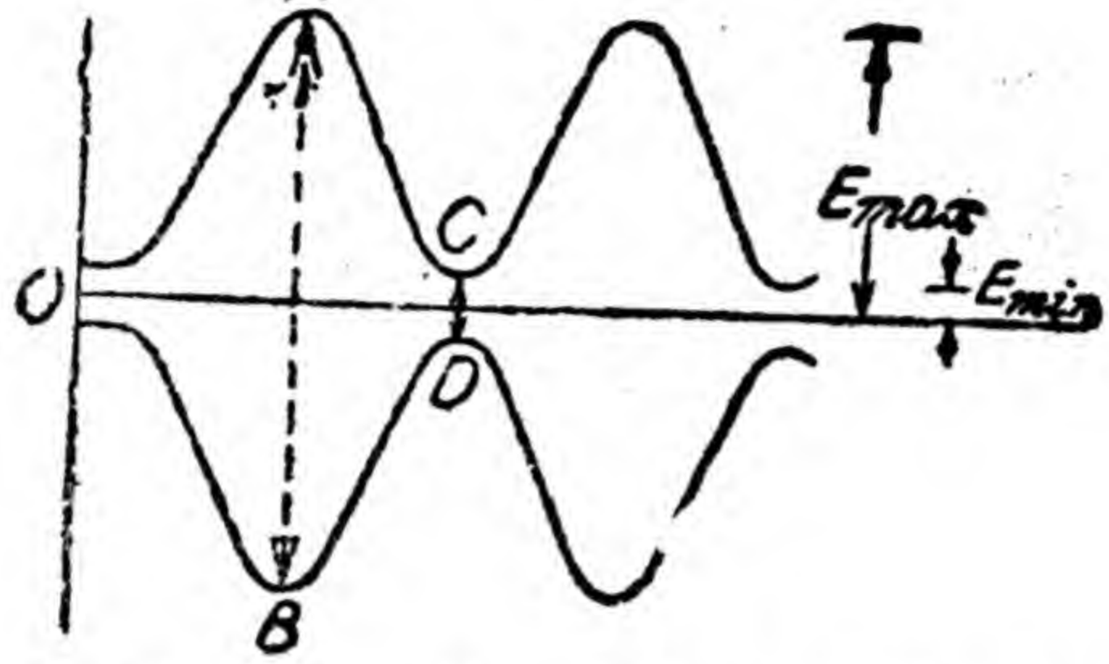


Fig. 31. Stationary pattern of an amplitude modulated wave on C.R.O. screen.

(*iii*) Now by varying the oscillator output the modulation is increased in steps and at each step the values of m and current I are noted.

If P_c is power in the unmodulated carrier given by $P_o=I_c^2.R$, then the power in two side bands

$$P_{sideband}=P-P_c, \qquad \text{where } P=I^2R.$$

Two graphs are plotted. One graph is plotted in P, the total power of the modulated carrier and modulation m as shown in Figure 32. The other graph is plotted in the power of side bands and m^2 as shown in figure 33.

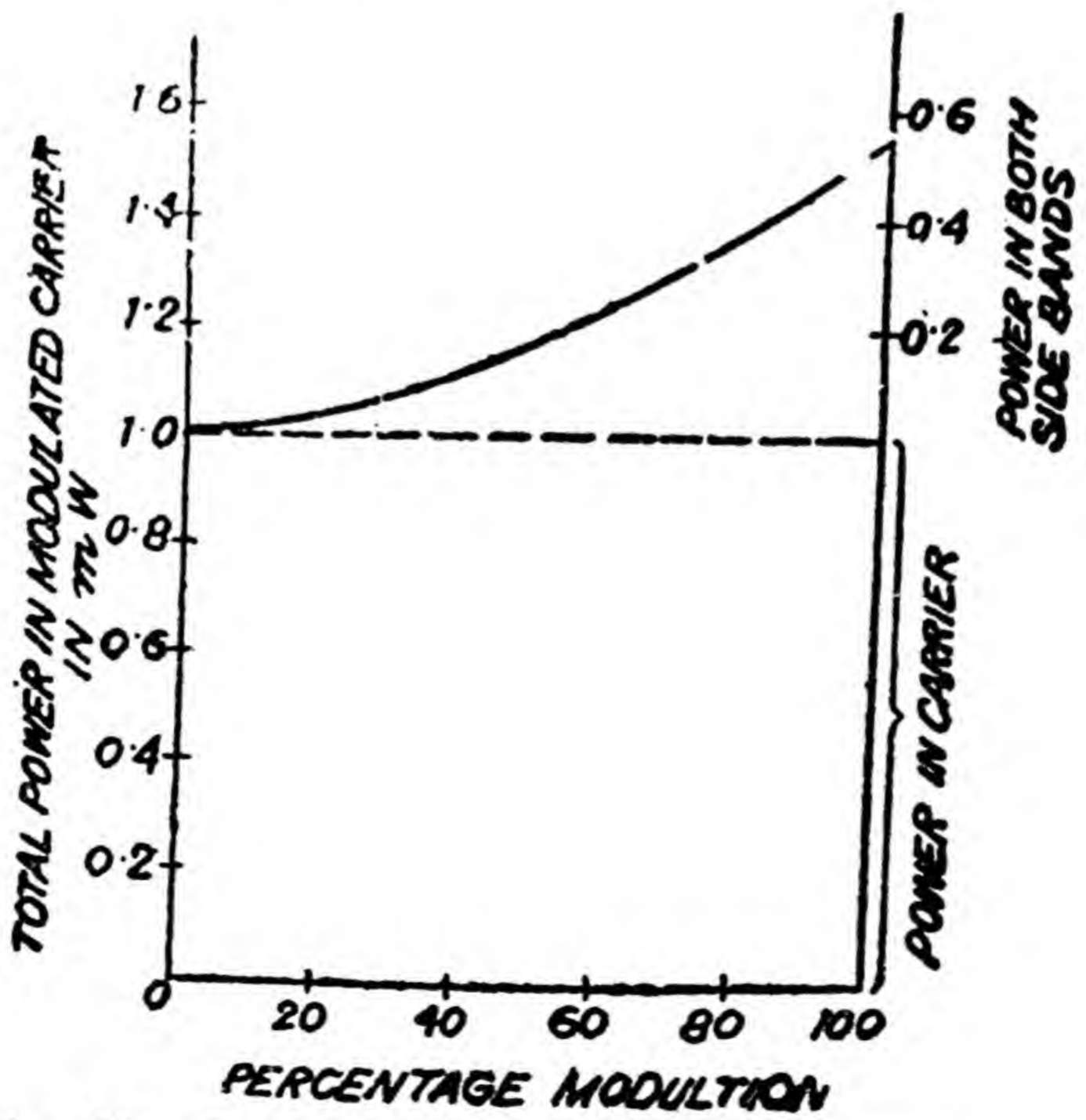


Fig. 32. Total power in modulated carrier.

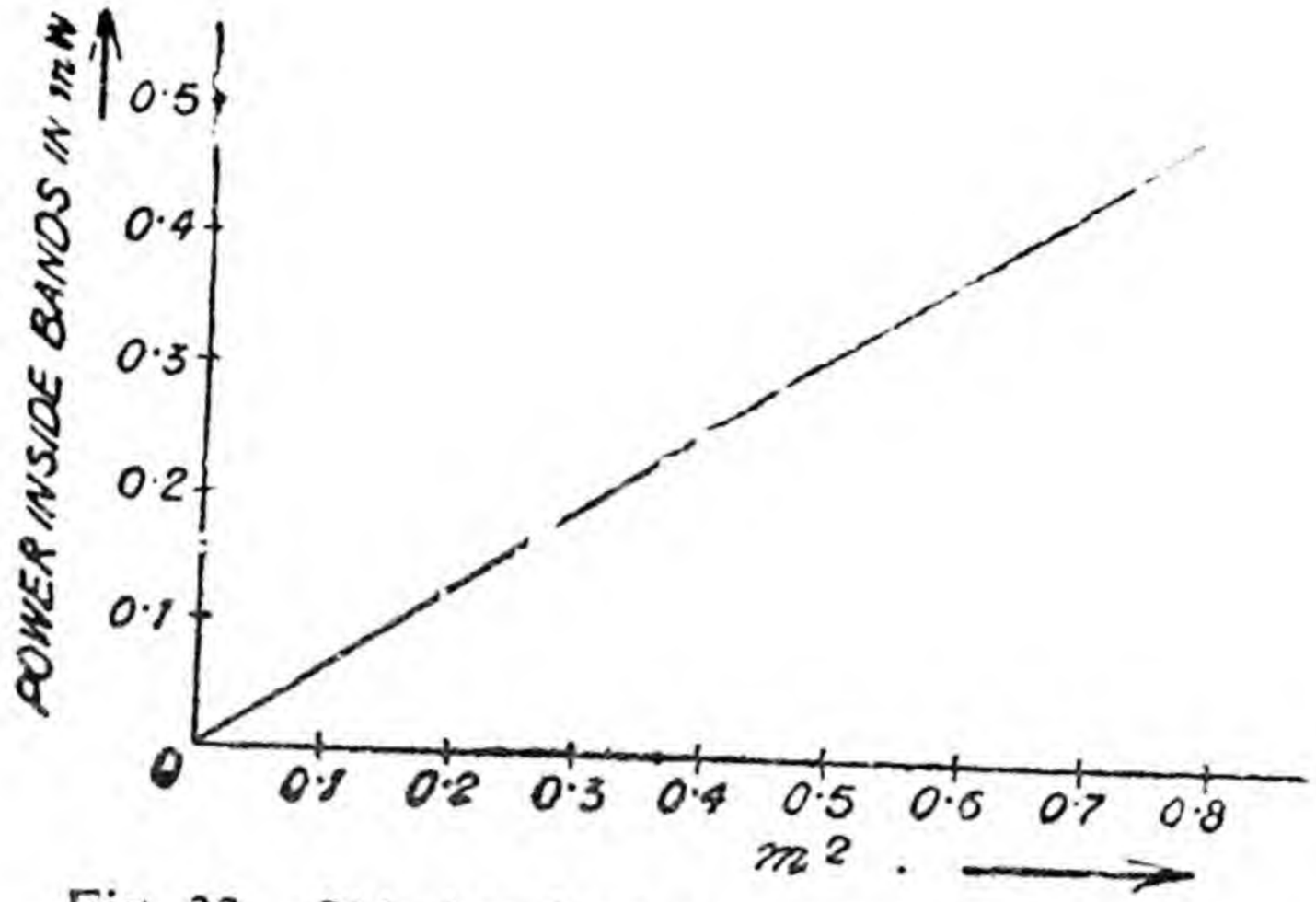


Fig. 33. Side band power as a function of m^2.

NUMERICAL EXAMPLES

1. *If the r.m.s. value of the aerial current before modulation is 12·5 A, and during modulation is 14 A. Calculate the percentage of modulation employed assuming no distortion.*

We can write the relation in the current before and after the modulation as

$$I = I_0 \sqrt{\left(1+\frac{m_a^2}{2}\right)},$$

$$14 = 12{\cdot}5 \sqrt{\left(1+\frac{m_a^2}{2}\right)}.$$

$$\therefore \quad \frac{m_a^2}{2} = \left(\frac{14}{12{\cdot}5}\right) - 1 = \frac{159}{625}$$

or

$$m_a^2 = \frac{318}{625}$$

or

$$m_a^2 = 0{\cdot}7132.$$

Percentage of modulation $= 0{\cdot}713\,2 \times 100$

$$= 71{\cdot}32\%.$$

2. *A wireless transmitter radiates 4KW, with an unmodulated carrier wave and 5·2 KW, with modulated wave. Assuming negligible distortion, calculate percentage of modulation.*

Relation in modulation and carrier wave power can be expressed as

$$P_m = P_o \left(1+\frac{m_a^2}{2}\right),$$

$$5{\cdot}2 = 4\left(1+\frac{m_a^2}{2}\right),$$

$$\frac{m_a^2}{2} = \frac{5{\cdot}2}{4} - 1 = 0{\cdot}3,$$

$$m_a^2 = 0{\cdot}6,$$

$$m_a = 0{\cdot}7746.$$

Percentage of modulation

$$=0{\cdot}7746\times 100$$
$$=\mathbf{77{\cdot}46\%}.$$

3. *What should be the power in an amplitude modulated wave when modulation is 10% and carrier power is 10 watts.*

Carrier power $P_c=E_c^2/2R=10$

and $m_a=1.$

$\therefore$ Power in modulated wave will be given by

$$P_m=10\ (1+\tfrac{1}{2})=\mathbf{15\ watts.}$$

4. *In a plate modulated class C amplifier, the output obtained is 1·5 KW, with an efficiency of 70%. The power input from the source is 1000W. Modulation is 60%. Determine :*

(*a*) *Power supplied by the modulating amplifier,*

(*b*) *Plate dissipation of the amplifier,*

(*c*) *Power supplied to the modulated amplifier.*

Given that

$$P_0=1500\ W,$$
$$\eta=0{\cdot}7,$$
$$P_{bb}=1000\ W,$$
$$m_a=0\ c.$$

(*a*) Power supplied by the modulating amplifier is

$$P_m=P_{bb}\frac{m_a^2}{2}$$
$$=1000\times\frac{26}{200}$$
$$=\mathbf{180\ W.}$$

(*b*) Plate dissipation is given by

$$P_p=P_{bb}\left(1+\frac{m_a^2}{2}\right)(1-\eta)$$
$$=1000\left(1+\frac{0{\cdot}36}{2}\right)\times 0{\cdot}3$$
$$=\mathbf{708\ W.}$$

(*c*) Power supplied to the modulated amplifier

$$P_i=\frac{P_0}{\eta}=\frac{1500}{0{\cdot}7}=\mathbf{2142{\cdot}82\ W.}$$

5. *A sinusoidal carrier voltage of frequency 1 mc/s. and amplitude 100 volts amplitude-modulated by a sinusoidal voltage of frequency 5 kc/s. producing a modulation index of 50%. Calculate the frequency and amplitude of upper and lower side band terms.*

Frequency of upper side band

$$=\omega_0+\omega_m$$
$$=1.10^6+5.10^3$$
$$=\mathbf{1005\ kc/s.}$$

Frequency of lower side band

$$=\omega_c-\omega_m$$
$$=10^6-6.10^3$$
$$=\mathbf{995\ kc/s.}$$

Amplitude of each side band term

$$=\frac{m_a}{2}.E_c$$
$$=\frac{0{\cdot}5\times 100}{2}=\mathbf{25\ volts.}$$

6. *A grid modulated class C amplifier is designed to give 2 kw modulated output power at a modulation depth of 80%. Find the rating of the valve. Assume an unmodulated efficiency of 25%.*

Refer to eq. (24) art. 30·3 (ii), the rating of the valve is

$$(P_p)_{max}=P_c\left(\frac{1}{\eta_{unmod.}}-1\right)$$
$$=P_c\left(\frac{1}{0{\cdot}25}-1\right)=3P_c$$

From eq. (22), we get

$$P_c=\frac{P_0}{1+\frac{m_a^2}{2}}=\frac{2000}{1{\cdot}32}$$
$$=1515{\cdot}1515.$$

Therefore

$$(P_p)_{max.}=4\times 1{,}515{\cdot}15=\mathbf{4545{\cdot}45\ Watts.}$$

7. *In a grid modulated class C amplifier the rating of the r. f. amplifier tube (maximum plate dissipation) is 3000 W and depth of modulation is 50%. Find (a) r. f. power output P_c, (b) percent increase in power contained in sidebands. $P_{sideband}$ and in modulated power output, P_0, due to modulation. If the depth of modulation is changed so that efficiency increases by 6%, then find the new depth of modulation and power dissipated in the valve. Given that unmodulated efficiency is 25%.*

Given $(P_p)_{max.}=3000$ W,

$m_a=0{\cdot}5$,

$\eta_{unmod.}=0{\cdot}25$.

(*a*) r. f. power output

$$P_c=\frac{(P_p)_{max.}}{\frac{1}{\eta_{unmod.}}-1}=\frac{3000}{\frac{1}{0{\cdot}25}-1}=\mathbf{1000\ Watt.}$$

(*b*) $P_{sideband}=P_c\frac{m_a^2}{2}=\frac{1000\times 0{\cdot}25}{2}=150$ Watt.

Total modulated output power is

$$P_0 = P_c\left(1+\frac{m_a^2}{2}\right) = 1000\times\left(1+\frac{0{\cdot}25}{2}\right)$$

$$= 1000\times 1{\cdot}125$$

$$= 1125 \text{ Watt.}$$

Therefore $P_{sideband}$ as percent of total output P_0 is

$$= \frac{P_{sideband}}{P_0}\times 100 = \frac{150}{1125}\times 100$$

$$= 13{\cdot}33\%.$$

Percent increase in total output power is

$$= \frac{P_0}{P_c}\times 100 = \frac{1125-1000}{1000}\times 100$$

$$= \frac{125}{1000}\times 100$$

$$= 12{\cdot}5\%.$$

(*c*) Efficiency at 50% modulation

$$= \eta_{unmod.}\left(1+\frac{m_a^2}{2}\right)$$

$$= 0{\cdot}25\left(1+\frac{0{\cdot}25}{2}\right)$$

$$= 0{\cdot}25\times 1{\cdot}125$$

$$= 0{\cdot}28125$$

$$= 28{\cdot}125\%.$$

Increase is by 6%. Therefore, new efficiency

$$= (28{\cdot}125+6)\%$$

$$= 34{\cdot}152\% = 0{\cdot}34125.$$

Let new modulation index be m_a', then

$$0{\cdot}341_{\cdot}5 = 0{\cdot}25\left(1+\frac{m_a'^2}{2}\right)$$

$$0{\cdot}8544 = m_a'.$$

Therefore new modulation is 85·44%.

Power output at new modulation

$$= P_c\left[1+\frac{(0{\cdot}8544)^2}{2}\right]$$

$$= 1000\times 1{\cdot}365$$

$$= 1365 \text{ Watt.}$$

Power dissipated at the valve

$$P_p = P_c\left(1+\frac{m_a^2}{2}\right)\left(\frac{1}{\eta}-1\right)$$

$$= 1365\left(\frac{1}{0{\cdot}34}-1\right)$$

$$= 1365 \times 1{\cdot}941$$

$$= \mathbf{2649{\cdot}70\ Watt.}$$

8. *The following data refers to typical F.M. and P.M. systems respectively ;*

F. M. system :

Carrier voltage $e_c = 100 \sin 2\pi 10^6 t$

Modulating voltage $e_m = 5 \cos 2\pi 1000 t$

$\Delta f = 50$ *kc/s.*

P. M. System :

Carrier voltage $e_c = 100 \sin 2\pi 10^6 t$

Modulating voltage $e_m = 2 \cos 2\pi 500 t$

$\Delta\phi = 20$ *rad.*

Find the instantaneous frequencies and instantaneous voltage in both the cases.

F. M. System : Refer to article 30·5.

Instantaneous frequency of f. m. wave is given by

$$f = f_o + \Delta f \cos \omega_m t$$

$$= 10^6 + 50 \times 10^3 \times \cos 2\pi . 1000 t$$

The instantaneous angular frequency

$$\omega_i = \omega_c + K_f E_m \cos \omega_m t$$

$$= 2\pi . f_o + \frac{K_f E_m}{2\pi} . 2\pi \cos 2\pi f_m t$$

$$= 2\pi . 10^6 + \Delta f . 2\pi \cos (2\pi . 1000 t)$$

$$= 2\pi\ [10^6 + 50 \times 10^3 \cos (2\pi . 1000 t)].$$

Also $$\phi' = \omega_c t + \frac{K_f E_m}{\omega_m} \sin \omega_m t + \theta_0$$

$$= 2\pi f_c . t + \frac{\Delta f}{f_m} \sin 2\pi f_m t + \theta_0$$

$$= 2\pi . 10^6 t + \frac{50 . 10^3}{1000} \sin (2\pi . 1000 t) + \theta_0$$

$$= 2\pi . 10^6 t + 50 \sin (2\pi . 1000 t) + \theta_0,$$

The expression for f. m. wave is

$$e = E_c \sin (\omega_c t + m_f \sin \omega_m t)$$

$$= E_o \sin\left(2\pi . f_c . t + \frac{\Delta f}{f_m} \sin 2\pi f_m t\right)$$

$$= 100 \sin\left(2\pi . 10^6 t + \frac{50 . 10^3}{10^3} \sin 2\pi . 1000 t\right)$$

$$= 100 \sin (2\pi . 10^6 . t + 50 \sin 2\pi . 1000 t).$$

P. M. System : Refer to art. 30·8.

Instantaneous frequency of phase modulated wave is given by

$$f = f_c + \Delta f \cos \omega_m t$$

$$= f_c + \Delta\phi\, f_m \cos 2\pi f_m t$$

$$= 10^6 + 2 \times 500 \cos (2\pi\, 500 t)$$

$$= 10^6 + 10^4 \cos (2\pi . 500 t).$$

The expression for p.m. wave is

$$e = E_o \sin (\omega_c t + \Delta\phi \sin \omega_m t)$$

$$= 100 \sin (2\pi . 10^6 . t + 20 \sin 2\pi . 500 t).$$

EXERCISES AND PROBLEMS

1. A carrier wave is amplitude modulated at audio frequency. Deduce an expression to show that two side-bands are produced. What is the relative power in each sideband carrier when modulation coefficient is 0·5 ?
2. What do you mean by the process of modulation ? Why is it necessary for the transmission of intelligence ?

   Describe briefly plate modulated class *C* amplifier for the production of amplitude modulated wave.
3. Explain with the help of suitable diagram how a reactance tube is used to produce frequency modulation.
4. Discuss various methods of modulating the carrier amplitude of broadcast transmitter, with special reference to the method which is used commonly.
5. Describe the main features of amplitude modulation. Describe a plate modulated class *C* amplifier and discuss in detail a method for the measurement of percentage of modulation.
6. Why is single side-band transmission preferred ? Discuss how it is produced by means of balanced modulators. What are its advantages over the conventional amplitude modulation systems ?
7. A carrier having an amplitude of 100 volts and a frequency of 100 mc/s, is frequency modulated with a modulating frequency equal to 1000 cycles The resulting frequency deviation $\triangle f$ is 75 kc/s.
   (*a*) Write an expression for the instantaneous voltage of the modulated wave.
   (*b*) Write an expression for the instantaneous frequency of the modulated wave.
8. The r.m.s. value of radio frequency voltage before modulation by a sine-wave audio frequency voltage is 50 volts, and after modulation 5 volts. Calculate the percentage of modulation.

9. What are tne advantages of f.m., over a.m. ? Derive the expression for the spectrum of f m wave. Under what condition does tLe spectrum not contain carrier frequency ?

10. Compare frequency modulation with phase modulation in terms of the frequency and phase deviations produced by modulation with sinusoidal tones of constant amplitude and different frequencies. Hence show that :
(*a*) a frequency modulator can be used to generate a phase modulated wave
(*b*) a phase modulator can be used to generate a frequency modulated wave.

11. A high frequency oscillation of frequency f is modulated by a frequency p so that the resultant current is represented by the expression

$$i=A\ (1-B\ \sin\ 2\pi pt)\ \sin\ 2\pi ft.$$

Resolve this expression into the carrier and two side frequencies. What is meant by percentage of modulation, and what is the percentage modulation in above case ? A transmitter supplies $10kW$ to the aerial when modulated, what will be its output when modulated to 100%.

[Hint : $i=A\ \sin\ 2\pi ft-AB\ \sin\ 2\pi pt\ \sin\ 2\pi ft$

$$=A\ \sin\ 2\pi ft-\frac{AB}{2}\ [\cos\ 2\pi\ (p-f)\ t-\cos\ 2\pi\ (p+f)\ t].$$

which consists carrier and two side bands. Comparing it with eq. 4 art. 30·1, we find that (across some resistance R)

$$A=\frac{E_c}{R},\quad \frac{AB}{2}=\frac{m_a E_c}{2R}$$

so that power dissipated across R is

$$P=\frac{E_c^2}{R^2}\left(1+\frac{m_a^2}{4}+\frac{m_a^2}{4}\right)=\frac{E_c^2}{R^2}\left(1+\frac{m_a^2}{2}\right)$$

$$=A^2\left(1+\frac{B^2}{2}\right)$$

with 100% modulation $B=1$, so that

$$P=\frac{3}{2}A^2$$

$$=\frac{3}{2}\times 10=15\ kw.$$

because when unmodulated, power A^2 is $10kw$.]

12. Describe the principle and operation of a suppressor grid class C modulated amplifier, what are the advantages of this type of modulation. If a class C modulated R. F. amplifier generates a carrier frequency power of 1·6 kW at an efficiency of 66%. Calculate the output of the modulator to modulate the carrier with 100% modulation.

[Hint : The modulated output of r. f. amplifier is given by

$$P_o=\eta\ P_{bb}\left(1+\frac{m_a^2}{2}\right)$$

It is given that

$P_0 = 1{\cdot}6kW$, $\eta = 0{\cdot}6$ and $m_a = 1$ for 100% modulation,

so that

$$P_{bb} = \frac{1{\cdot}6}{0{\cdot}6} \times \frac{2}{3} kW.$$

Average power supplied by the modulator (modulating amplifier) will be

$$P_m = P_{bb} \frac{m_a^2}{2} = \frac{1{\cdot}6}{0{\cdot}6} \times \frac{2}{3} \times \frac{1}{2} kW$$

$$= \frac{16000}{18}$$

$$= 808 \text{ watts}]$$

13. Discuss in detail the advantages and disadvantages of frequency modulation.
14. What are the advantages of pulse modulation system. Discuss pulse width and pulse code modulation systems. Illustrate with circuit diagrams.

———

31 DEMODULATION (DETECTION)

Demodulation is the process in which the original modulating voltage is recovered from the modulated wave. Circuits employed for detection are called detectors or demodulators and depends upon the type of modulation. They are broadly classified as :

(*i*) Amplitude modulated detectors.

(*ii*) Frequency modulated detectors.

31·1. PRINCIPLE OF A. M. DETECTION AND CLASSIFICATION OF A. M. DETECTORS :

In amplitude modulation envelope, the carrier amplitude has symmetrical variations at its positive and negative half cycles that correspond to modulating voltage. As this envelope is symmetrical, the mean value of detected or rectified r.f. current comes out to be zero. Therefore non-linear devices are used which make the modulated wave unsymmetrical either by wiping out the negative half r.f. cycles or by making positive and negative half cycles unequal. Consequently d.c. or mean value of detected (or rectified) r.f. current then rises and falls at modulating signal rate inaccordance with the amplitude of r.f. signal and may be separated by employing suitable filters. Thus audio frequency variations are extracted from the modulated wave. There are mainly two types of detectors:

(*i*) small signal or square law detectors (non-linear detectors),

(*ii*) large signal or linear detectors.

In square law detectors the output amplitude (detected output voltage) is not a linear function of input amplitude. The detecting tube is operated at the non-linear portion of its characteristic. Since current through the detecting circuit elements flows continuously, plate current of the detecting tube can be expressed in the form of power series expansion. This type of detection is suited for small signals.

In linear detection output amplitude (detected output voltage) is proportional to the input voltage *i.e.* their relationship is linear. In such a type of detection, rectification properties of the diode is used so that the current in diode flows in pulses. The behaviour of such a device, therefore, can not be analysed by the use of power series representation. This type of detection is suited for large signals.

31·2. SQUARE LAW DETECTION :

Square law detection can be classified as :

(a) Square law diode detection.

(b) Square law triode detection. his type can be further divided into two classes :

(*i*) anode bend detection or grid detection,
(*ii*) grid leak detection.

31·2-1. SQUARE LAW DIODE DETECTION :

The circuit for square law diode detection is shown in figure 1.

In this process of detection, we utilise the non linear portion of dynamic volt-ampere characteristic of diode.

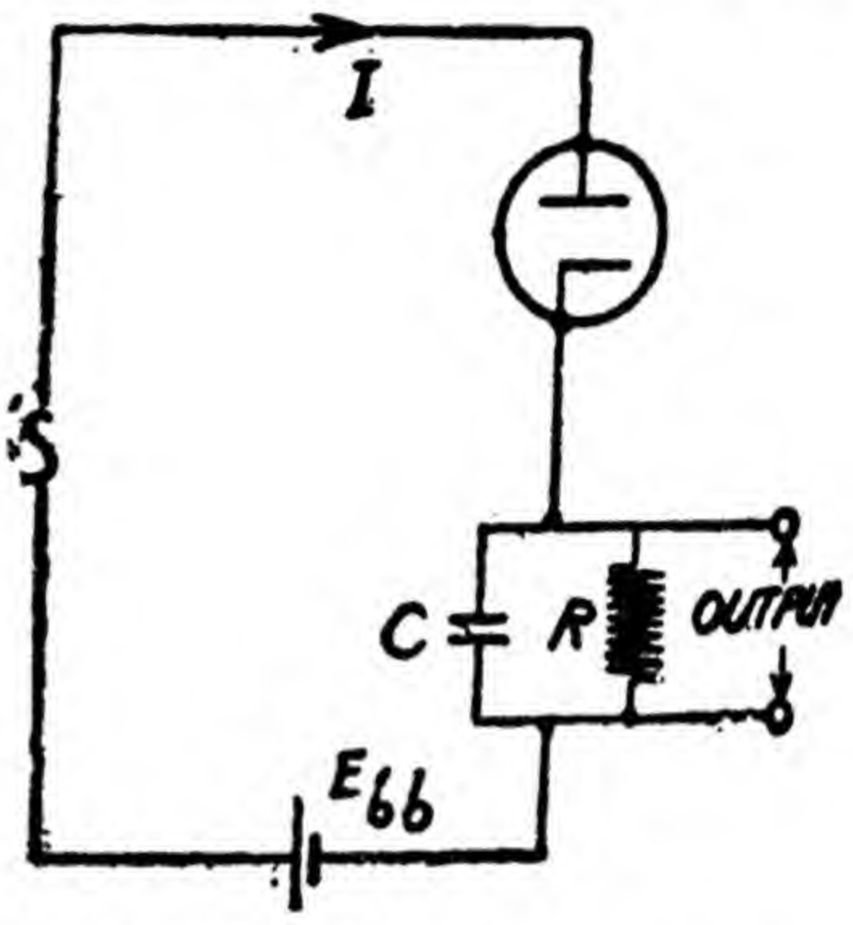


Fig. 1. Circuit of square law diode detector.

Since in the case of square law detector the input carrier voltage is of very small magnitude, the operating point will be on excessively non-linear portion of the characteristic. On the other hand, in the case of linear diode detector this operating point shifts towards linear portion of characteristic because of large input carrier voltage and most of the operation takes place over the linear region of characteristic.

Battery E_{bb} is applied to bias the plate of diode positively in order to shift the working point to a non-linear portion on the characteristic as shown in fig. 2.

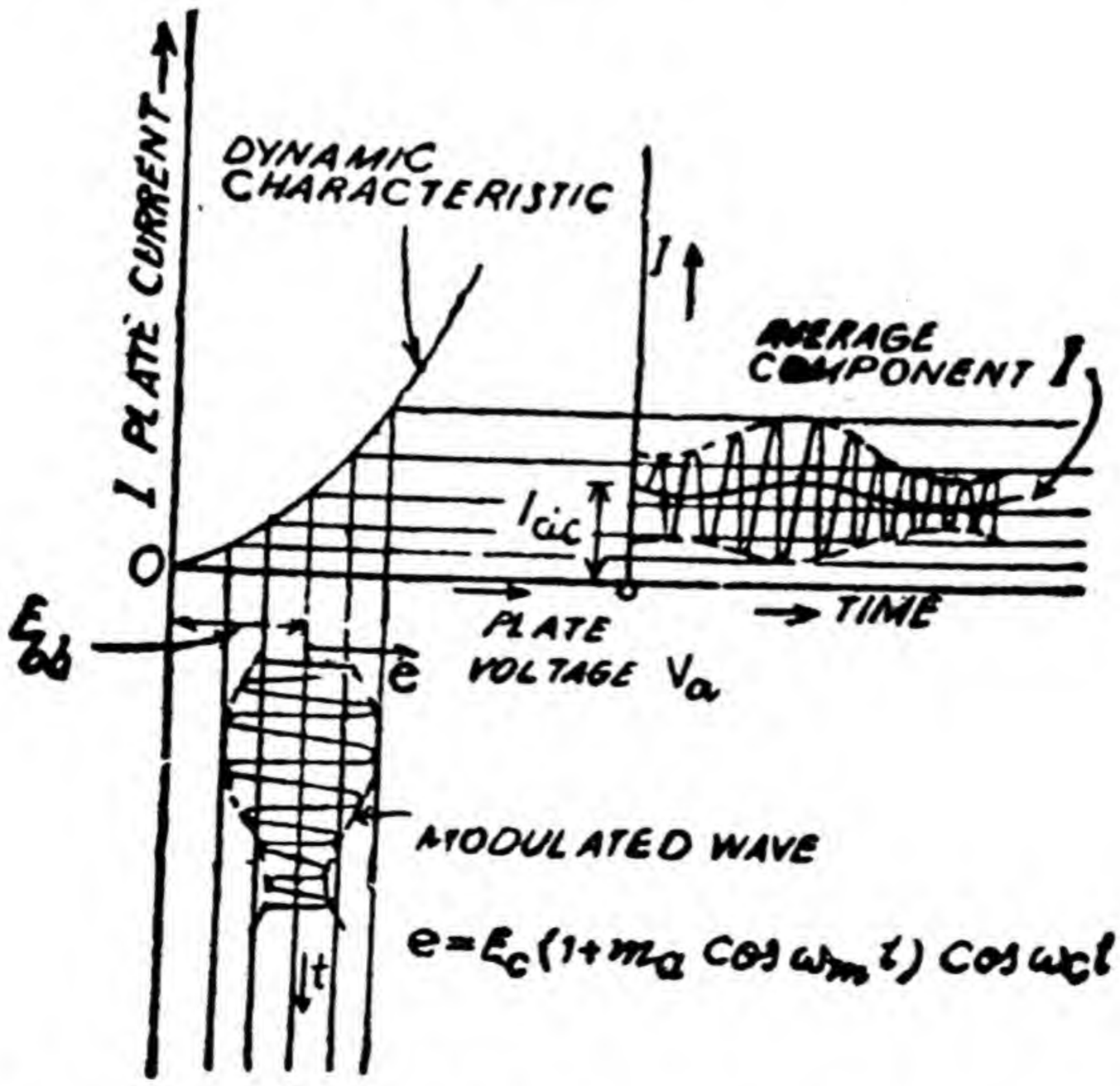


Fig. 2. Output wave of detector with resistive load.

When a modulated wave is applied to the detector input then since the operation takes place over the non-linear portion, lower half of the input waveform will be compressed in the output as shown. Due to unsymmetrical shape of output waveform, varying average current will be produced. This average current will be made up of two components.

The capacitor C offers low reactance to carrier frequencies so that all carrier frequencies are bypassed and only average current component flows in load resistance R.

The performance of square law diode detector can be represented by a power series. The plate current is related to plate potential by a square law relation

$$I = a_1 V_a + a_2 V_a^2. \qquad \ldots (1)$$

where $V_a = E_{bb} + e$

$$= E_{bb} + E_c (1 + m_a \cos \omega_m t) \cos \omega_c t$$

$$= E_{bb} + E_c \cos \omega_c t + \frac{m_a E_c}{2} \cos (\omega_c - \omega_m) t + \frac{m_a E_c}{2} \cos (\omega_c + \omega_m) t.$$

Putting this value of V_a in current relation (1), we shall get terms containing frequencies $2\omega_c$, $2(\omega_c + \omega_m)$, $2(\omega_c - \omega_m)$, ω_m and $2\omega_m$. The radio frequency terms are bypassed through the shunting capacitor C, while ω_m and $2\omega_m$ produce output voltage V_0 across load resistance R. ω_m term gives desired modulating voltage, while $2\omega_m$ is a harmonic term producing distortion.

31·2·2. SQUARE LAW TRIODE DETECTION :

We shall first discuss grid detection or anode bend detection.

(a) **Grid detection** (*Anode bend detection*) : The method of detection makes use of non-linear portion of the dynamic transfer characteristic of a grid controlled tube, *i.e*, triode, tetrode or pentode.

Circuit arrangement is shown in figure 3. The grid is biased such as to shift the operating point to the non linear portion of the dynamic transfer characteristics.

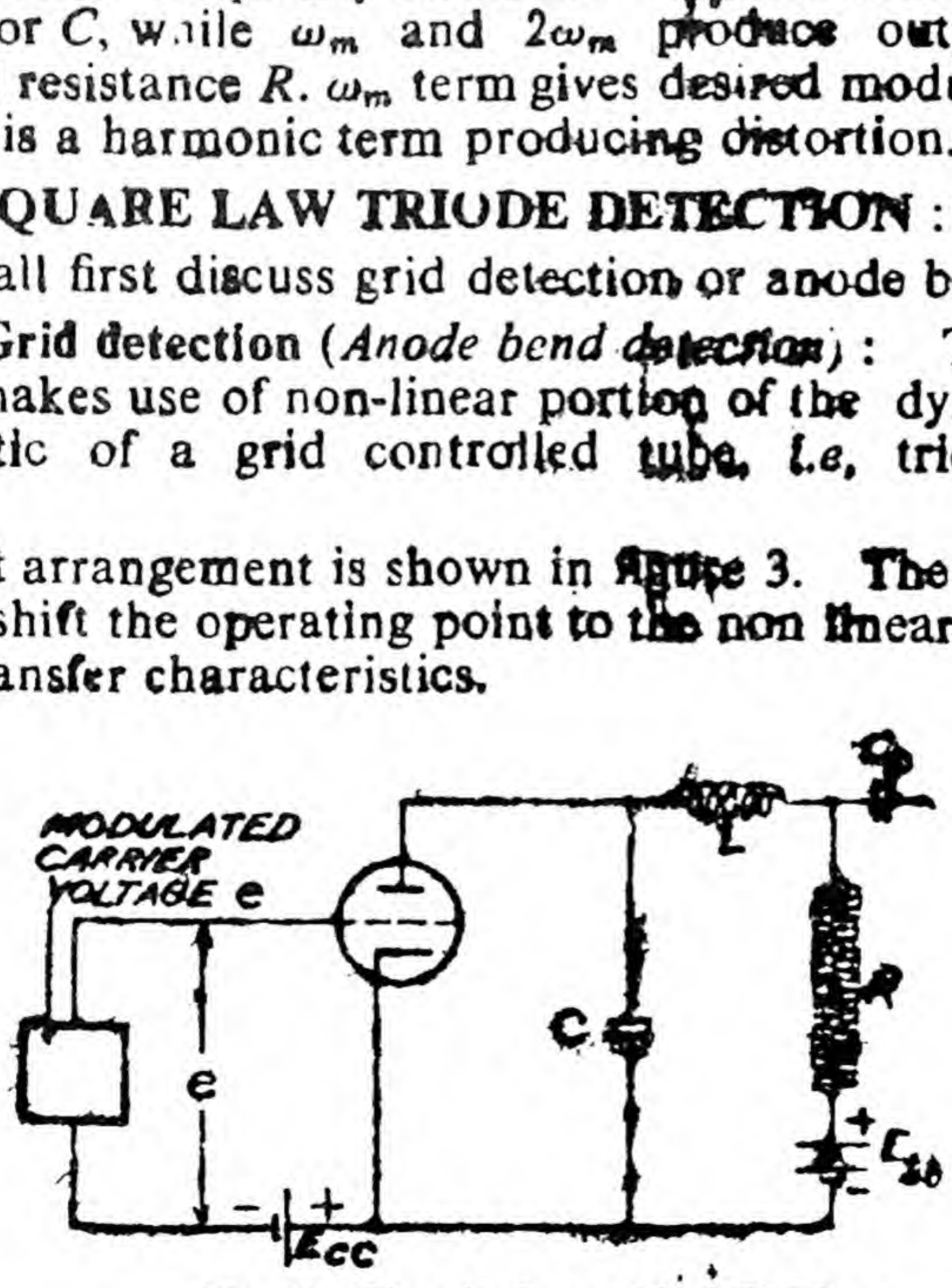


Fig. 3. Circuit for grid detection.

As shown in figure 4 average value of plate current over one cycle of the carrier wave, fluctuates with the signal amplitude and therefore, potential drop across the load resistance R has modulation frequency component which follows the modulation envelope of the signal. Condenser C bypasses radio frequency component.

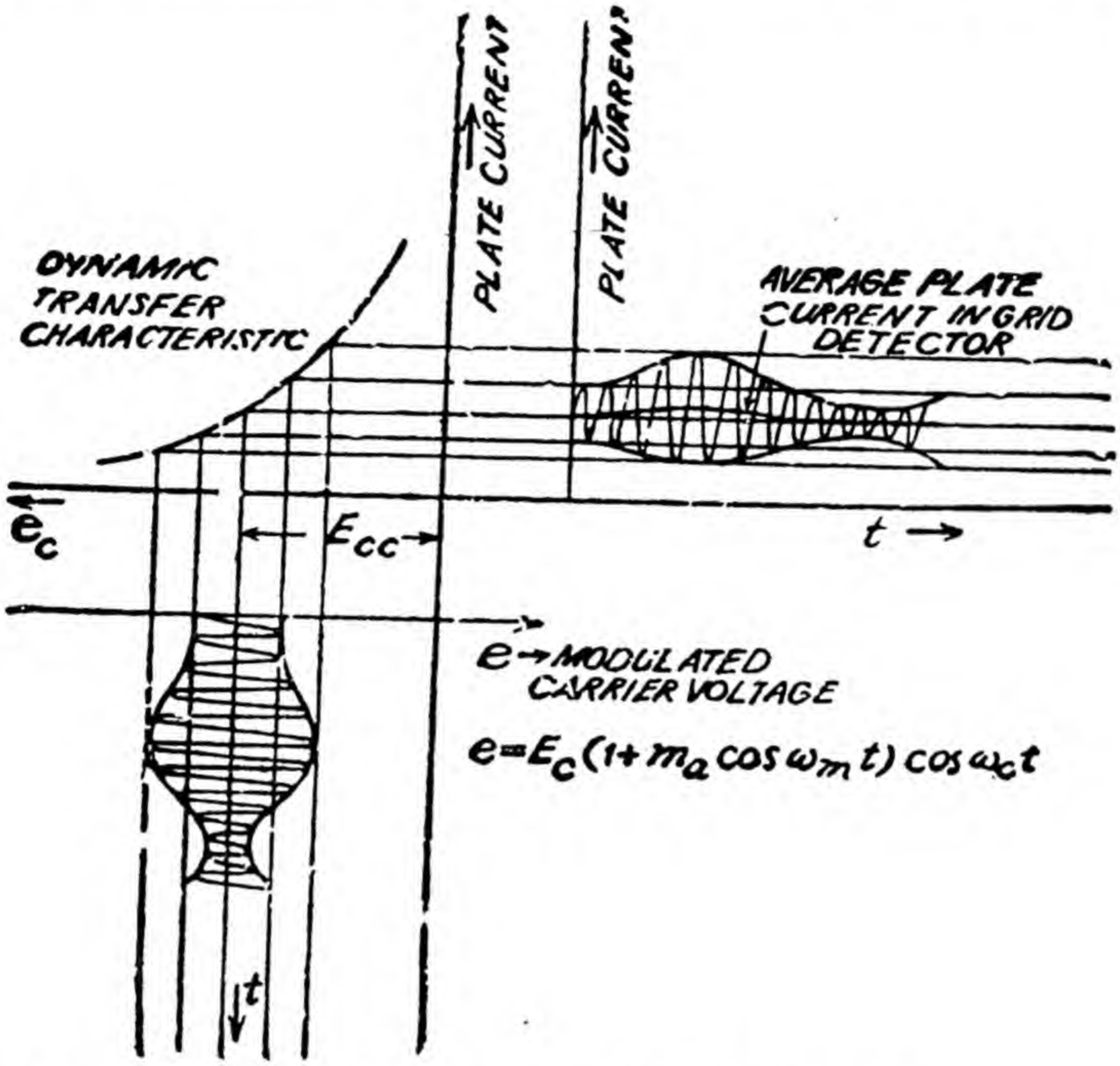


Fig. 4. Illustration of the operation of detector.

Analysis of anode bend detection : Let a.c. plate current i be represented approximately by the power series

$$i = a_1 e + a_2 e^2. \qquad \ldots(1)$$

neglecting higher order terms, where

$$e = E_o (1 + m_a \cos \omega_m t) \cos \omega_c t.$$

Therefore equation (1) becomes

$$i = a_1 E_c (1 + m_a \cos \omega_m t) \cos \omega_c t$$

$$+ \frac{a_2}{2} E_o^2 (1 + m_a \cos \omega_m t)^2 (1 + \cos 2\omega_c t)$$

$$= a_1 E_c \left[\cos \omega_c t + \frac{m_a}{2} \cos (\omega_c + \omega_m) t + \frac{m_a^2}{2} \cos (\omega_c - \omega_m) t \right]$$

$$+ \frac{a_2}{2} E_o^2 \left[1 + 2m_a \cos \omega_m t + \frac{m_a^2}{2} (1 + \cos 2\omega_m t) \right]$$

$$+ \frac{a_2}{2} E_c^2 \left[\cos 2\omega_c t + 2m_a \cos \omega_m t \cos 2\omega_c t + \frac{m_a^2}{2} \cos 2\omega_c t \right.$$

$$\left. + \frac{m_a^2}{2} \cos 2\omega_m t \cos 2\omega_c t \right] \qquad \ldots(2)$$

Since radio frequency components are bypassed by the shunt condenser, the remaining terms in the plate current are given by

$$i_1=\frac{a_2}{2}E_c^2\left[\ 1+\frac{m_a^2}{2}+2m_a\cos\omega_m t+\frac{m_a^2}{2}\cos 2\omega_m t\ \right]$$

$$=\frac{a_2}{2}E_c^2\left(\ 1+\frac{m_a^2}{2}\right)+a_2E_c^2m_a\cos\omega_m t$$

$$+\frac{a^2}{4}E_c^2m_a^2\cos 2\omega_m t. \qquad ...(3)$$

In the equation (3) the terms :

(*i*) $\frac{a_2}{2}E_c^2\left(\ 1+\frac{m_a^2}{2}\right)$ constitutes the d.c. component of plate current ($I_{d.c.}$).

(*ii*) The second term $a_2E_c^2m_a\cos\omega_m t$ constitutes the desired modulation frequency component.

(*iii*) The third term $\frac{a_2}{4}E_c^2.m_a\cos 2\omega_m t$ constitutes the distortion term in frequency, $2\omega_m$ being a harmonic of modulation frequency ω_m.

The d.c. component is blocked from being transmitted to next stage by means of coupling condenser C_c. The modulation frequency component in the output plate current, i, is therefore

$$i_2=a_2E_c^2\left(\ m_a\cos\omega_m t+\frac{m_a^2}{4}\cos 2\omega_m t\ \right). \qquad ..(4)$$

The second harmonic component should be kept as low as possible. If it is to be kept low, say 10% of the fundamental component, then it is necessary that

$$\frac{m_a^2}{4}<\frac{m_a}{10}$$

or $$m_a<0{\cdot}4.$$

That is, the modulation index should be less than 0·4. Ordinary broadcast is performed generally at 40%. Therefore, detectors can be used to act satisfactorily.

Example : Find the second harmonic distortion in audio output if the depth of modulation exceeds 20% in an A M. signal applied to square law detector.

From equation (4) the amplitude of the second harmonic component is $m_a^2/4$. The amplitude of desired modulation frequency component is m_a. Therefore, percentage of second harmonic is

$$=\frac{m_a^2/4}{m_a}\times 100=25m_a\%$$

If depth of modulation exceeds 20% *i.e.* $m_a>0{\cdot}2$ then percentage of second harmonic will be greater than $25\times 0{\cdot}2$ or 5. Thus second harmonic distortion will exceed 5%.

Advantages : Main advantage of anode bend detection is that when signal amplitude is not large, no current flows in the control grid circuit, and thus power consumption in grid circuit is very small. Superiority of anode bend detection over linear diode detection is that it does not require any power from the modulated carrier voltage.

(b) Grid Leak Detection : Circuit for grid leak detection is shown in fig. 5. The method makes use of the non-linear portion of the dynamic grid current-grid voltage characteristic of a grid controlled tube shown in fig. 6.

The complete circuit can be understood as a combination of diode detector and a stage of triode amplification for detected signal. Grid and a cathode act as elements of a diode, rectifying the input modulated signal and *RC* combination acts as its load impedance.

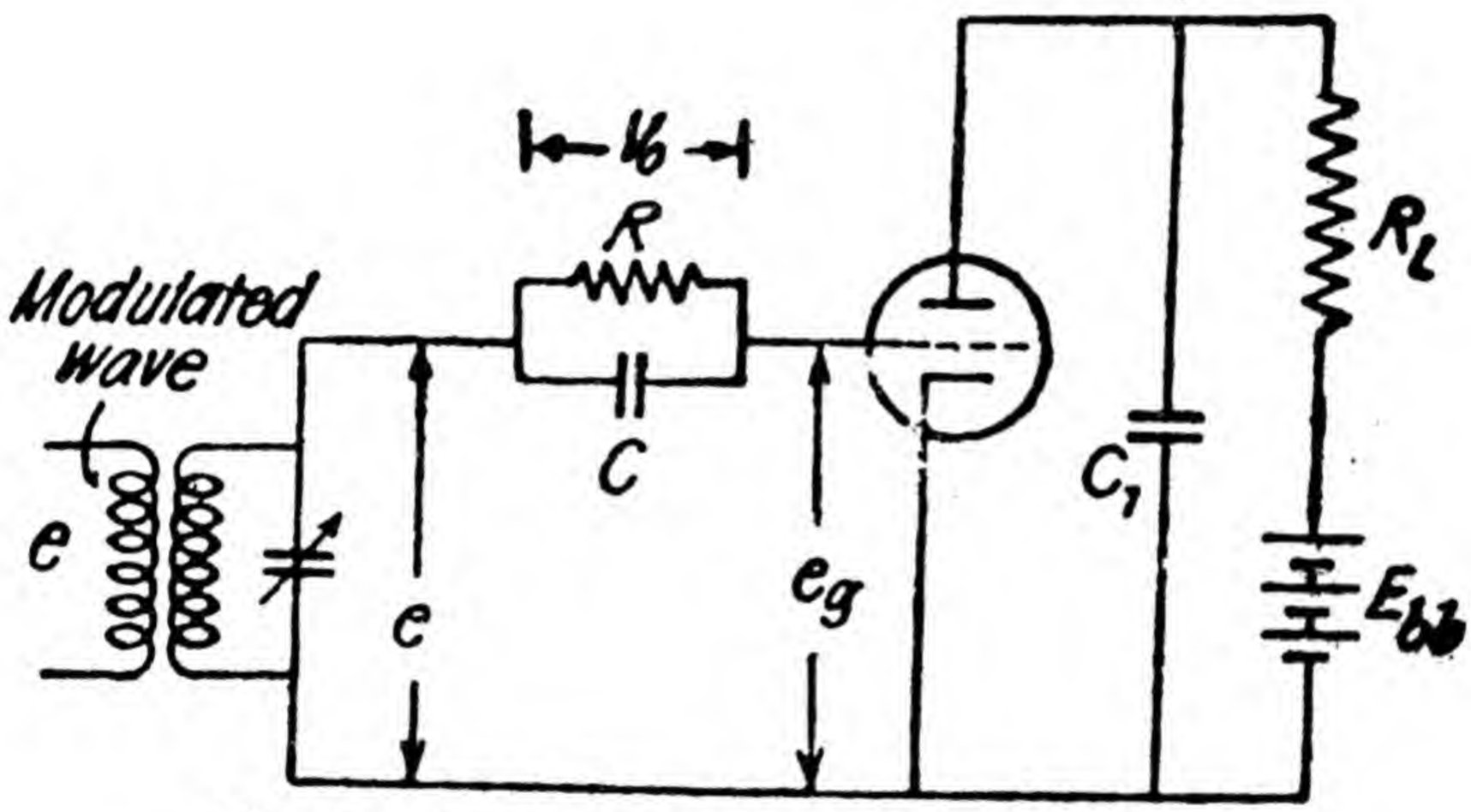


Fig. 5. Circuit of triode grid detection.

At positive peaks of input signal, electrons flow from cathode to the grid of the triode, charging the condenser, C, to a potential difference V_0. Similar to diode detector, the potential difference V_0 has three components :

(*i*) a direct component, nearly equal to the amplitude of the carrier wave.

(*ii*) an audio frequency component having amplitude nearly equal to the amplitude of the modulation envelope and

(*iii*) a radio frequency ripple.

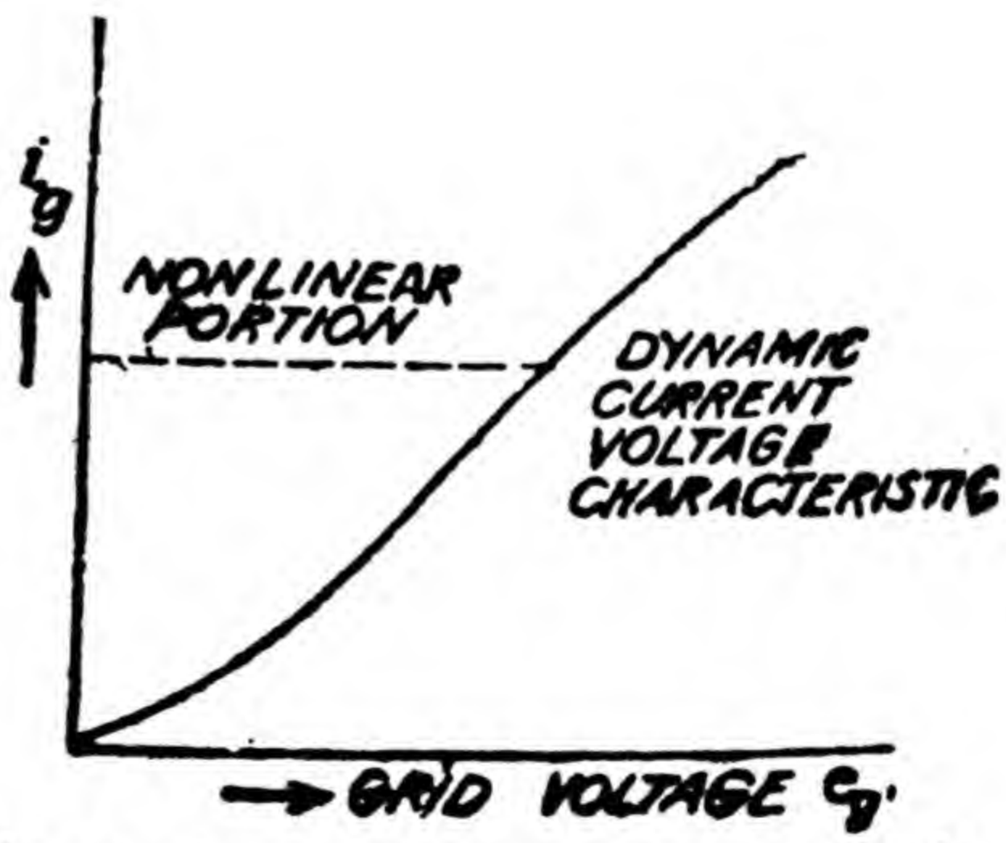


Fig. 6. Dynamic i_g–e_g characteristic of a grid-controlled tube.

This detected signal across the resistor R is also available between grid and cathode of the amplifier tube. Since grid potential

varies inaccordance with modulation frequency, the anode current will also follow the modulation envelope. The current sets up a modulating voltage across R_L and r.f. components are bypassed by shunt capacitor C_1.

Advantage : The important characteristics of grid leak detection method is that it provides high sensitivity, *i.e.* a comparatively large output voltage at modulat on frequency may be obtained for a given magnitude of input modulated voltage.

Disadvantage : When the input signal is small, current in grid circuit flows continuously for the complete r.f. cycle resulting in high consumption of power in grid circuit.

31 3. LINEAR DIODE DETECTORS :

This type of detectors are being widely used in home receivers because of satisfactory operation, although the modulated input to the detector must be large. These detect r circuits are of two types :

(*a*) average detection type, and

(*b*) envelope or peak detection type.

31·3-1. LINEAR AVERAGE TYPE DIODE DETECTOR :

Circuit is shown in fig. 7. The modulated signal is fed to the diode through a resonant circuit tuned to the carrier frequency. Resonant circuit has a bardwith sufficient to accommodate the sidebands of the signal and thus signal input to diode is without distortion. As shown in fig. 7 (*b*), input signal is rectified and

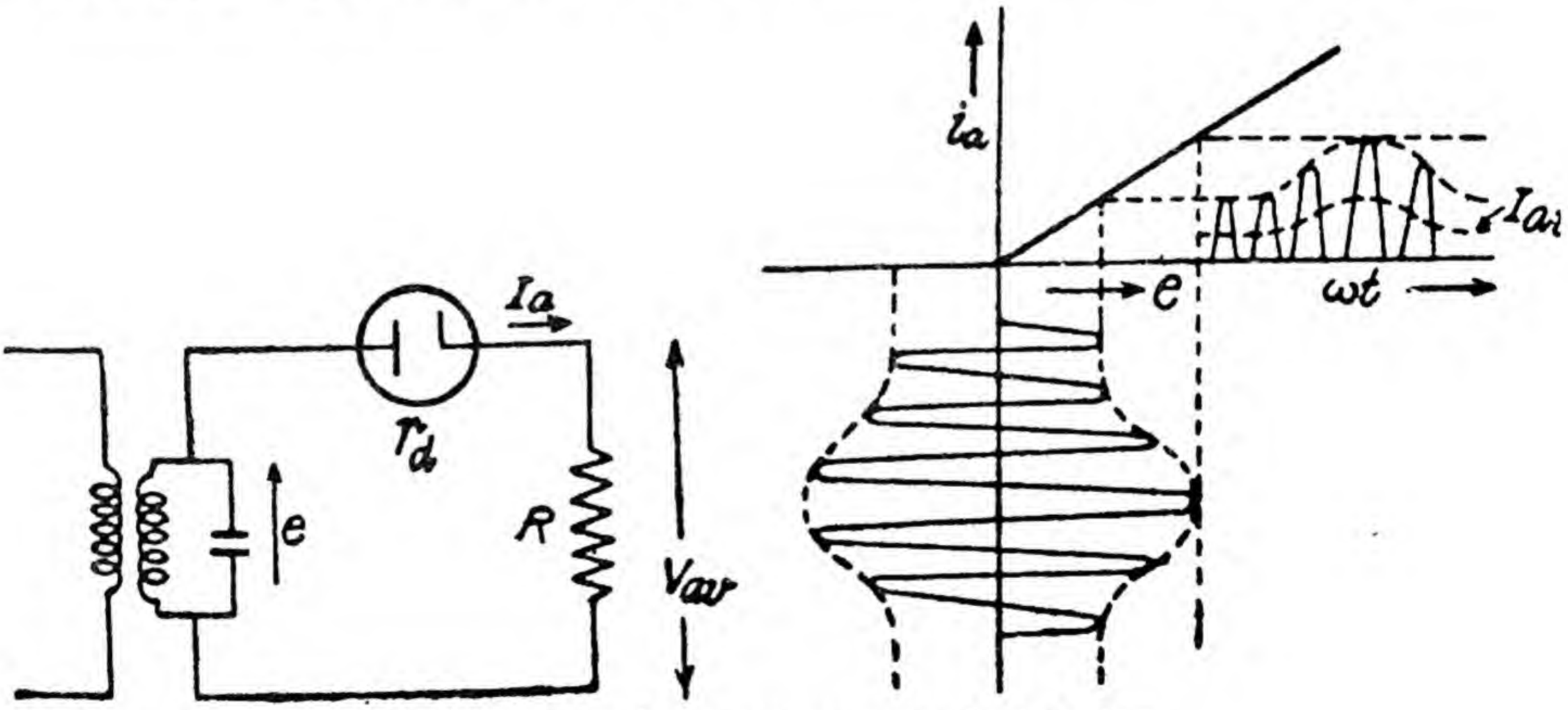


Fig. 7. (*a*) Linear average dicde detector.
(*b*) operational characteristic

average current, I_{av}, in the diode circuit is obtained which follows the variation of modulating voltage. We can show mathematically that output V_{av} contains modulating signal component. Let the modu'ated voltage at input be

$$e = E_c (1 + m_a \cos \omega_m t) \cos \omega_o t.$$

The current pulse during conduction period is then

$$i_a=\frac{e}{r_d+R}=\frac{E_c}{r_d+R}\cdot(1+m_a\cos\omega_m t)\cos\omega_c t$$

for $-\pi/2<\omega_c t<\pi/2$ r_d is diode resistance and R is the load resistance.

The average current over a period of carrier frequency is

$$I_{av}=\frac{1}{2\pi}\int_{-\pi}^{\pi} I_a\, d(\omega_c t)$$

$$=\frac{1}{2\pi}\int_{-\pi/2}^{+\pi/2}\frac{E_c}{r_d+R}(1+m_a\cos\omega_m t)\cos\omega_c t\, d(\omega_c t)$$

If $\cos\omega_m t$ term is assumed to be constant (since $\omega_m \ll \omega_c$, $\cos\omega_m t$ does not appreciably change during one cycle of carrier frequency) over an interval $-\pi/2$ to $\pi/2$ of carrier frequency then

$$I_{av}=\frac{E_c}{r_d+R}(1+m_a\cos\omega_m t).\frac{1}{2\pi}\int_{-\pi/2}^{+\pi/2}\cos\omega_c t\, d(\omega_c t)$$

$$=\frac{1}{\pi}\cdot\frac{E_c}{r_d+R}(1+m_a\cos\omega_m t), \qquad ...(1)$$

Therefore average output voltage is

$$V_{av}=\bar{V}=\frac{1}{\pi}\frac{E_c.R}{r_d+R}(1+m_a\cos\omega_m t), \qquad ...(2)$$

which shows that average component of output voltage contains a d c. component

$$\frac{E_c R}{\pi(r_d+R)}$$

and the modulating signal component with no distortion (*i.e.* no harmonics of modulating frequency, ω_m, is

$$\frac{E_c.R}{\pi(r_d+R)}m_a\cos\omega_m t.$$

The detection efficiency is given by

$$\eta=\frac{\text{Average value of output power}}{\text{Peak value of input voltage}}$$

Therefore, for average detector, we have

$$\eta=\frac{\bar{V}}{E_c(1+m_a\cos\omega_m t)}$$

$$=\frac{1}{\pi}\cdot\frac{R}{(r_d+R)}, \qquad ...(3)$$

using eq. 2.

or

$$\eta=\frac{1}{\pi}\cdot\frac{R/r_d}{1+R/r_d}. \qquad ...(4)$$

which, when $\frac{R}{r_d} \gg 1$, reduces to

$$\eta = \frac{1}{\pi} \qquad \ldots(5)$$

and is maximum.

This type of diode detector is not very efficient be cause the modulating signal component is one-third of the input voltage ($1/\pi$ term in eq. 2). Further carrier frequency component along with its harmonics is also present at the output wh ch is not desirable. For filtering this high frequency component and improving the efficiency, a capacitor is connected across load resistor R as detailed in next article.

31·3-2. ENVELOPE DIODE DETECTOR :

They are most commonly employed in all types of home radio receivers because of their large detection efficiency and also being practically free from distortion.

Let us first discuss the effect of introducing condenser and resistor in a diode circuit.

(*i*) *Introduction of a condenser* : Circuit is shown in fig. 8 (*a*).

Let us suppose that as soon as we switch the supply, positive half cycle starts. A current will then flow through the diode charging the condenser C and making its plate connected to cathode-positive. The potential difference across the condenser, V_c appears as a negative bias [shown by a line in fig. 8 (*b*)] on the anode of the valve due to which the anode potential will be

$$V_a = e - V_c,$$

where e is supply voltage.

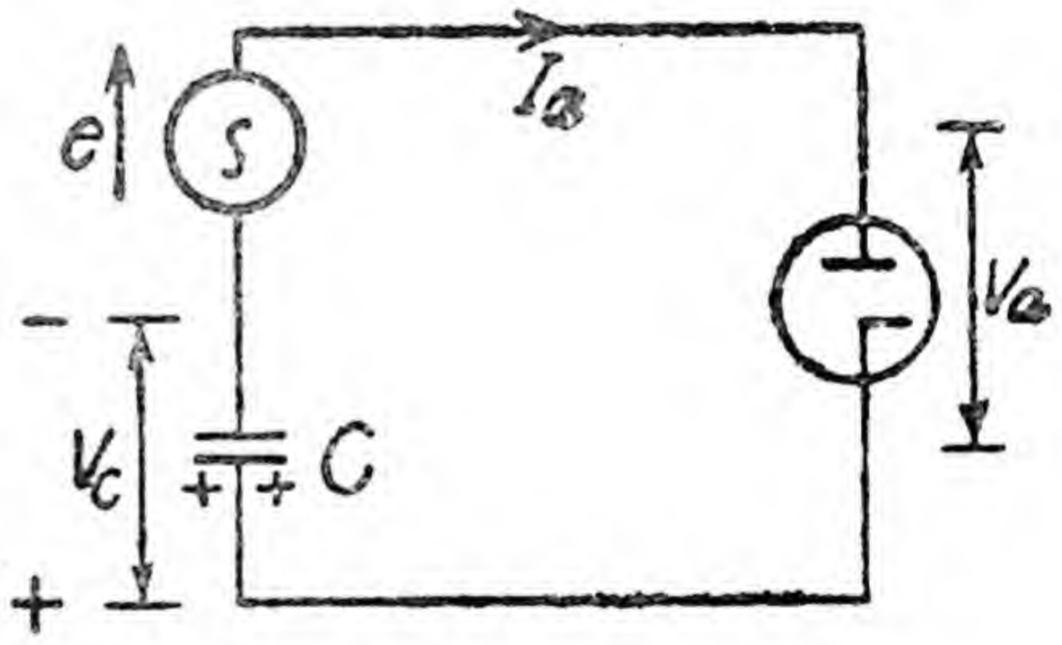


Fig. 8. (*a*) Condenser in a diode circuit.

This potential across condenser goes on increasing with the flow of current. Now towards the end of the positive half cycle, a stage comes when falling magnitude of the supply voltage V_1. becomes equal to the rising magnitude of bias voltage and hence the current stops to flow. Now during the rest of this cycle, since the current does not flow, diode is open, condenser cannot discharge but retains the previous charge.

On next positive half cycle, current will start flowing only when the supply voltage reaches the value V_1. With the flow of current, charge on condenser will again increase and hence the potential V_c Since this basis voltage increases, the current will rise for a smaller period as compared to first positive half cycle and stops

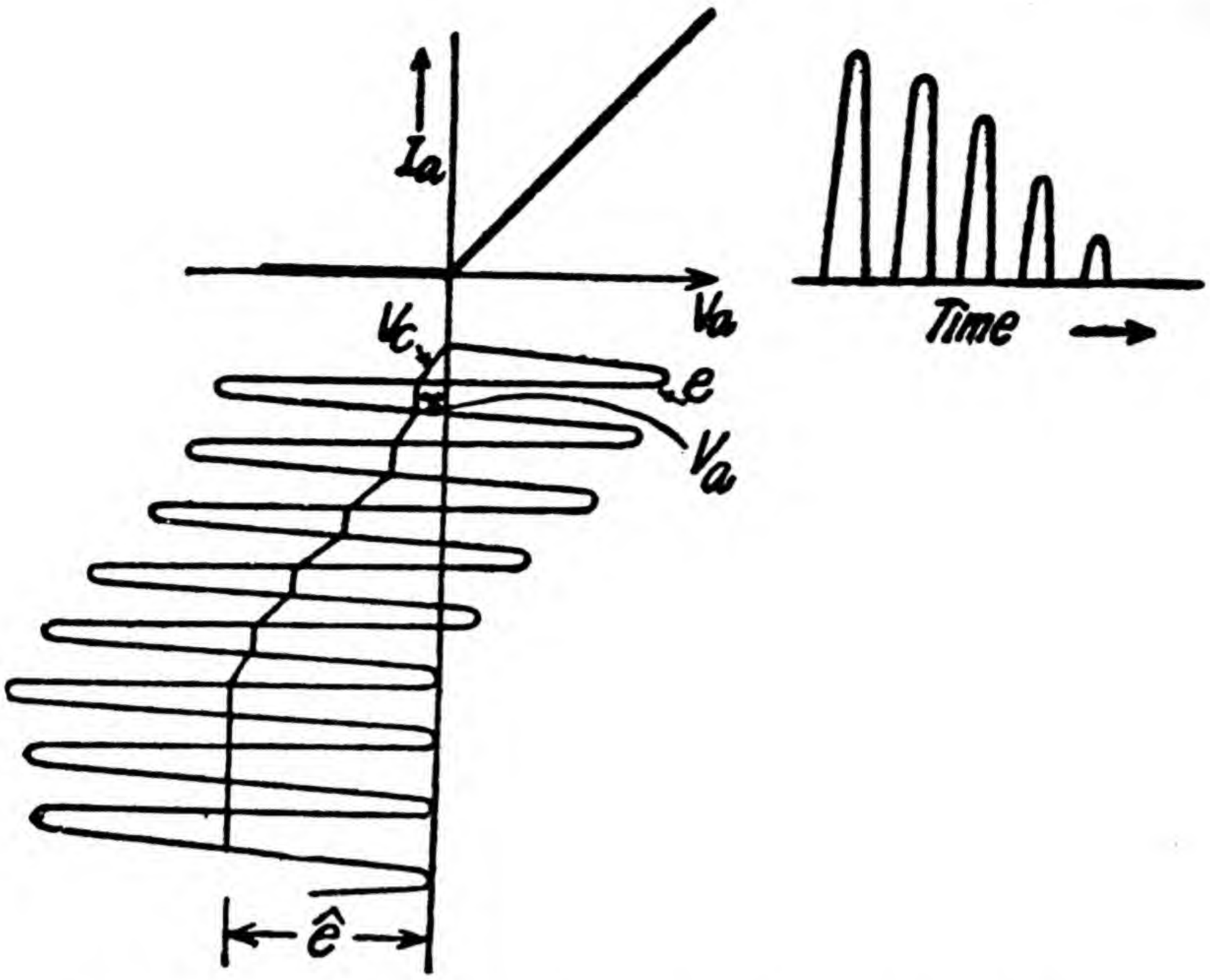


Fig 8 (*b*) Action of diode with condenser in series.

sooner. Finally a stage comes, when V_o rises to the peak value, *e*, of supply voltage and then no current flows through the valve at any time.

(*ii*) *Introduction of resistance in previous circuit of fig.* 8 (*a*) : Due to the introduction of resistance, at negative half cycle, the condenser will discharge through it and hence V_o will decrease. If time constant *CR* is sufficiently large as compared with the period of supply voltage, condenser will discharge very little between two positive cycles. Thus, ultimately, the potential across condenser will rise on switching the supply. But due to the increase in this potential, current pulses become lower and thus give less charge to the condenser. Further the interval between the current pulses, in which condenser discharges, becomes longer. Finally a stage comes when the voltage gained by the condenser during charging process becomes equal to the voltage lost during discharge and thus average pontential difference across condenser becomes constant {V_c' in fig. 8 (*d*)} which is less than peak value of supply voltage. In this case, therefore, current will always flow at the positive peaks of supply voltage.

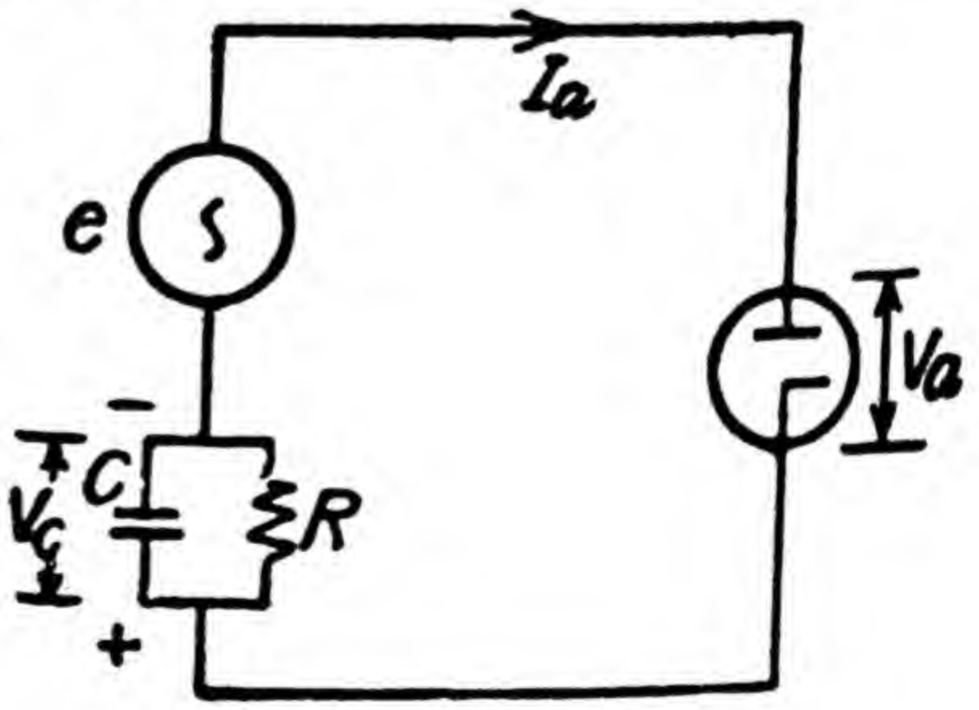


Fig. 8 (*c*)

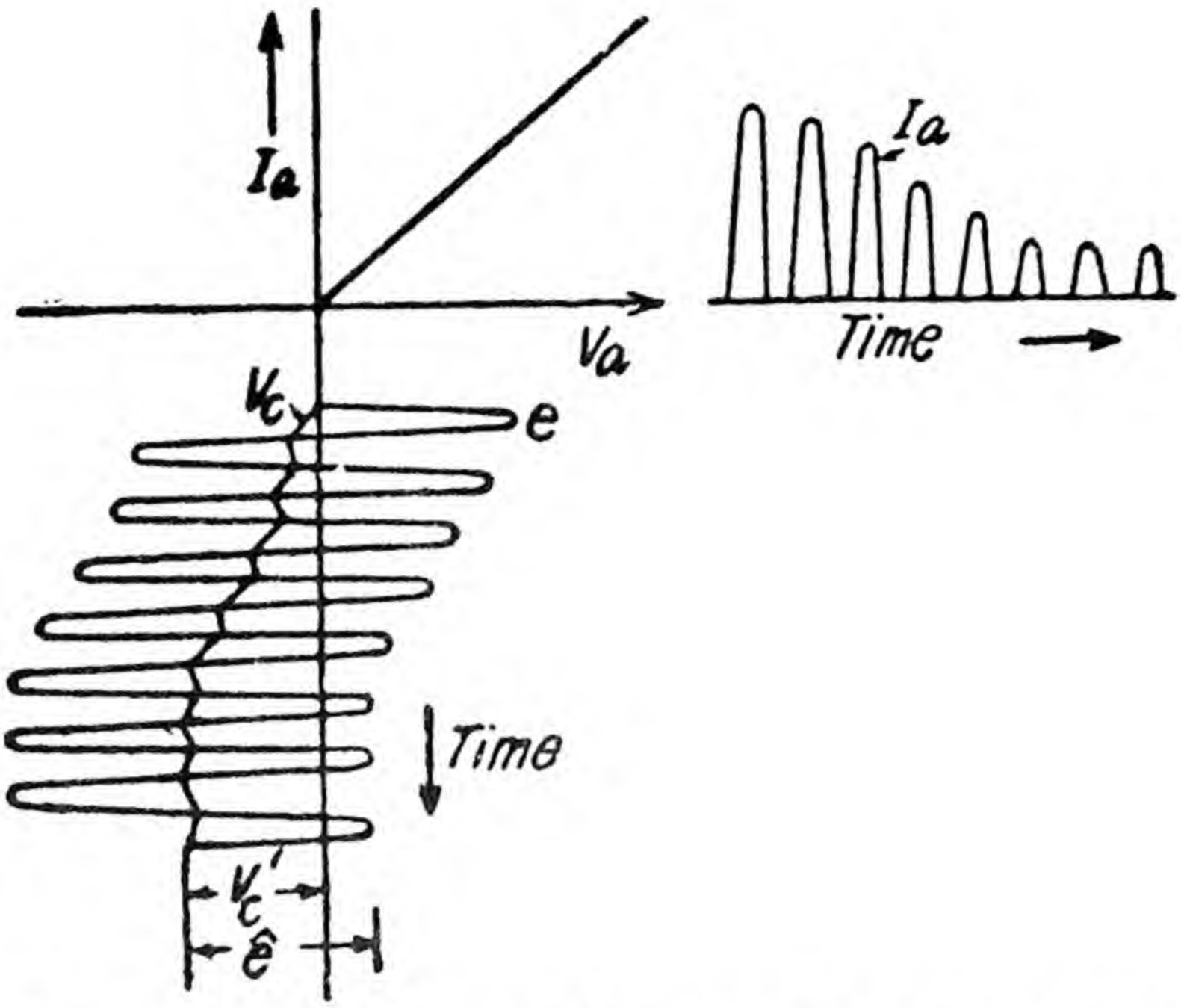


Fig. 8. (*d*) Action of diode with a *C-R* circuit in series.

Linear Diode Detector Circuit : Linear diode detection utilises the rectification characteristic of a diode. In fig. 9 (*a*) and 9 (*b*) basic circuit of linear diode, detector is shown.

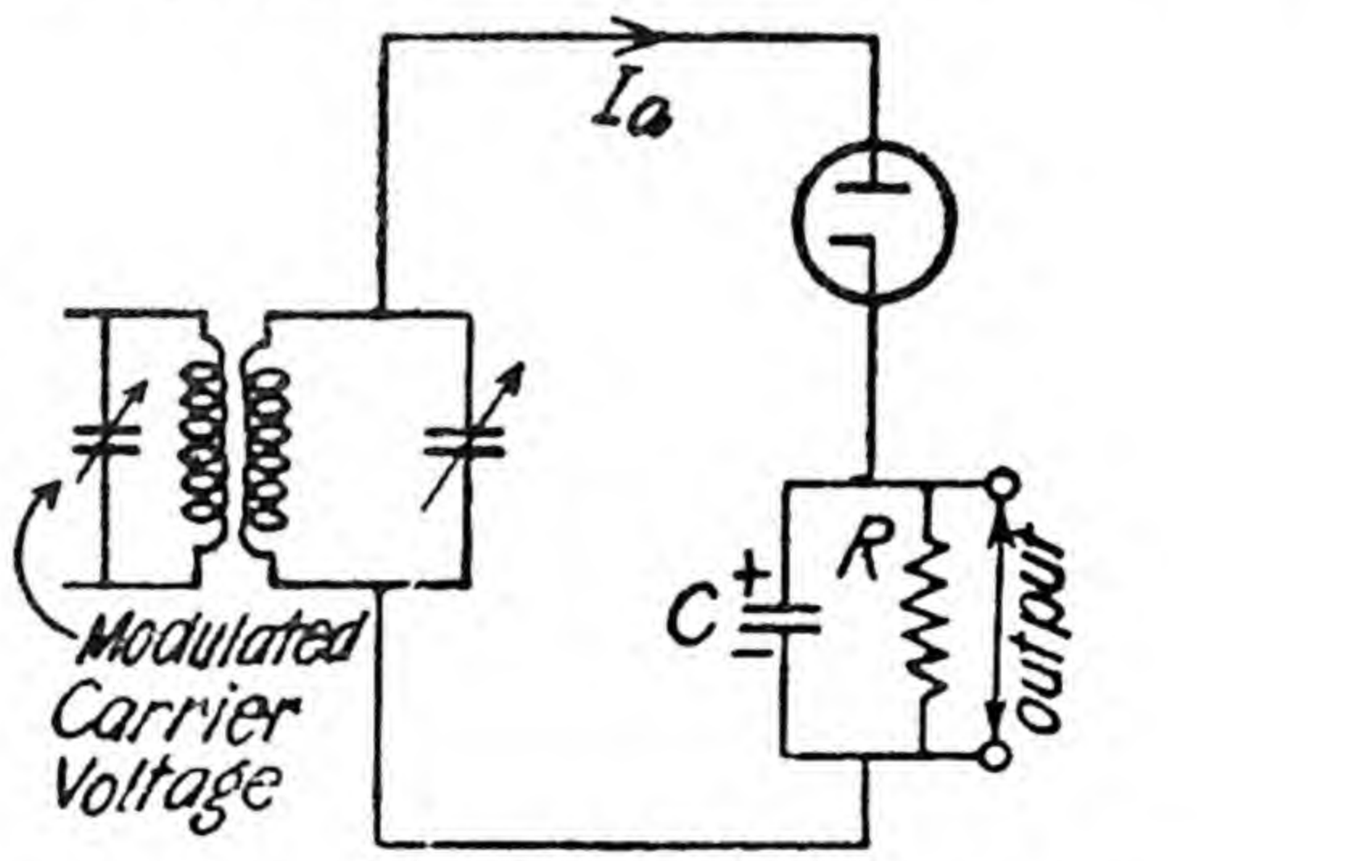


Fig. 9 (*a*). Basic circuit of linear diode detector.

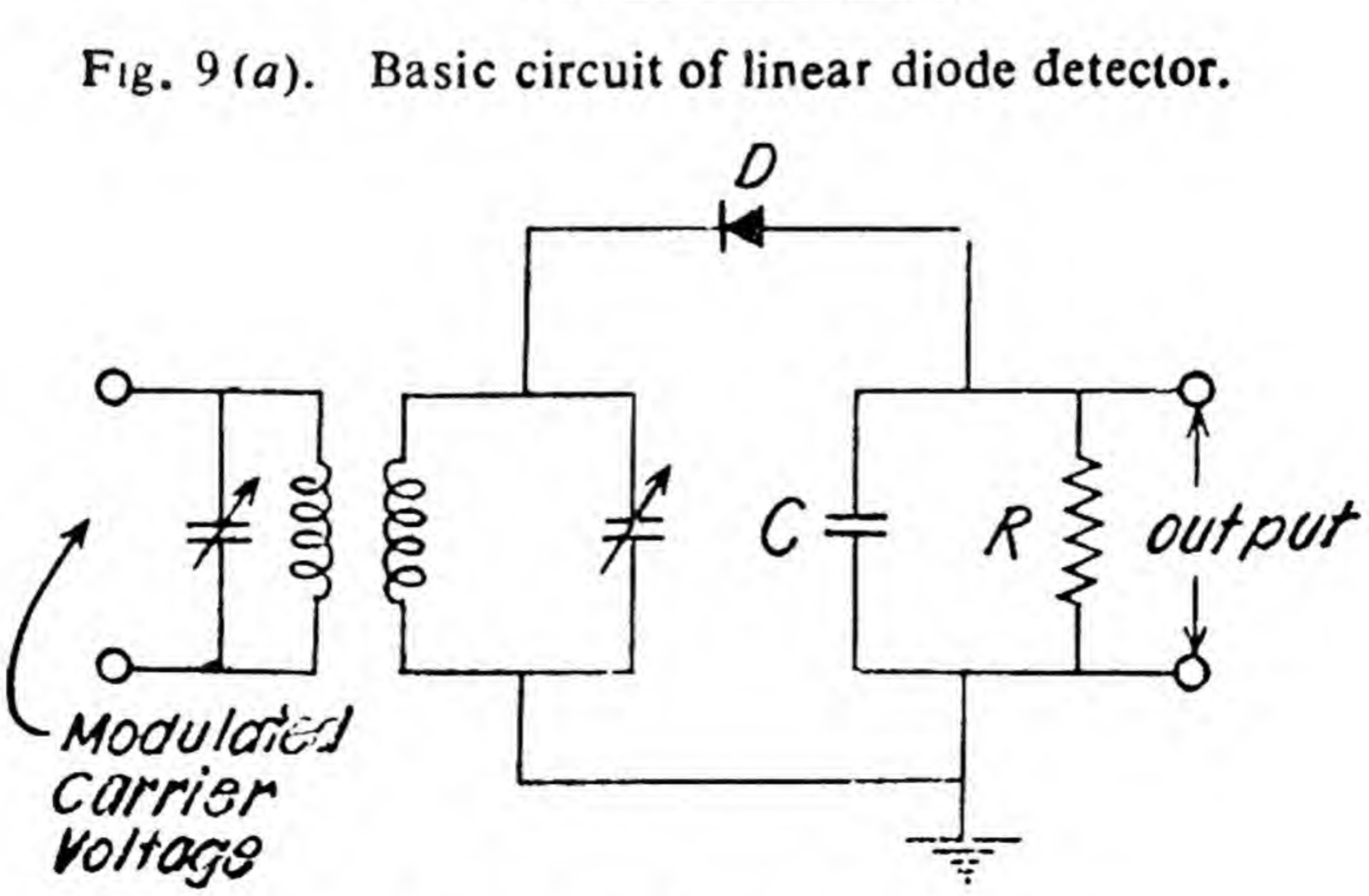


Fig 9 (*b*). Basic circuit linear diode detector.

The modulated carrier voltage is applied to a series combination of diode and load impedance consisting of a resistance R in parallel with a capacitance C. As shown in figure 9, (*a*) and (*b*) I is the current through diode V_a is the potential across the diode, V_0 is output voltage

Satisfactory Conditions of Detection : The condenser with a capacitance C should be such that its reactance at radio frequency f_c is much less than R, *i.e.*

$$\frac{1}{2\pi f_c C} \ll R,$$

while its reactance as audio frequency f_m is much greater than R, *i.e.*

$$\frac{1}{2\pi f_m C} \gg R.$$

We see from above conditions that impedance of load circuit is small at radio frequency while large at modulation frequency. Therefore, when a modulated signal is applied to the circuit, the radio frequency components of the current through diode sets up only a small potential difference across the load but audio frequency and direct components will set up large potential differences (Refer to fig. 10) Thus a unidirectional potential difference V_0,

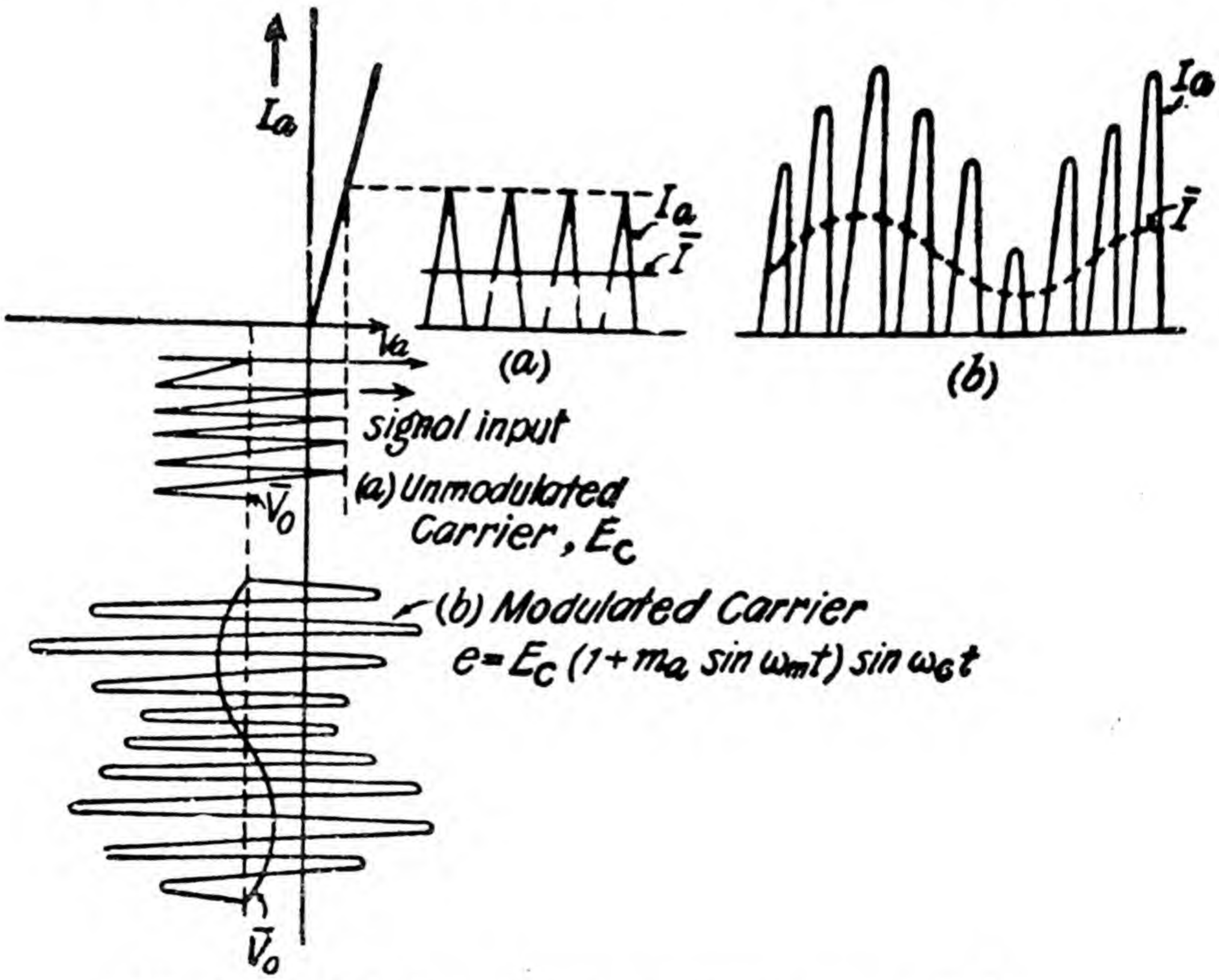


Fig. 10. Variation of average current through diode
(*a*) when input is unmodulated
(*b*) when input is modulated.

varying at modulation frequency will appear across the load and the ripple be very small.

When the carrier wave is not modulated, the average current $\bar{I}$ through diode will be constant; but when the wave is modulated, then the average value of the current $\bar{I}$ through the diode will fluctuate at audio frequency rate over one cycle of carrier wave. Therefore, if the period of modulation T_m is greater than CR, then potential difference across the load fluctuates with average current through diode because condenser will not bypass audio frequency.

The radio frequency is bypassed for the period of carrier wave is much less than time constant CR.

Therefore the condition of satisfactory detection is

$$T_m \gg CR \gg T_o$$

or

$$\frac{1}{f_m} \gg CR \gg \frac{1}{f_c}$$

Theory of detection : For the analysis of diode detection, we assume that signal applied to the diode is having a *large* amplitude and is not deeply modulated. For this reason, changes in the peak current through the diode will remain confined to straight portion of volt-ampere characteristic of the diode. In otherwords, we, during the analysis, can take the characteristic of the diode to be a straight line, the slope of which can be taken as

$$r_d = \frac{V_a}{I_a}. \qquad \ldots(1)$$

With carrier as input : Let us first consider the case when unmodulated signal (carrier only)

$$e_c = E_c \cos \omega_c t \qquad \ldots(2)$$

is input to the diode. As shown in fig. 10(*a*), the average current

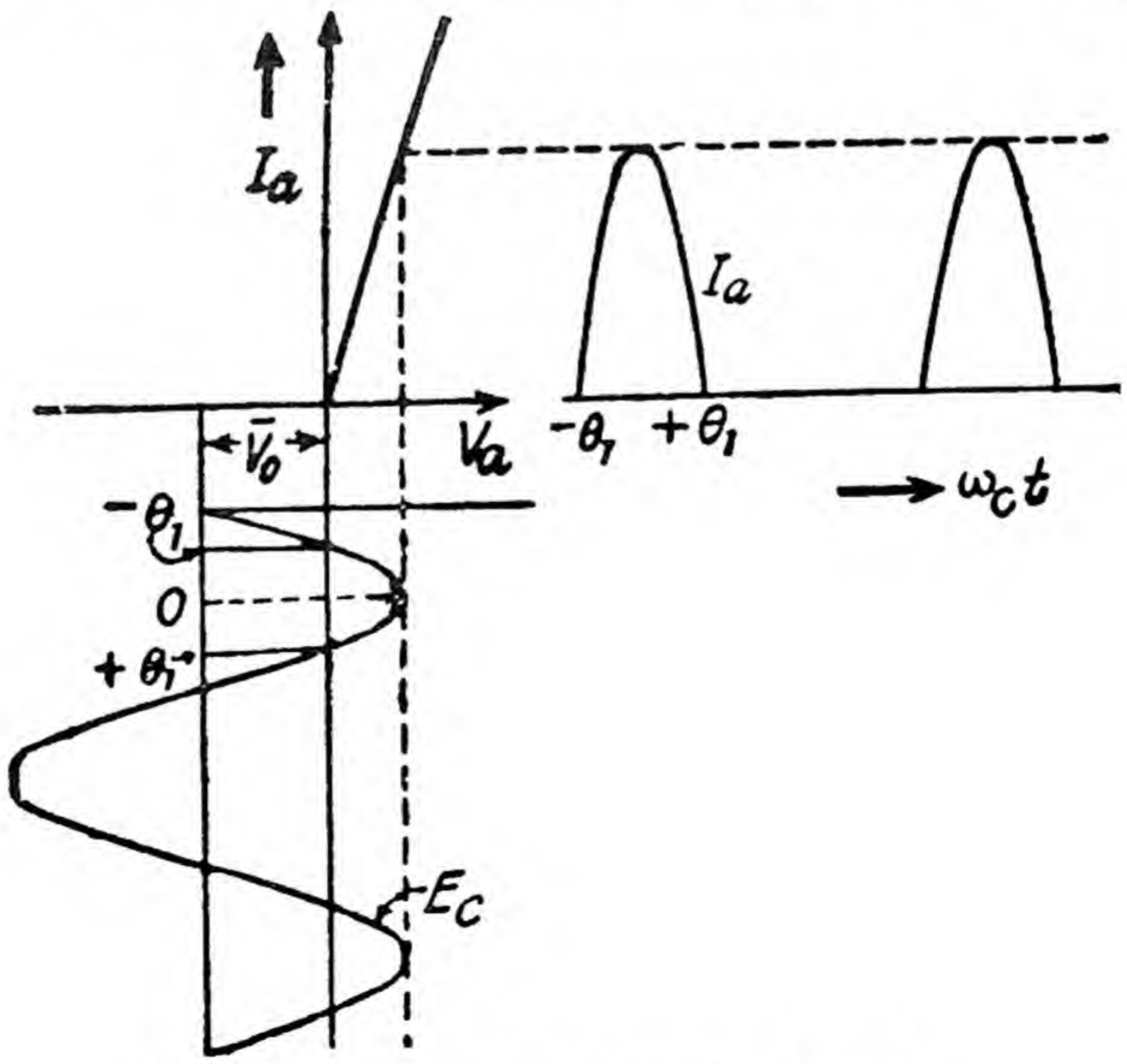


Fig. 11. Conduction angles of diode

$\bar{I}$, through the diode will be constant in this case. This current will develop the average potential

$$\bar{V}_0 = R\bar{I} \qquad \ldots(3)$$

across the load, R. We ignore the ripples in output voltage because they will be very small when CR time constant is long compared with the periodic time of the carrier voltage.

It is obvious from fig. 11, that plate current flows in pulses starting from angle $-\theta_1$ to $+\theta_1$. At $\omega_c t_1 = -\theta_1$, signal voltage (which is carrier here), e_o becomes equal to the potential difference across the load. Then

$$e_c = E_c \cos \omega_o t_1 = \bar{V}_0 \qquad \ldots(4)$$

or
$$\omega_c t_1 = -\theta_1 = \cos^{-1} \frac{\bar{V}_0}{E_o} \qquad \ldots(5)$$

The average anode current is given by

$$\bar{I} = \frac{1}{2\pi}\int_{-\theta_1}^{+\theta_1} I_a \, d(\omega_o t)$$

$$= \frac{1}{\pi}\int_{0}^{\theta_1} \frac{V_a}{r_d} \, d(\omega_c t)$$

$$= \frac{1}{\pi r_d}\int_{0}^{\theta_1} (E_c \cos \omega_c t - \bar{V}_0) \, d(\omega_o t)$$

$$= \frac{1}{\pi r_d} (E_c \sin \theta_1 - \bar{V}_0 \theta_1). \qquad \ldots(6)$$

Putting this value of $\bar{I}$ in eq. (3), we get the mean output voltage

$$\bar{V}_0 = \frac{E_c R}{\pi r_d}\left(\sin \theta_1 - \frac{\bar{V}_0}{E_c} \theta_1 \right) \qquad \ldots(8)$$

The ratio of the average potential difference across the load to the peak signal voltage is called efficiency of detection. Denoting it by β we have

$$\beta = \frac{\bar{V}_0}{E_c} = \cos \theta_1 \text{ (from eq. 5)} \qquad \ldots(8)$$

Putting eq. (8) into eq. (7), we get

$$\cos \theta_1 = \frac{R}{\pi r_d} (\sin \theta_1 - \theta_1 \cos \theta_1)$$

or
$$\tan \theta_1 - \theta_1 = \frac{\pi r_d}{R}. \qquad \ldots(9)$$

Since angle θ_1 is small, we can expand $\tan \theta_1$ and taking only first two terms of expansion

$$\left(\theta_1 + \frac{\theta_1^3}{3} \right) - \theta_1 = \frac{\pi r_d}{R}$$

or $$\frac{\theta_1^3}{3}=\frac{\pi r_d}{R}$$

or $$\theta_1^2=\left(\frac{3\pi r_d}{R}\right)^{2/3} \quad ...(10)$$

For small values of angle of flow $2\theta_1$ and hence for small value of θ_1, the detection, β, is given by

$$\beta=\cos\theta_1=' -\frac{\theta_1^2}{2}$$

$$=1-\frac{1}{2}\left(\frac{3\pi r_d}{R}\right)^{2/3}. \quad ...(11)$$

From this eq. we note that :

(*i*) detection efficiency increases as load resistance R increases, approaching unity for considerably large values of load resistance compared to diode resistance.

(*ii*) detection efficiency is large for small values of diode resistance r_d. In actual diode, r_d varies because of curvature in volt-amp. characteristic and has a minimum value over the linear range of the characteristic *i.e.* for large values of anode potential. Consequently detection efficiency increases with the increase of input signal amplitude.

Thus for efficient operation of the detector, input signal must be large and load resistance must be kept much greater than the diode resistance.

With modulated wave as input : Let us now consider the case when the signal input to the diode is modulated. The signal is expressed as

$$e=E_c(1+m_a\cos\omega_m t)\sin\omega_c t,$$

where ω_m and ω_c are audio and carrier frequencies respectively. As discussed previously and depicted in fig. 10, when signal input to detector circuit is modulated then average value of current $\bar{I}$ through diode and so the average potential $\bar{V}_0$, across load will vary at modulation frequency following modulation envelope. Detection efficiency

$$\beta=\frac{\text{average potential difference across load}}{\text{peak signal input voltage}}$$

$$=\frac{\bar{V}_{output}}{E_c(1+m_a\sin\omega_m t)}$$

or $$V_{output}=\beta E_c+\beta E_c m_a\sin\omega_m t. \quad ...(12)$$

Eq. (12) has two components in the output voltage. The direct component of potential difference across load is

$$V_b=\beta E_c \quad ...(13a)$$

and modulation frequency component is

$$V_1=m_a\,\beta E_c\sin\omega_m t. \quad ...(13b)$$

Eq. (13*b*) represents the required expression for modulating voltage output (audio output). In order to obtain appreciable audio output, efficiency of detection should be high which, inturn, requires that load resistance R should be much greater than the diode resistance r_d. Further we notice that the signal amplitude

$$E_c(1+m_a \sin \omega_m t)$$

will be different when $\sin \omega_m t = +1$ than when $\sin \omega_m t = -1$. Consequently diode resistance r_d, which decreases when signal amplitude increases, will be different for these two values. This means detection efficiency β is greater when $\sin \omega_m t = +1$ than when $\sin \omega_m t = -1$. The net effect of this variation is that the modulation frequency output, V_1, of the detector is not sinusoidal but distorted. The degree of distortion increases when the depth of modulation is large, because then the diode resistance r_d varies over greater range. In order to reduce this distortion, a high value of load resistance, R, is chosen so that the ratio (r_d/R) is small. Then the term $\frac{1}{2}(3\pi r_d/R)^{1/3}$ in eq. (11) is small and detection efficiency β is little affected reducing, thereby, the distortion. The diode detectors thus operate most efficiently when the input voltage is large and load resistance is high enough compared to diode resistance.

Equations (13*a*) and (13*b*) can also be put in another form including diode resistance r_d and load resistance R.

From eq. (7), for unmodulated signal input having amplitude, E_c the output voltage is

$$\bar{V}_{output} = \frac{E_c R}{\pi r_d}\left(\sin \theta_1 - \frac{\bar{V}_{output}}{E_o}\theta_1 \right).$$

Now with modulated signal input having amplitude $E_c(1+m_a \sin \omega_m t)$, output voltage from eq. (12) is

$$\bar{V}_{output} = \beta E_o(1+m_a \sin \omega_m t)$$

Substituting these values of amplitude and output voltage in right hand side of the previous equation, we get

$$\bar{V}_{output} = \frac{E_o(1+m_a \sin \omega_m t)\, R}{\pi r_d}\left(\sin \theta_1 - \frac{\beta E_o(1+m_a \sin \omega_m t)}{E_c\,(1+m_a \sin \omega_m t)}.\theta_1 \right)$$

$$= \frac{RE_c(1+m_a \sin \omega_m t)}{\pi r_d}(\sin \theta_1 - \beta\theta_1)$$

$$= \frac{RE_o}{\pi r_d}(\sin \theta_1 - \theta_1 \cos \theta_1) + \frac{RE_c m_a \sin \omega_m t}{\pi r_d}(\sin \theta_1 - \theta_1 \cos \theta_1)$$

in which we have put $\beta = \cos \theta_1$.

Therefore d.c. output voltage is

$$V_b = \frac{RE_c}{\pi r_d}(\sin \theta_1 - \theta_1 \cos \theta_1) \qquad \ldots(14a)$$

and modulation frequency component *i.e.* a.c. output

$$|V_1| = \frac{RE_c m_a}{\pi r_d}(\sin \theta_1 - \theta_1 \cos \theta_1) \qquad \ldots(14b)$$

The effect of bypass condenser : Choice of time constant RC in linear diode detector : Diagonal Clipping :

The time constant RC cannot be chosen too high or too low. If time constant is very small, then during the non-conducting intervals discharge curve is almost vertical which produces large fluctuations in the output voltage. On the other hand, if RC is high, then the discharge of condenser during the non-conducting intervals will be too slow. The discharge curve now will be almost horizontal so as to miss several peaks of the rectified output voltage during negative peaks as shown in fig. 12.

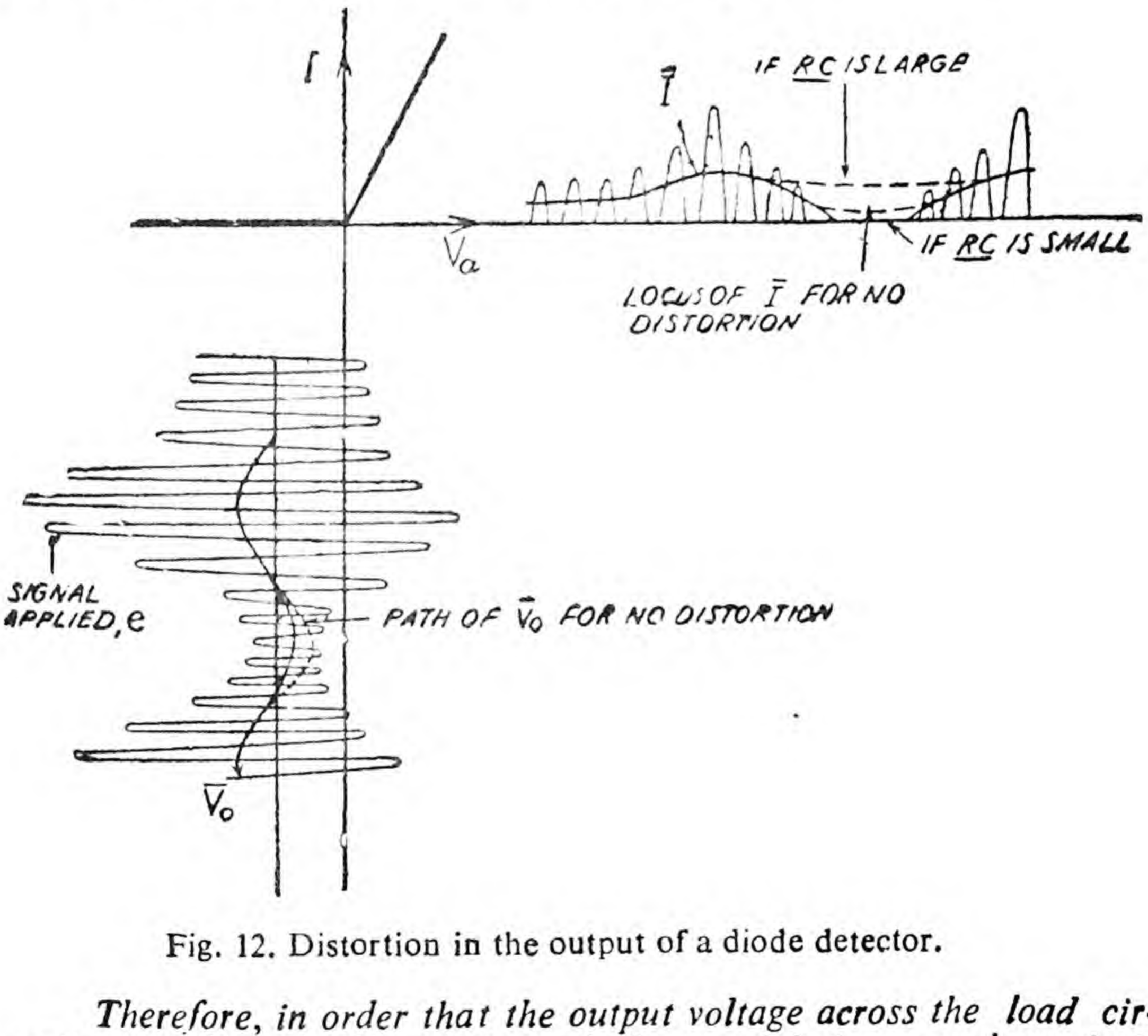


Fig. 12. Distortion in the output of a diode detector.

Therefore, in order that the output voltage across the load circuit may follow the fluctuations of the modulation envelope, it is necessary that the natural rate of fall of potential across load should be equal to or greater than the highest rate at which it is required to fall in the process of detection.

In general, if a condenser of capacitance C is charged to potential difference $\hat{V}$, then its discharge through a resistance R can be represented by the equation

$$V = \hat{V} e^{-t/RC},$$

so that the natural rate of fall of potential is

$$-\frac{dV}{dt}=\frac{\hat{V}}{RC}e^{-t/RC}=\frac{V}{RC}.$$

In the present case across load the average potential difference is $\overline{V}_0$: therefore natural rate of fall will be given by

$$\frac{\overline{V}_0}{RC}. \qquad ...(15)$$

For no distortion, the curve of $\overline{V}_0$ in fig. 12 must follow the equation

$$\overline{V}_0=\beta E_o\,(1+m_a \sin \omega_m t).$$

The required rate of fall of the average potential difference across load is given by

$$-\frac{d\overline{V}_0}{dt}=-\beta E_c m_a \cos \omega_m t. \qquad ...(16)$$

As we have discussed previously that for no distortion, natural rate of fall of potential across load must be equal to or greater than the required rate, we must have from equations (15) and, (16) that

$$\frac{\overline{V}_0}{RC} \geqslant -\beta E_o m_a \omega_m \cos \omega_m t.$$

Putting the value of $\overline{V}_0$,

$$\frac{\beta E_o\,(1+m_a \sin \omega_m t)}{RC} \geqslant -\beta E_c m_a \omega_m \cos \omega_m t,$$

$$\frac{1}{RC} \geqslant -\frac{m_a \omega_m \cos \omega_m t}{1+m_a \sin \omega_m t}. \qquad ..(17)$$

Inequality (17) will be maximum if

$$\sin \omega_m t=-m_a.$$

Putting this value of sin $\omega_m t$ in eq. (17), we get

$$\frac{1}{RC} \geqslant \pm \frac{m_a \omega_m \sqrt{(1-m_a^2)}}{1-m_a^2}.$$

Omitting negative sign as having no physical significance,

$$\frac{1}{RC} \geqslant \frac{m_a \omega_m}{\sqrt{(1-m_a^2)}} \qquad ...(18)$$

or $$\frac{1}{\omega_m RC} \geqslant \frac{m_a}{\sqrt{(1-m_a^2)}}$$

or $$\frac{X_C}{R} \geqslant \frac{m_a}{\sqrt{(1-m_a^2)}}$$

where $X_C=\frac{1}{\omega_m . C}$ is the reactance of condenser.

Therefore, the condition for no distortion due to the shunting capacitance can be expressed as

$$\frac{X_C}{R} \geqslant \frac{m_a}{\sqrt{(1-m_a^2)}} \qquad ...(19)$$

As modulation index m_a increases, $\frac{m_a}{\sqrt{(1-m_a^2)}}$ also increases and time constant decreases. When modulation is 100%, ratio $\frac{m_a}{\sqrt{(1-m_a^2)}}$ approaches infinity and RC, the time constant, becomes zero. With zero time constant, modulation frequency and carrier frequency both will remain in the output as there will be no filtration of carrier frequency. The output thus largely fluctuates when the time constant is zero. This is the cause for which we do not prefer linear diode detection when the wave is 100% modulated.

For 90% modulation, we see that

$$\frac{X_C}{R} \geqslant 2 \text{ approximately.} \qquad \ldots(20)$$

Relation (20) is the design formula for the load circuit of a diode detector.

Example : Design the load circuit for a linear diodc detector if a signal of 4kc/s with 70% modulation is to be detected satisfactorily.

We shall calculate the values of R and C. The limiting values for them are given by

$$\frac{X_C}{R} = \frac{m_a}{\sqrt{(1-m_a^2)}}$$

or
$$RC = \frac{\sqrt{(1-m_a^2)}}{\omega_m m_a}$$

Given, $m_a = 0{\cdot}7$

$f_m = 4.10^3$

Therefore, $RC = \frac{\cdot\sqrt{\{1-(0{\cdot}7)^2\}}}{2\pi \times 4 \times 10^3 \times 0{\cdot}7}$

$$= \frac{0{\cdot}714}{2\pi \times 4 \times 10^3 \times 0{\cdot}7}$$

$$RC = \frac{714 \times 10^{-3}}{628 \times 4 \times 7} = 4{\cdot}06 \times 10^{-5}$$

or $R = 1000$ ohms.

$C = 4{\cdot}06 \times 10^{-8}$ farad.

$= 0\ 04 \mu f.$

Input resistance of diode detector : When a signal voltage is applied to the diode detector then the ratio of the signal voltage to the current drawn from the source is called the input impedance of the detector, the resistive component of this impedance, or the effective input resistance may be defined in terms of an equivalent resistor which would absorb the same power as the detector circuit when connected across the same. This may be obtained as follows :

Let the signal

$$E = E_c \cos \omega_c t$$

be applied to the diode detector. Then the average value of the power absorbed by the detector circuit is

$$\bar{P}=\frac{1}{2\pi}\int_{-\pi}^{+\pi}(I_a E_o \cos \omega_o t)\, d(\omega_o t)$$
$$=\frac{1}{2\pi}\int_{-\theta_1}^{+\theta_1}(I_a E_o \cos \omega_c t)\, d(\omega_o t)$$
$$=\frac{1}{\pi}\int_{0}^{\theta_1}(I_a E_c \cos \omega_o t)\, d(\omega_o t),$$

as the diode conducts from $-\theta_1$ to $+\theta_1$. I_a is the current through the diode and $2\theta_1$ is the angle of flow.

If it is assumed that the diode characteristic is straight and V_a be the direct potential difference across the load, then

$$\bar{P}=\frac{1}{\pi}\int_{0}^{\theta_1} E_o \cos \omega_o t.\ \frac{(E_o \cos \omega_o t-\bar{V}_0)}{r_a}.d(\omega_o t)$$
$$=\frac{E_o^2}{\pi r_a}\int_{0}^{\theta_1}\left(\cos^2 \omega_o t-\cos \omega_o t\ \frac{\bar{V}_0}{E_o}\right) d(\omega_o t)$$
$$=\frac{E_o^2}{\pi r_a}\int_{0}^{\theta_1}(\cos^2 \omega_o t-\cos \omega_o t \cos \theta_1)\, d(\omega_o t)$$

on using eq. (8). Therefore

$$\bar{P}=\frac{E_c^2}{\pi r_a}\left(\theta_1+\frac{\sin 2\theta_1}{2}-2 \cos \vartheta_1.\sin \theta_1\right)$$
$$=\frac{E_o^2}{2\pi r_a}(\theta_1-\cos \theta_1.\sin \theta_1) \qquad ...(21)$$

If R_{in} denotes the effective input resistance, then by definition,

$$\bar{P}=\frac{E_o^2}{2R_{in}} \qquad ...(22)$$

thus

$$\frac{E_o^2}{2R_{in}}=\frac{E_o^2}{2\pi r_a}(\theta_1-\cos \theta_1.\sin \theta_1)$$

or $$R_{in}=\frac{\pi r_a}{\theta_1-\cos \theta_1.\sin \theta_1} \qquad ...(23)$$

$$=\frac{\pi r_a}{\theta_1-\cos \theta_1\ \sqrt{(1-\cos^2\theta_1)}}=\frac{\pi r_a}{\cos^{-1}\beta-\beta\sqrt{(1-\beta^2)}} \qquad ...(24)$$

Eq. (23) can be put in the form

$$R_{in}=\frac{\pi r_a}{\theta_1-\frac{\sin 2\theta_1}{2}}$$

Since the angle of flow, $2\theta_1$, is small we can put

$$\sin 2\theta_1=2\theta_1-\frac{(2\theta_1)^3}{3\ !}$$

so that

$$R_{in}=\frac{\pi r_d}{\theta_1-\theta_1+\frac{4\theta_1^3}{3!}}$$

$$=\frac{3\pi r_d}{2\theta_1^3} \qquad \ldots(25)$$

Putting from eq. (10).

$$\theta_1^3=\frac{3\pi r_d}{R},$$

we get from eq. (25),

$$R_{in}=\frac{R}{2} \qquad \ldots(26)$$

The detector circuit is usually coupled to the preceding tuned amplifier stage. The effect of input resistance of the detector is to shunt an equivalent load across the tuned circuit of that amplifier and will decrease its dynamic impedance and selectivity. A decrease in the dynamic impedance of the tuned circuit also decreases the gain of the tuned amplifier. To minimize this shunting effect, therefore, input resistance of the detector should be as large as possible which, in view of eq. (26), require a large value of load resistance R.

Another form of relation (26) interms of detection efficiency, β can be deduced using eqs. (9) and (24). According to these relations

$$\pi r_d=R\,(\tan\theta_1-\theta_1)$$

and

$$\pi r_d=R_{in}\,(\theta_1-\cos\theta_1\sin\theta_1)$$

so that

$$\frac{R_{in}}{R}=\frac{\tan\theta_1-\theta_1}{\theta_1-\cos\theta_1\sin\theta_1}.$$

When efficiency β is high, θ_1 is small so that

$$\sin\theta_1=\theta_1-\frac{\theta_1^3}{6},\quad \cos\theta_1=1-\frac{\theta_1^2}{2}=\beta$$

Therefore

$$\frac{R_{in}}{R}=\frac{1}{\cos\theta_1}\cdot\frac{\sin\theta_1-\theta_1\cos\theta_1}{\theta_1-\cos\theta_1\sin\theta_1}$$

$$=\frac{1}{\beta}\,\frac{\left(\theta_1-\frac{\theta_1^3}{6}\right)-\theta_1\left(1-\frac{\theta_1^2}{2}\right)}{\theta_1-\left(\theta_1-\frac{\theta_1^3}{6}\right)\left(1-\frac{\theta_1^2}{2}\right)}$$

$$=\frac{1}{\beta}\,\frac{\left(\frac{\theta_1^3}{2}\cdots\frac{\theta_1^3}{6}\right)}{\left(\frac{\theta_1^2}{2}+\frac{\theta_1^3}{6}\right)}$$

$$= \frac{1}{2\beta}$$

or
$$R_{in} = \frac{R}{2\beta} \qquad ...(27)$$

Since β is already taken to be quite high, input resistance can be made large by choosing a large value of load resistance in order to minimize the aforesaid shunting effect.

31·4. DETECTION CHARACTERISTICS OF DIODE AND ITS USES :

Circuit for obtaining the detection characteristics, which are sometimes called as dynamic characteristics, is shown in fig. 13.

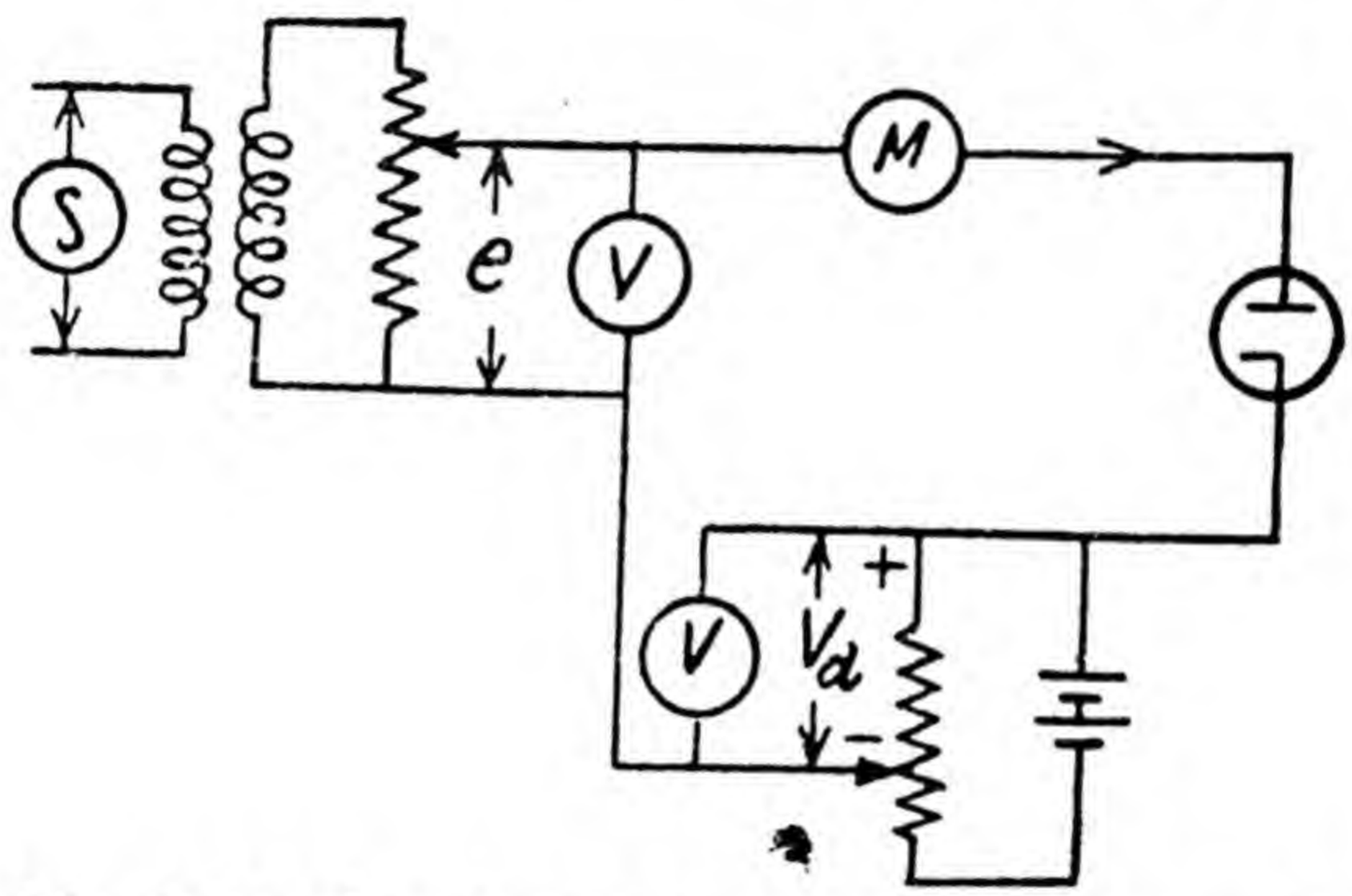


Fig. 13. Circuit for drawing dynamic characteristics of diode.

A direct component V_d to anode potential is given by a battery and potential divider. This direct component V_d takes the place of average potential difference $\bar{V}_0$ across the load circuit of the detector. A signal voltage e is taken from the mains. We keep the potential difference V_d constant and is independent of signal voltage e. The direct current microammeter M measures the average current I through diode.

The dynamic characteristics are drawn by plotting the average current $\bar{I}$ through the diode against the direct potential difference V_0 for a fixed signal amplitude $\hat{e}$. For different curves other values of $\hat{e}$ are chosen.

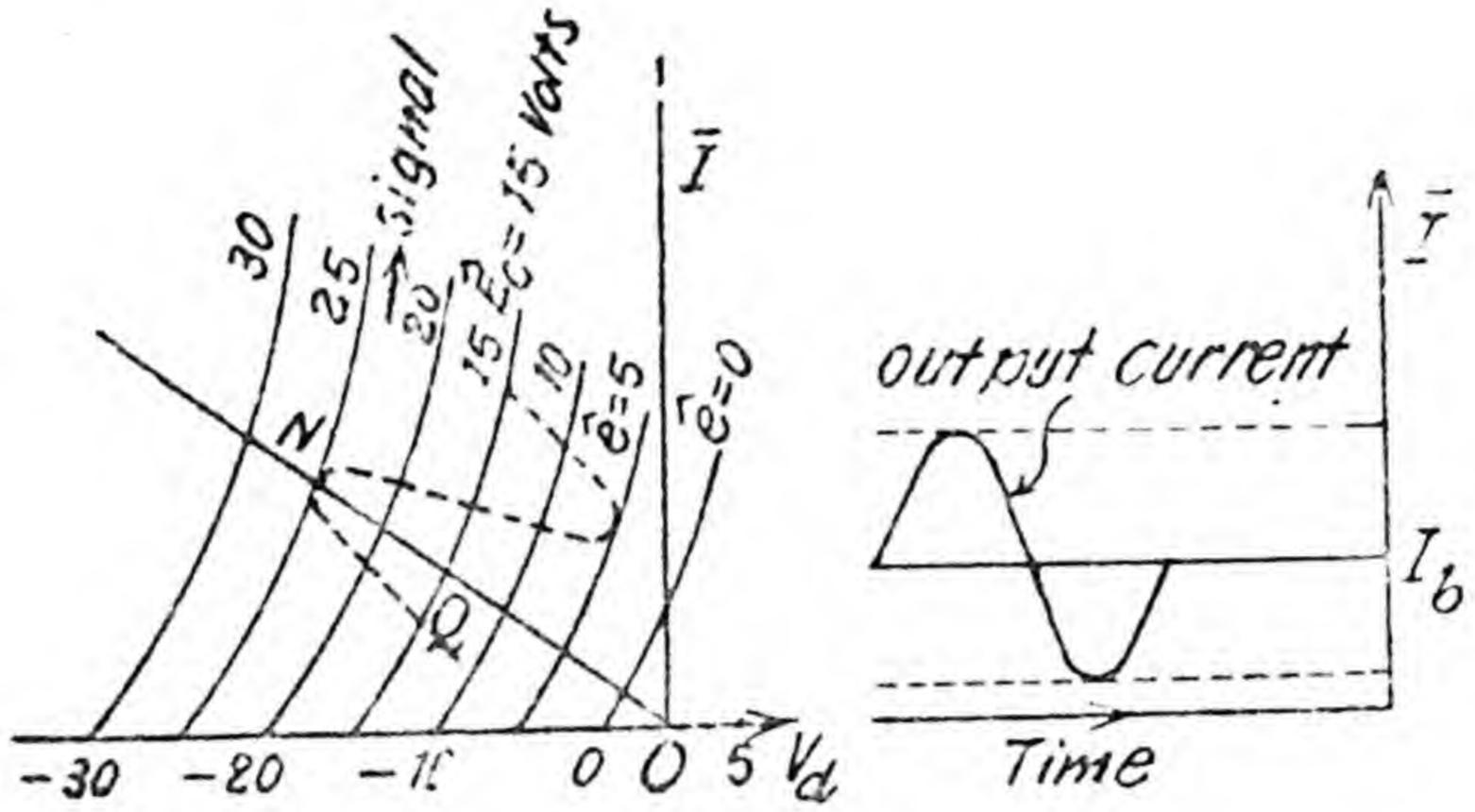
Figure 14. shows the dynamic characteristics of the diode.

Uses of detection characteristics : The performance of a detector can be analysed by means of these characteristics :

(*i*) Calculation of efficiency of detection (β) :

The average potential difference $\bar{V}_0$ across R will be given by

$$\bar{V}_0 = R\bar{I}.$$



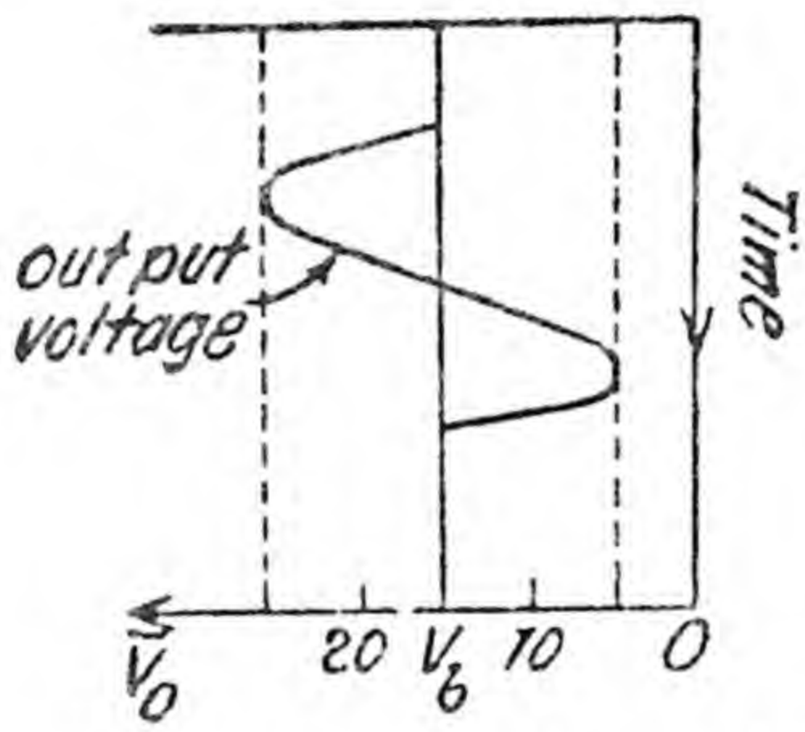


Fig. 14 Detection dynamic) characteristics of diode.

Since this gives a negative bias to the anode of the valve, it should correspond to direct component V_d of anode potential.

$$V_d = -R\bar{I}$$

The relation can be represented in figure 14 by drawing a load line for the resistance R. The different points of intersection with detection characteristics give different values of V_d or $\bar{V}_0$ and $\hat{e}$ From these values detection efficiency can be calculated since

$$\beta = \frac{\bar{V}_0}{\hat{e}}.$$

(*ii*) Point Q is called operating point of the detection for the carrier amplitude concerned. The average value of current $\bar{I}$ and voltage $\bar{V}_0$ moves on either side of Q along the load line. We can trace the waveforms of $\bar{I}$ and V_0 from the intersections of the load line with the detection characteristics as shown in fig. 14. These waveforms represent a true picture of the detector's performance when the time constant CR of load circuit is small.

(*iit*) We can estimate the distortion from these characteristics. For a fixed value of depth of modulation, distortion decreases with the increase of carrier amplitude $\hat{E}_c$.

Distortion in linear diode detector :

The two main reasons for the distortion in linear diode detector are :

(*i*) Improper selection of time constant *RC* introduces distortion. With a too low value of *RC*, radio frequency component will not be eliminated fully from the output voltage and thus causing large fluctuations in the output waveform. On the other hand, if *RC* is too high, distortion results due to clipping during the negative peaks of the modulation wave.

(*ii*) Distortion also arises due to curvature of the tube characteristics. Due to this curvature, the efficiency of detection varies inaccordance with the amplitude of modulation envelope. This distortion can be minimised if detection efficiency, β, is made sufficiently high for which

(*a*) load resistance should be large as compared to plate resistance of diode, and

(*b*) carrier voltage of large envelope amplitude should be applied.

31·5. DIODE DETECTOR FOR AUTOMATIC VOLUME CONTROL :

In order to minimise fading in receiving sets automatic volume control becomes necessary, In a.v.c., the gain of pre-amplifying stage is varied such as to give uniform response at the receiver output.

Automatic volume control is secured by using the direct component of the output voltage of diode detector to control the gain of r. f. amplifiers. These amplifiers use variable μ valves whose amplification decreases as their negative grid bias increases.

From equation (13*a*) of § 31·3-2 the direct component of potential difference across the load is given by

$$V_b = \beta E_c.$$

Thus V_b totally depends upon the amplitude of carrier and lowers the gain, when applied to the grid of the amplifier, as carrier amplitude rises.

Fig. 15 (*a*) shows the conventional a.v.c. circuit. The direct component V_b is derived from detector's output. R_2–C_2 acts as a

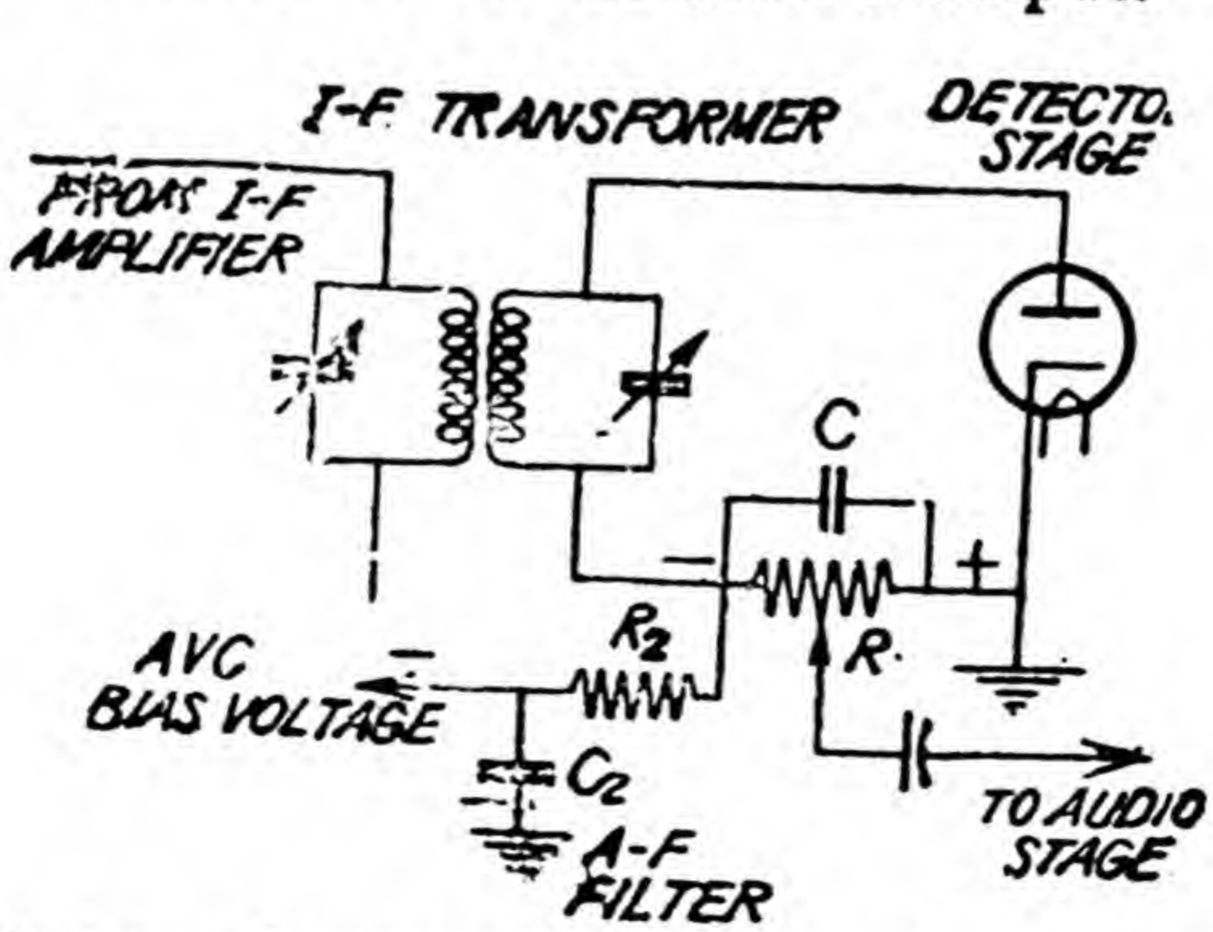


Fig. 15. (*a*) Circuit for automatic volume control.

low pass filter to separate the direct voltage component of the detector, V_b from the a c. components. The a.v c. voltage is developed across C_2. The values of R_2 and C_2 are large so as to minimise any change on the charge of capacitor C_2 due to a.c. component

When incoming signal is weak, then due to the presence of a.v.c. voltage r. f. amplifiers operate with a very weak output at which the sensitivity of the receiver is insufficient and low. Therefore it is rather better to use delayed a.v.c. circuit which acts only when the carrier wave has any amplitude above the delay level.

Delayed AVC : Circuit is shown in fig. 15 (*b*). A dudiode triode tube is used When the diode D_1 works, its output consists of modulation frequency component V_1 and rectified d.c. voltage V_b. V_1 and V_b are then applied to the control grid of triode T so that modulation frequency component is amplified and is available at the output, while the d.c. voltage makes the grid more negative, thereby, *decreasing* the d.c. component of plate current of triode, *i.e.* decreasing the d.c. voltage across R_K.

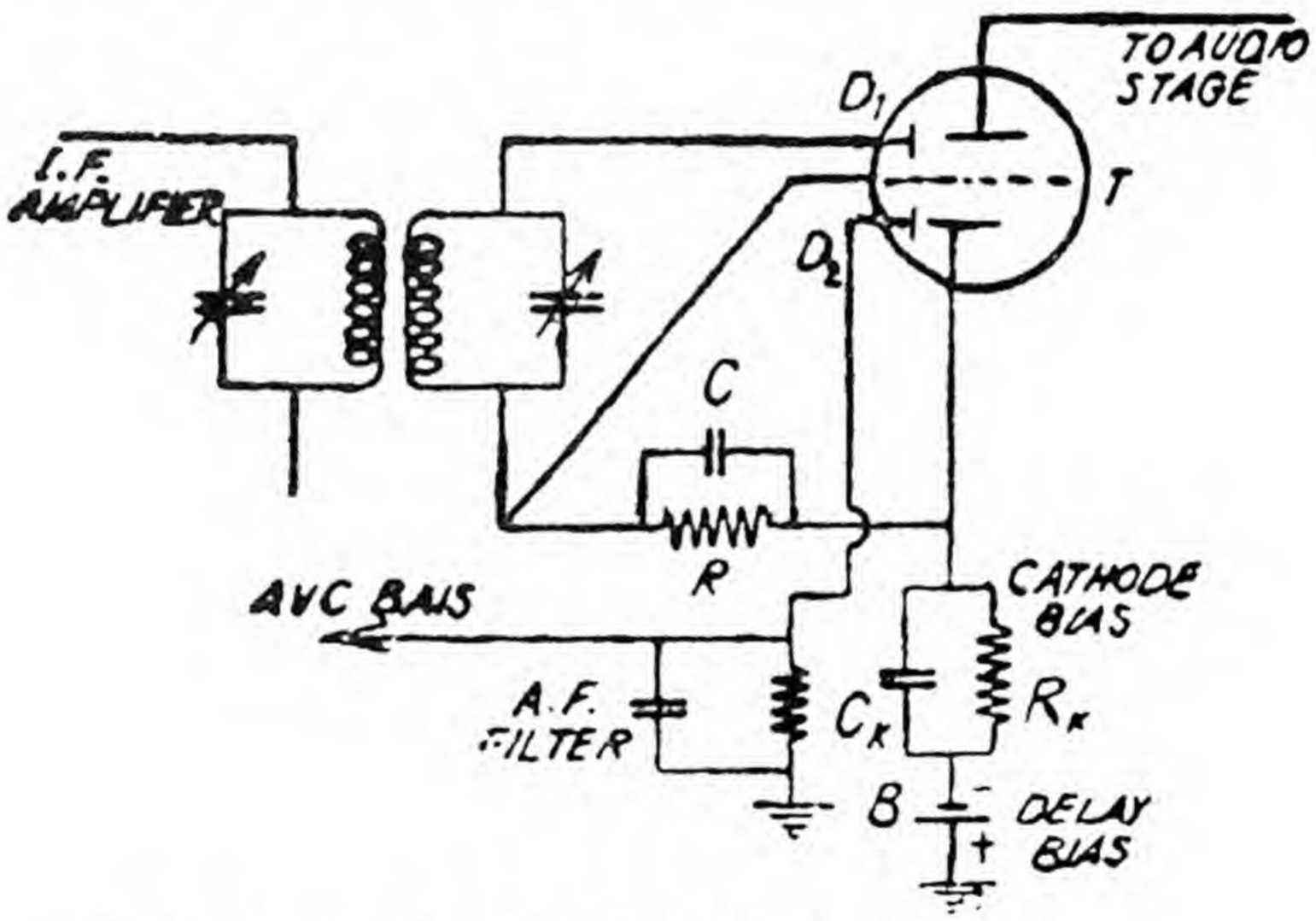


Fig. 15. (*b*) De'ayed AVC Circuit.

With zero or small carrier input, d.c. component of plate current develops sufficient voltage across R_K to make cathode positive to ground by an amount depending upon the delay level. Since plate of diode D_2 is grounded, it would not operate so long as the cathode is positive with respect to ground. With the increase in carrier input voltage, negative bias increases so that d.c. component of plate current of triode decreases resulting in the decrease of d.c. voltage across R_K and hence cathode becomes less positive with respect to ground. The process continues till the carrier amplitude rises to the level set by delay bias at which cathode becomes negative and then diode D_2 conducts. This sets up the normal *AVC* action. Obviously *AVC* acts only when carrier amplitude rises to a definite level.

A transistor AVC circuit: Refer to fig. 15(*c*). When the signal level increases, the emitter current increases and point *X* becomes more negative. This provides *AVC* voltage. Combination R_3 C_3 acts as filter to remove any modulation from the control

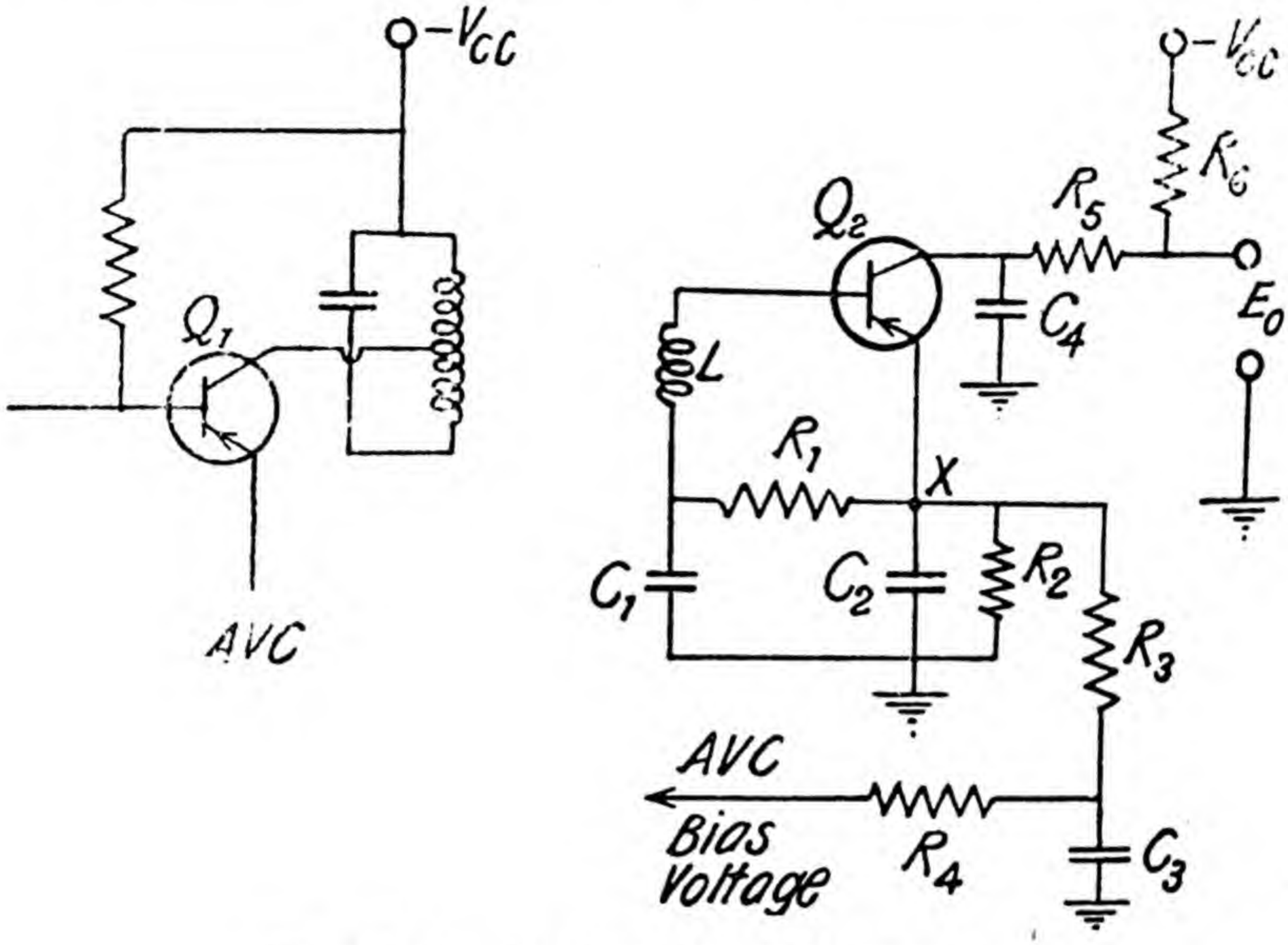


Fig. 15 (c). A transistor AVC circuit.

voltage. This voltage is applied to the emitter of Q_1 (or other *RF* or *IF* amplifier) through a resistor R_4, in order to provide a suitable control current. If signal level reaching the detector tends to increase, the emitter current of Q_2 increases. This makes point *X* more negative and thus emitter of Q_1 would become more negative. Since Q_1 is a *PNP* transistor, this action opposes the forward bias so that its emitter current and the corresponding stage gain are reduced, thus tending to maintain a constant level of the signal at the detector.

A diode AVC circuit and delayed AVC circuit : A diode detector circuit from where *AVC* voltage is taken is shown in fig. 15(*d*). *AVC* voltage is taken from diode load resistor, point 2, to ground. Rectified voltage at point 2, consists of a modulating signal voltage ($\beta E_c m_a \sin \omega_m t$, see eq. 12 art 31·3-2) and a d.c. component (βE_c). R_3 C_4 filter circuit is used to remove modulating signal from *AVC* voltage. Capacitor C_4 has a low reactance (compared to R_3) at *modulating* frequency so that it bypasses the modulating signal to ground, while blocks the d.c. component (βE_c). Thus only d.c. component voltage appears across R_3 and is available as *AVC* voltage to be applied to earlier stages. Note that function of C_3 capacitor is to prevent this d.c component from reaching the next stage (audio frequency amplifier) which needs modulating signal only

Addition of an *AVC* circuit has an adverse effect on the operation of diode detector, because R_3 is effectively in parallel with R_1

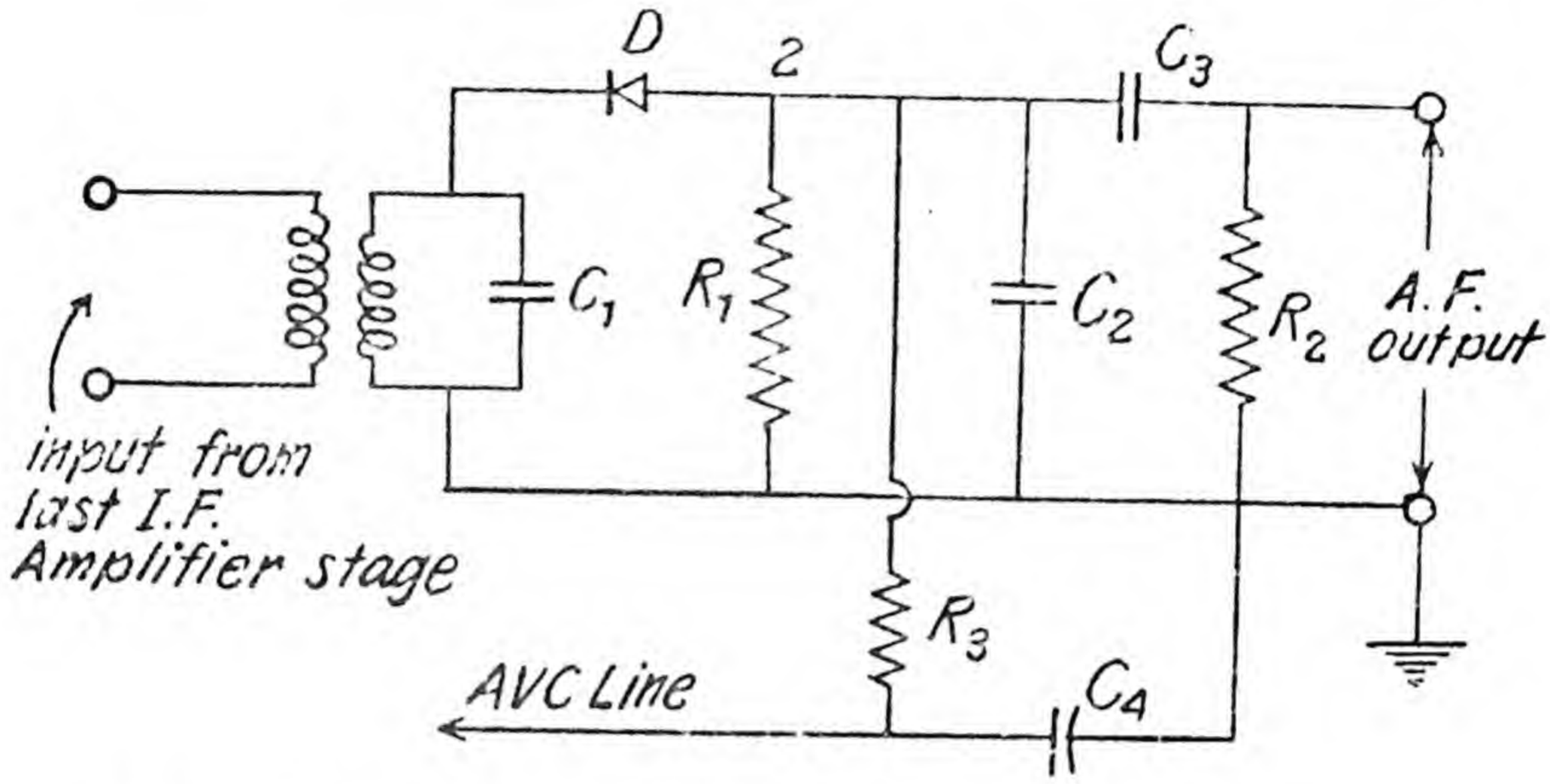


Fig. 15 (*d*) Diode detector stage of A.M. receiver from where AVC voltage is derived.

reducing *a.c.* load resistance. This can cause amplitude distortion at high modulation levels. It is therefore significant to keep R_3 large compared to R_1 but too large a value will increase the time constant $R_3\ C_4$ which will make *AVC* bias unable to respond to rapidly varying signal inputs. Generally $R_3\ C_4$ is chosen in the range 0·1—0·5 second. This effect of loading by *AVC* circuit can be removed if *AVC* function is separated from the detector circuit. Fig 15(*e*) shows a separate *AVC* circuit. Suppose the cathode of

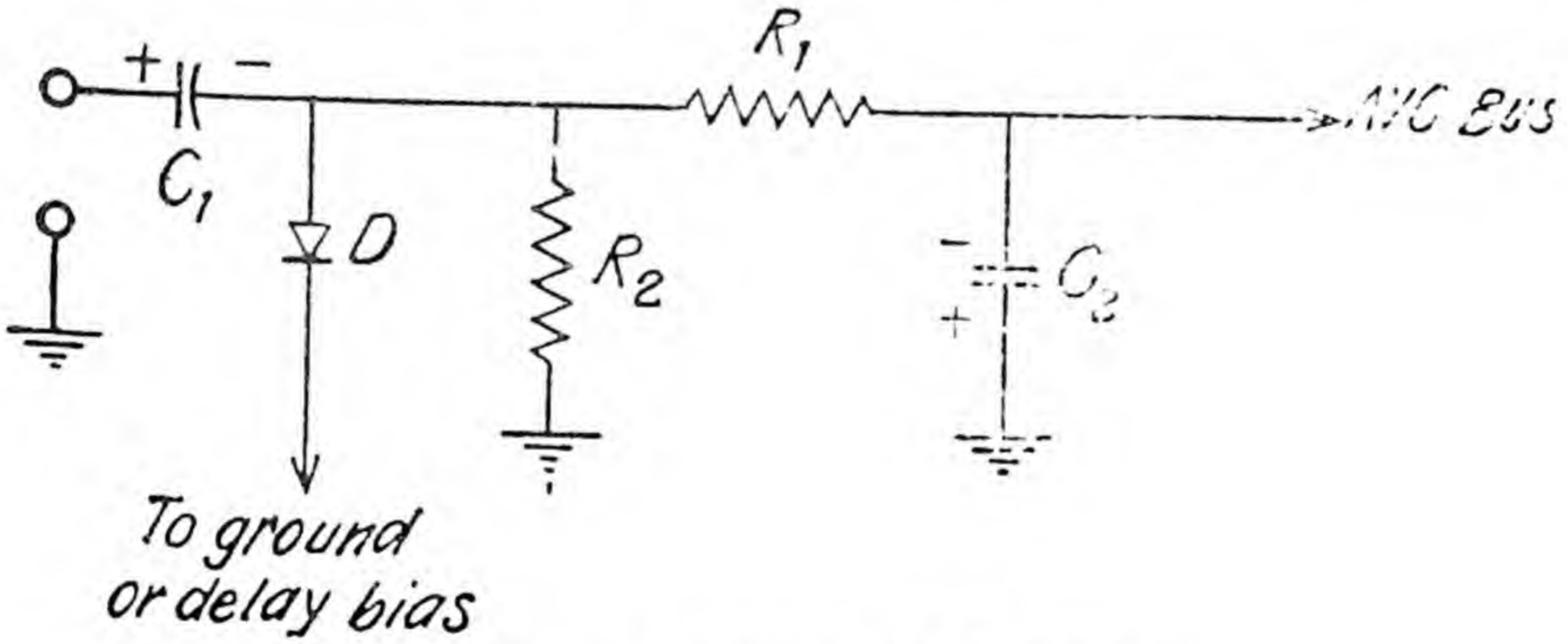


Fig. 15 (*e*). A separate AVC circuit.

diode is grounded (see delayed *AVC* action in which cathode is connected to a positive bias voltage above which *AVC* action is desired), then diode conducts during the first half of the positive cycle of r.f. input signal, and capacitor C_1 charges to peak r.f. value with polarity indicated. As the r f. input signal, decreases, capacitor *cannot* discharge through the diode but through R_2 (and also through $R_1\ C_2$). Because time constant $C_1\ R_2$ (or $R_1\ C_2$) is

large, condenser discharges only slightly. Due to the slight discharge of C_1, condenser C_2 pick up a small charge with polarity as shown in the fig. 15 (*e*). The whole action is repeated (again diode will conduct at the positive half cycle of input r.f. signal, capacitor C_1 charges to peak r.f. value, it then discharges giving more charge to capacitor C_2) and *after few cycles, charge on* C_2 *is built up to full value, producing AVC bias voltage.* Time constant $R_1 C_2$ is quite large and therefore r.f. ripples and modulating signal ($\beta E_c m_a \sin \omega_m t$) do not cause fluctuation in *AVC* bias voltage.

When delayed *AVC* is desired, cathode of diode is connected to positive bias voltage corresponding to delayed level overwhich we desired *AVC* to operate. Suppose it is connected to $+4V$. Then the diode cannot conduct if the input r.f. signal level is less than $4V$ at its peak. Meanwhile capacitor C_2 will discharge through R_1 and R_2, the *AVC* bias drops to zero and thus at weak input signals *AVC* does not operate and receiver remains sensitive for such weak signals. But for signals having peak value above $4V$, *AVC* action starts as explained before.

SOME SOLVED PROBLEMS

Example 1. *A diode detector circuit with an efficiency of* 80% *represents an effective load of* 0·3 $M\Omega$ *across the input source of modulated wave. Calculate the values of resistance and capacitance to be used with the diode for modulated signals having modulation frequencies upto* 5 *kc/s and a maximum depth of modulation of* 80%.

Given that $\beta = 80\%$, $R_{in} = 0{\cdot}3$ meg. ohm $m_a = 0{\cdot}8$, $f_m = 5$ kc/s

Load resistance R can be found from eq. (27) which is

$$R_{in} = \frac{R}{2\beta}$$

or
$$R = R_{in}\, 2\beta = 0{\cdot}3 \times 10^6 \times 2 \times 0{\cdot}8$$
$$= 0{\cdot}48 \times 10^6 \text{ ohms.}$$

Value of the capacitor can be found from relation (18) art. 3 ·3-2 which is, for maximum depth of modulation,

$$\frac{1}{\omega_m CR} = \frac{m_a}{\sqrt{(1-m_a^2)}}$$

or
$$C = \frac{\sqrt{(1-m_a^2)}}{m_a \omega_m R} = \frac{\sqrt{[1-(0{\cdot}8)^2]}}{0{\cdot}8 \times 2\pi \times 5 \times 10^3 \times 0{\cdot}48 \times 10^6}$$
$$= 49{\cdot}75\ pF$$

Example 2. *The components of an envelope diode detector are* $R = 200K\Omega$ *and* $C = 100pF$. *Determine the maximum percentage of modulation if diagonal clipping is to be avoided i.e. for no distortion for an audio frequency of* 10 *kc/s.*

Given ; $R = 200 \times 10^3\ \Omega$, $C = 100 \times 10^{-12}\ F$,
$f_m = 10 \times 10^3$ kc/s

Maximum depth of modulation is given by

$$\frac{1}{\omega_m CR} = \frac{m_a}{\sqrt{(1-m_a^2)}}$$

$$\frac{1}{2\pi \times 10^4 \times 200 \times 10^3 \times 100 \wedge 10^{-12}} = \frac{m_a}{\sqrt{(1-m_a^2)}}$$

$$\frac{10}{4\pi} = \frac{m_a}{\sqrt{(1-m_a^2)}}$$

or $257{\cdot}75\ m_a^2 = 100$

or $m_a^2 = 0{\cdot}387$

or $m_a \approx 0{\cdot}62.$

Example 3. *In an ideal envelope detector the tube has a conductance of 2mA/V. The load is parallel combination of a resistor and capacitor such that when a voltage 20 sin ωt is applied in the input, the conduction begins at cot 60°. Find the value of the resistor and the d.c. voltage developed.*

Using eq. (9) art. 31·3-32,

$$\tan\theta_1 - \theta_1 = \frac{\pi r_d}{R},$$

we can find the value of load resistor R. We are given that transconductance of the tube is $2mA/V$. Therefore

$$r_d = \frac{1}{g_m} = \frac{1}{2\times 10^{-3}} = 500\Omega.$$

Further conduction begins at 60° and will stop at 120° so that angle of flow

$$2\theta_1 = 120 - 60 = 60°$$

or $\theta_1 = 30° = 0{\cdot}5230$ in rad.

The load resistance is then

$$R = \frac{3{\cdot}14 \times 500}{\tan 30° - 0{\cdot}5230}$$

$$= \frac{3{\cdot}14 \times 500}{0\ 5774 - 0{\cdot}5230}$$

$$= 29{\cdot}20\ K\Omega.$$

From fig. 11, it is obvious that diode starts conduction only when the input signal voltage is just equal to the d.c. voltage across the load. Therefore

d.c. voltage developed across load $= 20 \sin 60°$

$= 17{\cdot}32\ V.$

Example 4. *The input voltage to an envelope diode detector is* $10(1+0{\cdot}5 \sin \omega_m t) \sin \omega_o t$ *volts. If conduction angle is 45°, find d.c. and a.c. output voltages.*

From equations (14*a*) and (14*b*) art. 31·3-2, we find that d.c. output is

$$V_b = \frac{RE_c}{\pi r_d}(\sin\theta_1 - \theta_1\cos\theta_1)$$

and a.c. output is

$$V_1 = \frac{RE_c m_a}{\pi r_d}(\sin\theta_1 - \theta_1\cos\theta_1).$$

It is given that modulated input

$$e = 10(1+0{\cdot}5\sin\omega_m t)\sin\omega_c t.$$

so that

$$E_c = 10 \text{ volts and } m_a = 0{\cdot}5.$$

In order to calculate $\frac{R}{r_d}$, we use eq. 2 art. 31·3-2 which is

$$\frac{R}{r_d} = \frac{\pi}{\tan\theta_1 - \theta_1},$$

with conduction angle $\theta_1 = 45^c/2 = 22^\circ\ 30''$, we get

$$\frac{R}{r_d} = \frac{3{\cdot}14}{\tan 22{\cdot}5^\circ - 0\ 3927}$$

$$= \frac{3{\cdot}14}{0{\cdot}4142 - 0{\cdot}3927} = \frac{3{\cdot}14}{{\cdot}0215}$$

$$= 146\ 05.$$

Therefore d.c. output

$$V_b = \frac{10\times 146{\cdot}05}{3{\cdot}14}(\sin 22{\cdot}5^\circ - 0{\cdot}3927\cos 22{\cdot}5^\circ)$$

$$= 465{\cdot}3\ (0{\cdot}3827 - 0{\cdot}3927\times 0{\cdot}9239)$$

$$= 465{\cdot}3\times 0{\cdot}0207$$

$$\approx 9{\cdot}31 \text{ volts.}$$

The a.c. output voltage is

$$V_1 = \frac{RE_c}{\pi r_d} m_a(\sin\theta_1 - \theta_1\cos\theta_1)$$

$$= V_b m_a$$

$$= 9{\cdot}31\times 0{\cdot}5 = 4{\cdot}655 \text{ volts or } \left(\frac{4{\cdot}655}{\sqrt{2}} \text{ r.m.s.}\right)$$

Example 5. *In a linear grid detector the range over which the grid can operate without undue distortion is 5 volts. If the carrier is modulated to a depth of 0·7, determine the greatest possible peak value of the audio input voltage.*

This implies that if grid is biased −2·5 volts, then maximum swing of the modulated wave for no distortion *about* grid bias should be only 2·5 volts. Then refer to fig. 2 chapter 30, we, therefore, put

$$E_{max} = 2{\cdot}5 \text{ volts.}$$

Further $$m_a = \frac{E_{max} - E_{min}}{E_{max} + E_{min}}$$

or $$0{\cdot}7 = \frac{2{\cdot}5 - E_{min}}{2{\cdot}5 + E_{min}}$$

or $$E_{min} = 0{\cdot}44 \text{ volts.}$$

Therefore greatest possible peak value of the modulating voltage will be

$$= \frac{E_{max} - E_{min}}{2}$$

$$= \frac{2{\cdot}5 \times 0{\cdot}44}{2} = 1{\cdot}03 \text{ volts.}$$

31·6 FREQUENCY DEMODULATION :

For the detection of f.m. waves, the method generally used is first to convert frequency modulation to amplitude modulation which is then subjected to conventional methods of detection. Circuits in which the instantaneous magnitude of output voltage varies inaccordance with the instantaneous frequency variat on of the input voltage, will serve the above purpose. Such circuits are named as discriminator.

31·7. DISCRIMINATOR :

Vacuum tube circuit is shown in fig. 16 (*a*).

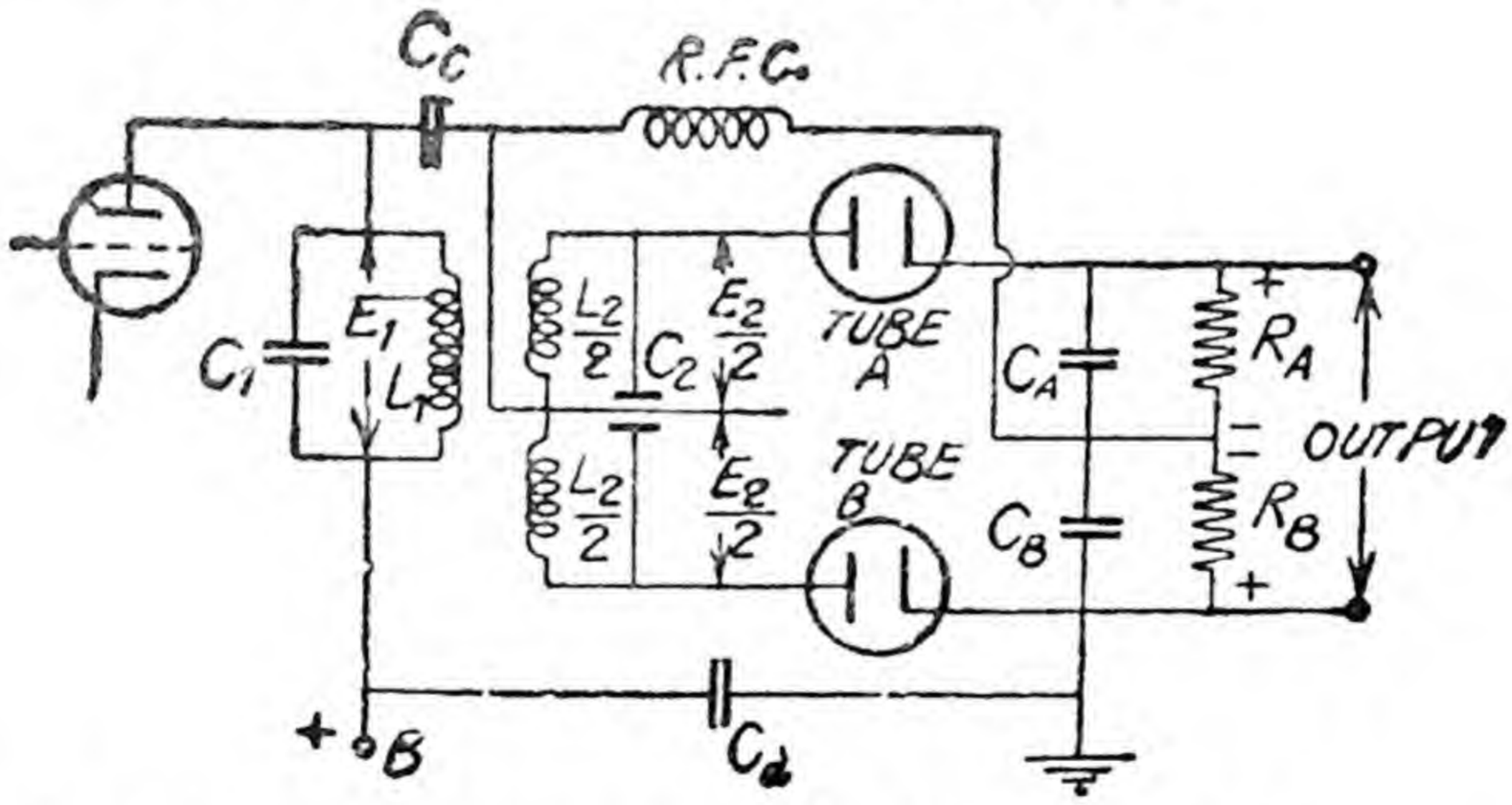


Fig. 16 (*a*) Discriminator circuit for the detection of f.m. wave.

Description of Components : C_1–L_1, which constitutes the tank circuit of an amplifier stage, is tuned to the centre frequency of modulated wave. This tuned circuit is coupled to another tank circuit C_2L_2 tuned to the same centre frequency. The centre tap of L_2 and the top of the tuned circuit L_1C_1 are kept at the same radio frequency potential with respect to the ground by means of a coupling condenser C_c.

Condenser C_A and C_B of output circuit are sufficiently large so as to keep the cathodes of both diodes at the same r.f. ground

potential. Resistances R_A and R_B are of equal value. The polarity of voltage across them is shown in fig. 16 (*a*).

Transistor discriminator circuit is shown in fig. 16 (*b*).

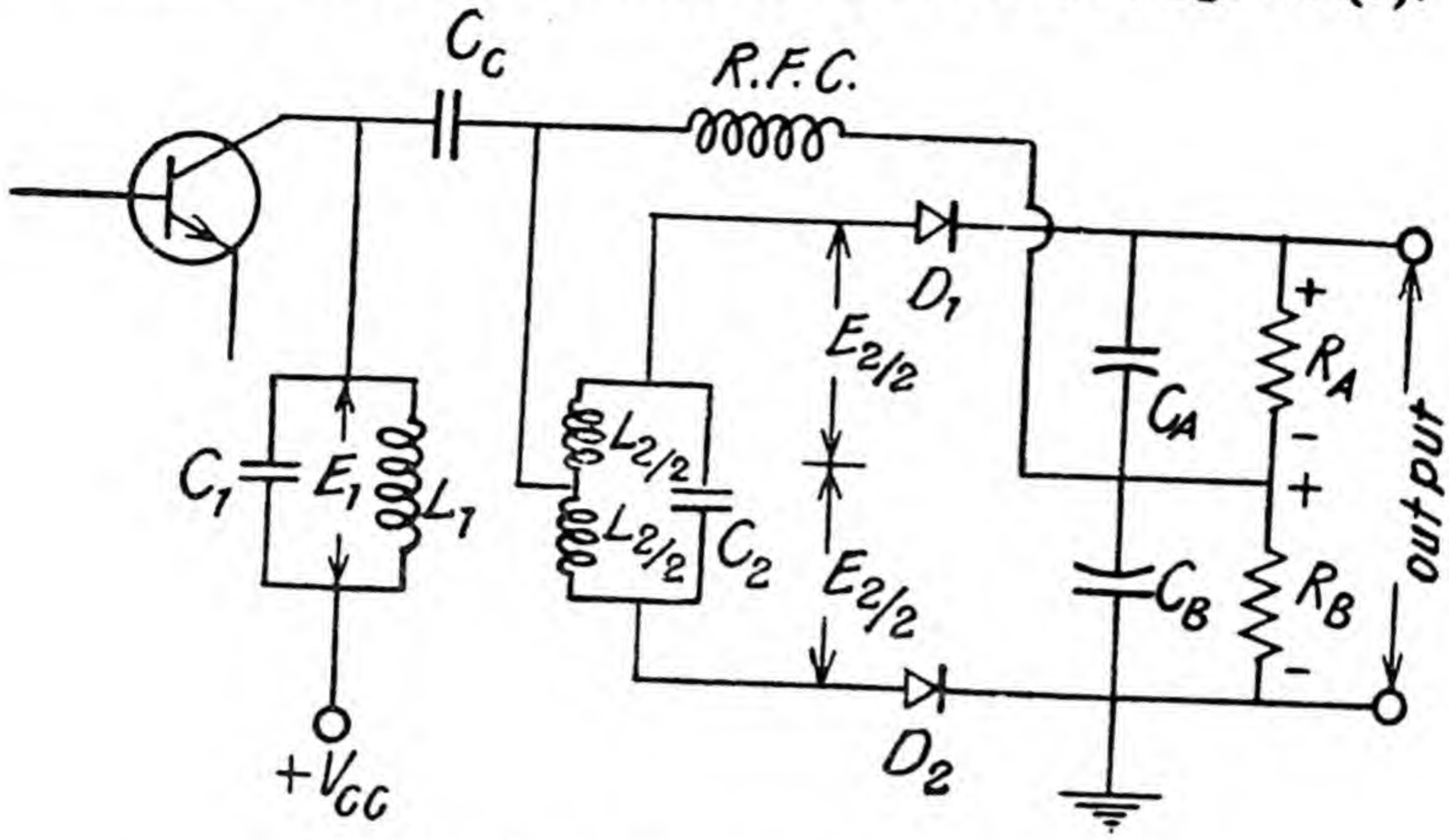


Fig. 16. (*b*) Discriminator.

Operation : When r.f. voltage is induced in L_2C_2 circuit, then the developed voltage E_2 across tank circuit will be applied to the plates of the diodes *A* and *B*. A rectified current from both diodes will flow in their respective output circuits. Three cases may arise:

(*a*) when the r.f. voltages E_A and E_B at the plates of tubes *A* and *B* with respect to ground are equal, the rectified current from both diodes will be of equal value. Therefore, the direct voltages across R_A and R_B will be equal. The polarity of voltages across R_A and R_B show that when $E_A = E_B$, the net output of discriminator is zero.

(*b*) if $E_A > E_B$, then voltage across R_A will be greater than that across R_B and an output voltage, positive with respect to ground, will be obtained.

(*c*) when $E_A < E_B$, the magnitude of voltage across R_B exceeds that across R_A and a negative output voltage will result.

Thus, if relative magnitudes of E_A and E_B be made to depend upon the instantaneous frequency of the input wave, then output voltage, as discussed above, will itself vary according to the frequency variations of input signal.

We shall discuss below the three cases in detail :

Case (a) *When input has only centre frequency* (*i.e., no modulation*).

Circuits C_1L_1 and C_2L_2 are tuned to centre frequency. The phase relations in voltages and currents in the two circuit components at the resonant frequency will be as shown in vector diagram 17.

With voltage E_1 in C_1L_1 circuit, the current in inductance L_1 (I_{L_1}) will lag E_1 by 90°. The induced voltage in C_2L_2, $-j\omega M I_{L_1}$ will be 90° out of phase with I_{L_1}. The current I_2 is flowing

through capacitor C_2 produces a voltage E_2 across C_2 which lags I_2 by 90°. Thus E_2 voltage across the terminals of circuit C_2L_2 leads E_1, the voltage across circuit C_1L_1, by 90°.

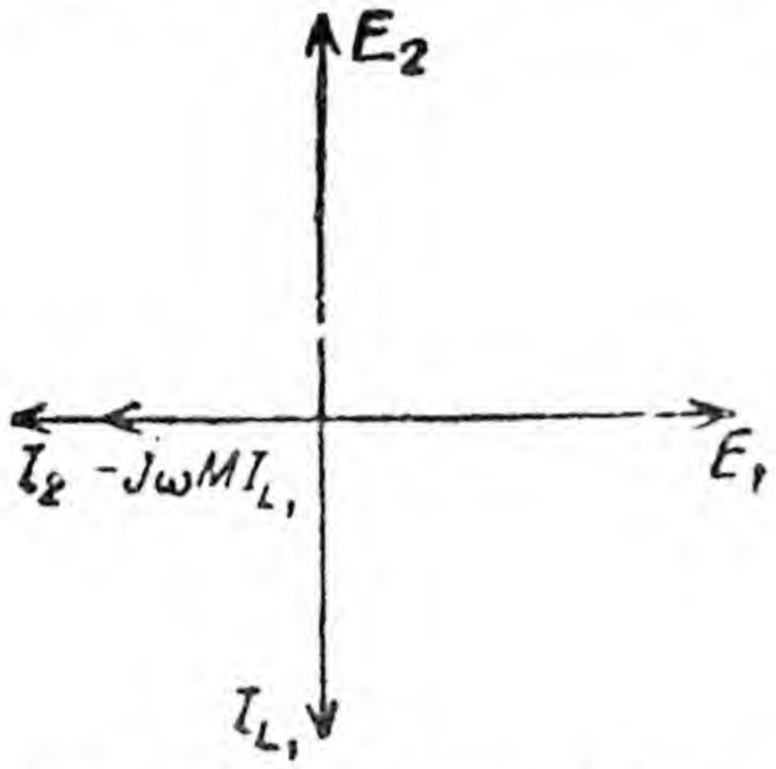


Fig. 17 (*i*) Phase relation in discriminator circuit for case (*a*).

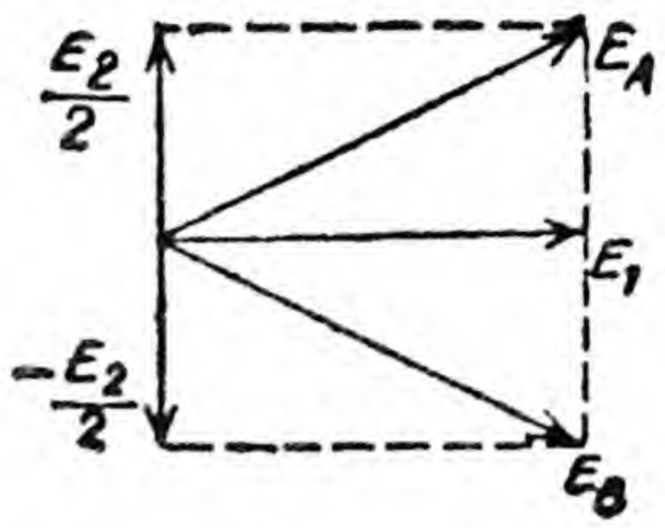


Fig. 17 (*ii*). Representation of relative magnitudes E_A and E_B.

It can be seen from fig 17 (*ii*) that r.f. voltages E_A and E_B across the two tubes will be made up of two components :

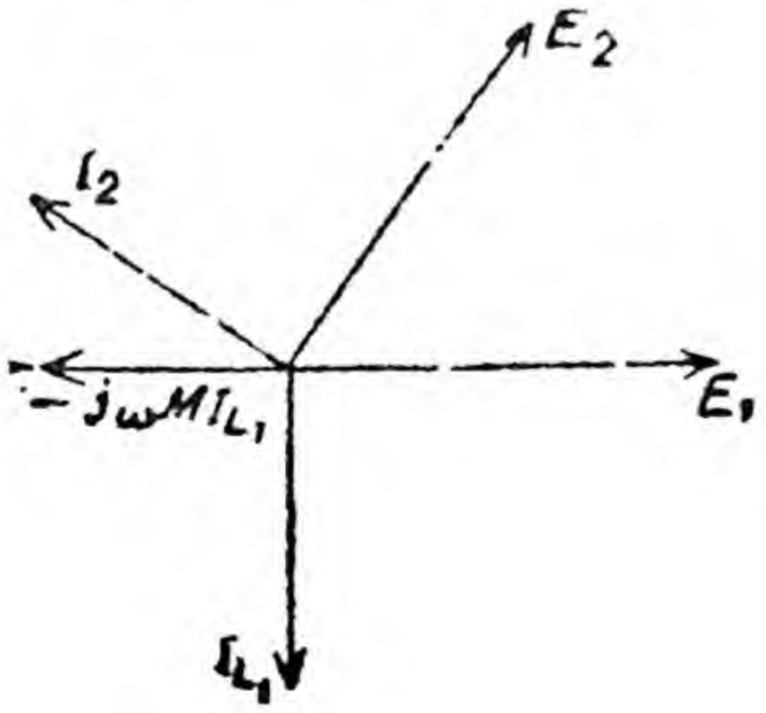
$$E_A = E_1 + \frac{E_2}{2},$$

and $$E_B = E_1 + \left(-\frac{E_2}{2}\right).$$

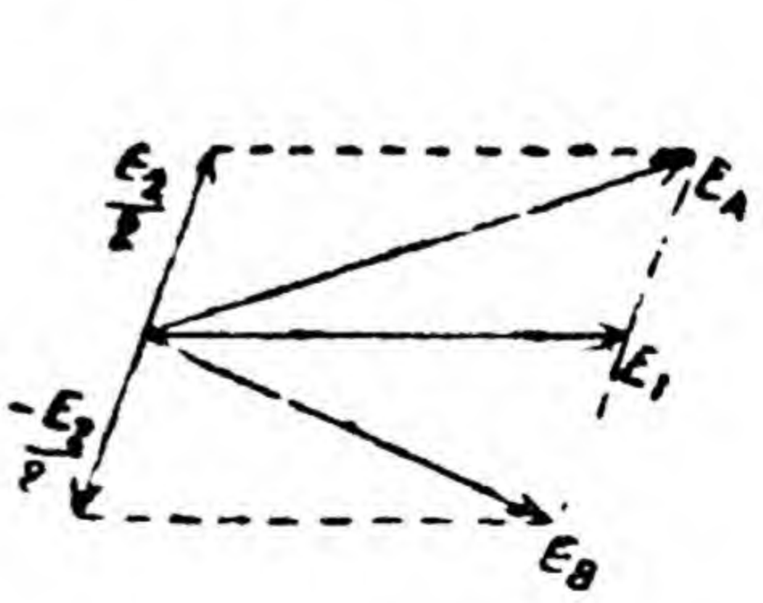
Since E_A and E_B are equal in magnitude, the output rectified currents from both tubes will be the same which set up equal potentials across R_A and R_B. Therefore, the discriminator output will be zero.

Case (b) : *When impressed frequency is greater than centre frequency of modulated wave :*

Above resonance, the secondary circuit represents an inductive reactance, so that I_2 lags $-j\omega MI_{L_1}$. The voltage E_2 across the capacitor C_2 still lags I_2 by 90°, but the phase angle between E_2 and E_1 is now less than 90° as shown in fig. 18 (*i*).



(*i*)



(*ii*)

Fig. 18. Discriminator phase relations when impressed frequency is greater than centre frequency.

It is obvious from fig. 18 (*ii*) that $E_A > E_B$ which means that the rectified current by the two tubes will be different and so the voltages across R_A and R_B, being different in magnitude, will develop a resultant positive voltage at this discriminator output.

Since the phase angle between E_1 and E_2 depends upon the deviation of the impressed frequency from the resonant frequency of the tuned circuits, the magnitudes of E_A and E_B will also vary inaccordance with the deviation. The output voltage which depends upon the relative magnitudes of E_A and E_B will also follow the deviation of the impressed frequency from the centre frequency or in other words the output of discriminator will vary at audio frequency rate.

Case (c) : *When impressed frequency is less than the centre frequency of modulated wave :*

In this case, the secondary circuit represents a capacitive reactance. Current I_2 will lead $-j\omega MI_{L_1}$ by an angle depending upon the deviation of impressed frequency from centre frequency of modulated wave. Therefore the angle between E_2 and E_1 will be more than 90° as shown in figure 19 (*i*). As a result, E_B will have a larger magnitude than E_A.

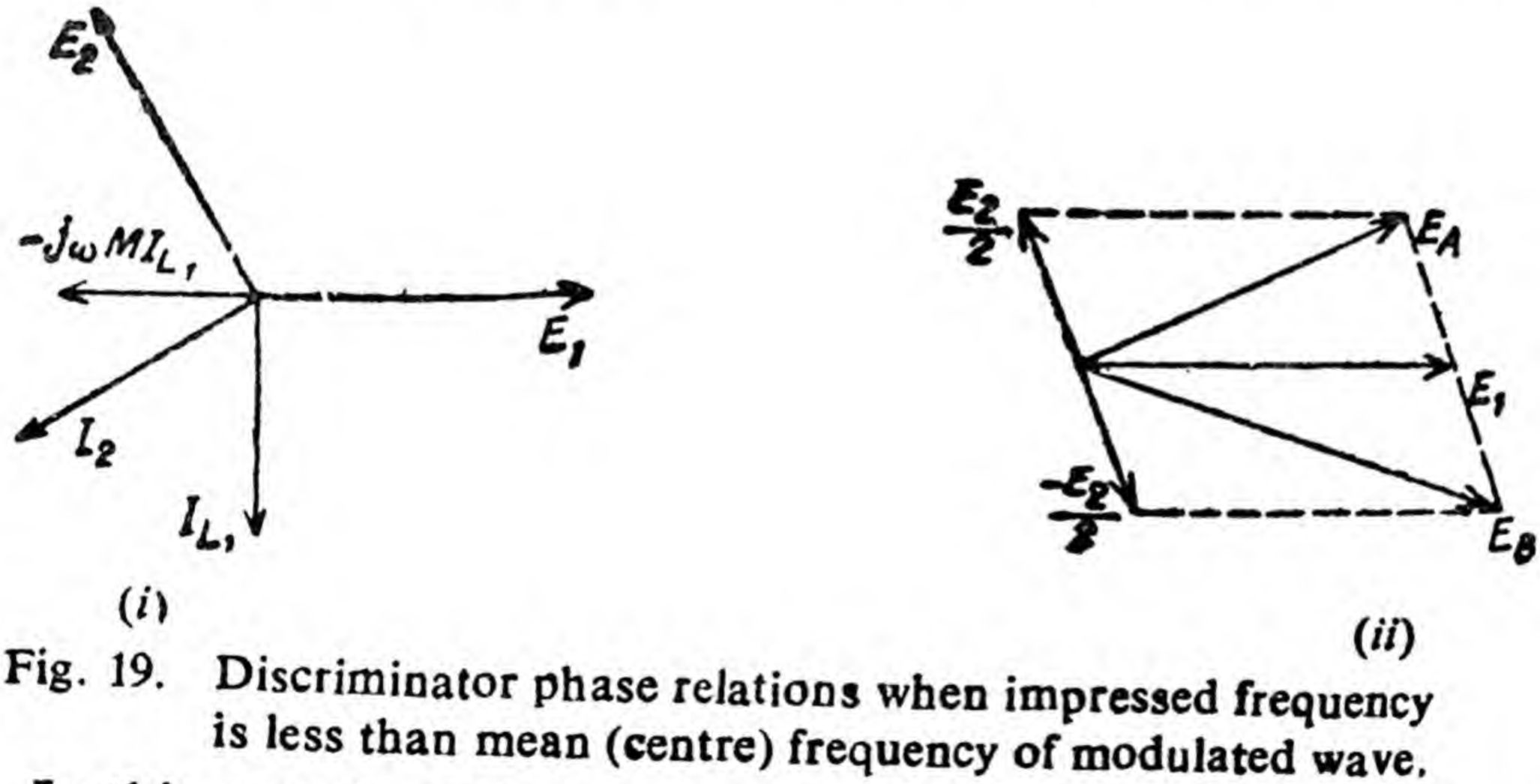


Fig. 19. Discriminator phase relations when impressed frequency is less than mean (centre) frequency of modulated wave.

In this case the voltage across R_B increases in comparison to voltage across R_A and the output of discriminator will be negative

Mathematical analysis of discriminator :

Fig. 20 gives the inductively coupled portion of the discriminator. R_1 and R_2 are the resistances associated with inductances L_1 and L_2.

The e.m.f. equations for the two circuits can easily be expressed as

$$j\frac{I_1}{\omega C_1}+\left[R_1+j\left(\omega L_1-\frac{1}{\omega C_1}\right)\right]I_1-j\omega MI_2=0. \qquad \ldots(1)$$

and $$\left[R_2+j\left(\omega L_2-\frac{1}{\omega C_2}\right)\right] I_2-j\omega M I_1=0. \qquad \ldots(2)$$

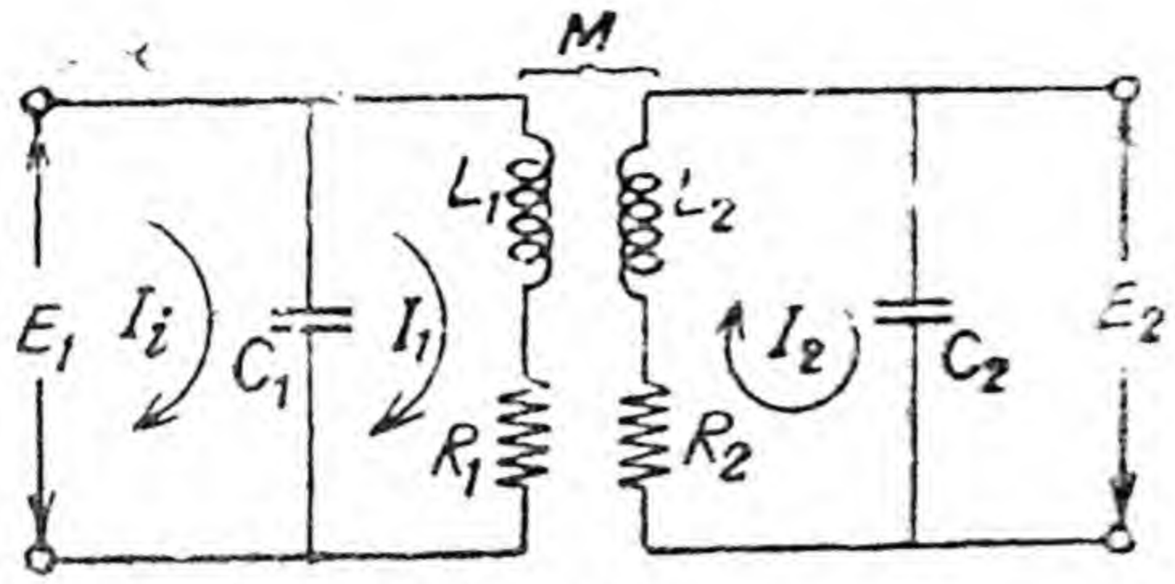


Fig. 20.

Let the centre frequency be ω_0 then at operation frequency ω, the fractional frequency variation δ can be written as

$$\delta=\frac{\omega-\omega_0}{\omega_0}$$

or $$\omega=\omega_0\,(1+\delta). \qquad \ldots(3)$$

Therefore, we can write

$$R_1+j\left(\omega L_1-\frac{1}{\omega C_1}\right)=R_1\left[1+\frac{j\omega_0 L_1}{R_2}(1+\delta)-\frac{j}{\omega_0 R_1 C_1\,(1+\delta)}\right]$$

$$=R_1\left[1+jQ_1\,(1+\delta)-\frac{jQ_1}{1+\delta}\right]$$

$$=R_1\,[1+jQ_1\,(1+\delta)-jQ_1\,(1+\delta)^{-1}]$$

$$=R_1\,[1+jQ_1.2\delta], \qquad \ldots(4)$$

where $$Q_1=\frac{\omega_0 L_1}{R_1}=\frac{1}{\omega_0 C_1 R_1} \qquad \ldots(5)$$

Similarly,

$$R_2+j\left(\omega L_2-\frac{1}{\omega C_2}\right)=R_2\,(1+j.Q_2 2\delta)$$

where $$Q_2=\frac{\omega_0 L_2}{R_2}=\frac{1}{\omega_0 C_2 R}$$

Hence equations (1) and (2) become

$$R_1\,(1+j.2\delta Q_1)-j\omega_0\,MI_2=-j\,\frac{I_1}{\omega_0 C_1} \qquad \ldots(6)$$

and $$R_2\,(1+j.2\delta\,Q_2)-j\omega_0 MI_1=0 \qquad \ldots(7)$$

From equations & (6) and (7), the expression for I_1 can be written as

$$I_1=\frac{-j\,\frac{1}{\omega_0 L_1}.\,R_2(1+j\cdot 2\delta.Q_2)}{R_1R_2\,(1+j.2\delta\cdot Q_1)(1+j\cdot 2\delta\cdot Q_2)+(\omega_0 M)^2}\,I_i$$

Dividing numerator and denominator by R_1R_2, we get

$$I_1=\frac{-j\,\dfrac{1}{\omega_0 L_1 R_1}(1+j\cdot 2\delta.Q_2)}{(1+j.2\delta.Q_1)(1+j.2\delta.Q_2)+\dfrac{\omega_0^2M^2}{R_1R_2}}\,I_i$$

$$=\frac{-j\,Q_1\,(1+i.2\delta.Q_2)}{(1+j.2\delta.Q_1)(1+j.2\delta.Q_2)+K^2Q_1Q_2}\,I_i \qquad ...(8)$$

where $\dfrac{\omega_0^2M^2}{R_1R_2}=K^2Q_1Q_2$; K being the coupling coefficient.

Now we also write from figure that

$$E_1=\frac{1}{j\omega_0C_1}(I_i-I_1)$$

But from equation (8), we can infer that $I_1 \gg I_i$ so that

$$E_1=-\frac{I_1}{j\omega_0C_1}=-\frac{I_1R_1}{j\omega_0R_1C_1}=-\frac{I_1R_1Q_1}{j}$$

$$=\frac{Q_1^2\,(1+j.2\delta.Q_2)}{(1+j.2\delta.Q_1)(1+j.2\delta.Q_2)+K^2Q_1Q_2}R_1I_i \qquad ...(9)$$

Assuming the mutual inductance M to be small, the impedance coupled from secondary into the primary circuit may be ignored so that $I_1\approx\dfrac{E_1}{R_1+j\omega L_1}$.

If Q of the primary is high, then $\omega L_1 \gg R_1$, due to which

$$I_1\approx\frac{E_1}{j\omega L_1}.$$

The e.m.f. induced in the secondary is given by

$$E_{ind}=\pm j\omega MI_1=\pm\frac{M}{L_1}E_1$$

If we neglect the loading effects of the diode rectifiers, then

$$E_2=\frac{-\dfrac{j}{\omega C_2}E_{ind}}{R_2+j\left(\omega L_2-\dfrac{1}{\omega C_2}\right)}=\pm\frac{\dfrac{j}{\omega C_2}\cdot\dfrac{M}{L_1}E_1}{R_2(1+j.2\delta.Q_2)}$$

or

$$\frac{E_2}{E_1}=\frac{\pm\dfrac{j}{\omega C_2}.K\sqrt{\left(\dfrac{L_2}{L_1}\right)}}{R_2(1+j.2\delta.Q_2)}$$

because $\dfrac{M^2}{L_2L_1}=K^2$ giving $\dfrac{M}{L_1}=K\sqrt{(L_2/L_1)}$. Choosing the negative sign (because E_1 and E_2 are 180° in phase opposition) and taking $Q_2=\dfrac{1}{\omega C_2R_2}$, we get

$$\frac{E_2}{E_1}=-\frac{jQ_2.K\sqrt{(L_2/L_1)}}{(1+j.2\delta.Q_2)}$$

Putting the value of E_1 from equation (9), we get

$$E=j\sqrt{\left(\frac{L_2}{L_1}\right)}.\frac{K.Q_2Q_1^2}{(1+j\,2\delta.Q_1)(1+j.2\delta.Q_2)+K^2Q_1Q_2}\,R_1I_1 \qquad \ldots(10)$$

Ordinarily $Q_1=Q_2=Q$ say Under this condition equations (9) and (10) reduce to the following forms

$$E_1=\frac{Q_2^2\,(1+j.2\delta.Q)}{(1+j.2\delta\,Q)^2+K^2Q^2}\,R_1I_1 \qquad \ldots(11)$$

$$E_2=j\sqrt{\left(\frac{L_2}{L_1}\right)}.\frac{KQ^3}{(1+j\,2\delta.Q)^2+K^2Q^2}R_1I_1.$$

Hence $$\frac{E_2}{E_1}=j\sqrt{\left(\frac{L_2}{L_1}\right)}.\frac{KQ}{1+j.2\delta.Q}. \qquad \ldots(12)$$

The discriminator is assumed to be driven by a pentode which is a constant current device (C_1L_1 tank circuit is a load for pentode and in a pentode after suitable plate voltage, plate current becomes constant and hence I_1 is constan.). Thus I_1 being assumed as constant, we can easily infer that E_1 varies with δ, *i.e.* fractional deviation of operating frequency from centre frequency. As E_2 is a function of E_1, output would thus surely follow the frequency variation of input voltage

Further from discussion of discriminator function, we find that voltages applied to the plate of diodes A and B are respectively :

$$E_A=\frac{E_2}{2}+E_1,$$

and $$E_B=-\frac{E_2}{2}+E_1$$

Let output voltage developed across R_A and R_B be E°_A and E°_B If detection efficiency be η, then

$$\eta=\frac{E^\circ_A}{E_A}=\frac{E^\circ_B}{E_B}$$

so that $$E^\circ_A=\eta E_A$$

and $$E^\circ_B=\eta E_B$$

or $$E^\circ_A=\eta E_A=\eta\left|E_1+\frac{E_2}{2}\right|=\eta E_1\left|1+\frac{1}{2}.\frac{E_2}{E_1}\right|$$

$$=\eta E_1\left|1+j.\tfrac{1}{2}.\sqrt{\left(\frac{L_2}{L_1}\right)}.\frac{KQ}{1+j.2\delta.Q.}\right|$$

and $$E^\circ_B=\eta\ E_B=\eta\left|E_1-\frac{E_2}{2}\right|=\eta E_1\left|1-\tfrac{1}{2}.\frac{E_2}{E_1}\right|$$

$$=\eta E_1 \left| 1-j.\tfrac{1}{2}.\sqrt{\left(\frac{L_2}{L_1}\right)}.\frac{KQ}{1+j.2\delta.Q} \right|$$

Therefore output of the discriminator will be

$$E^\circ = E^\circ_A - E^\circ_B$$

$$=\eta E_1 \left[\left| 1+j.\tfrac{1}{2}.\sqrt{\left(\frac{L_2}{L_1}\right)}.\frac{KQ}{1+j.2\delta.Q} \right| - \left| 1-j.\tfrac{1}{2}.\sqrt{\left(\frac{L_2}{L_1}\right)}.\frac{KQ}{1+j.2\delta.Q} \right| \right]$$

Putting the value of E_1 from equation (11), we get

$$E^0=\eta R_1 I_i \left|\frac{Q^2(1+j.2\delta.Q)}{(1+j.2\delta.Q)^2+K^2Q^2}\right| \cdot \left[\left| 1+j.\tfrac{1}{2}.\sqrt{\left(\frac{L_2}{L_1}\right)}\frac{KQ}{1+j.2\delta.Q} \right| - \left| 1-j.\tfrac{1}{2}.\sqrt{\left(\frac{L_2}{L_1}\right)}.\frac{KQ}{1+j.2\delta.Q} \right| \right] \quad \ldots(13)$$

Which is the desired expression for the discriminator output.

A plot in E^0 and instantaneous impressed frequency is shown in figure 21 and is called discriminator characteristic. Point to be

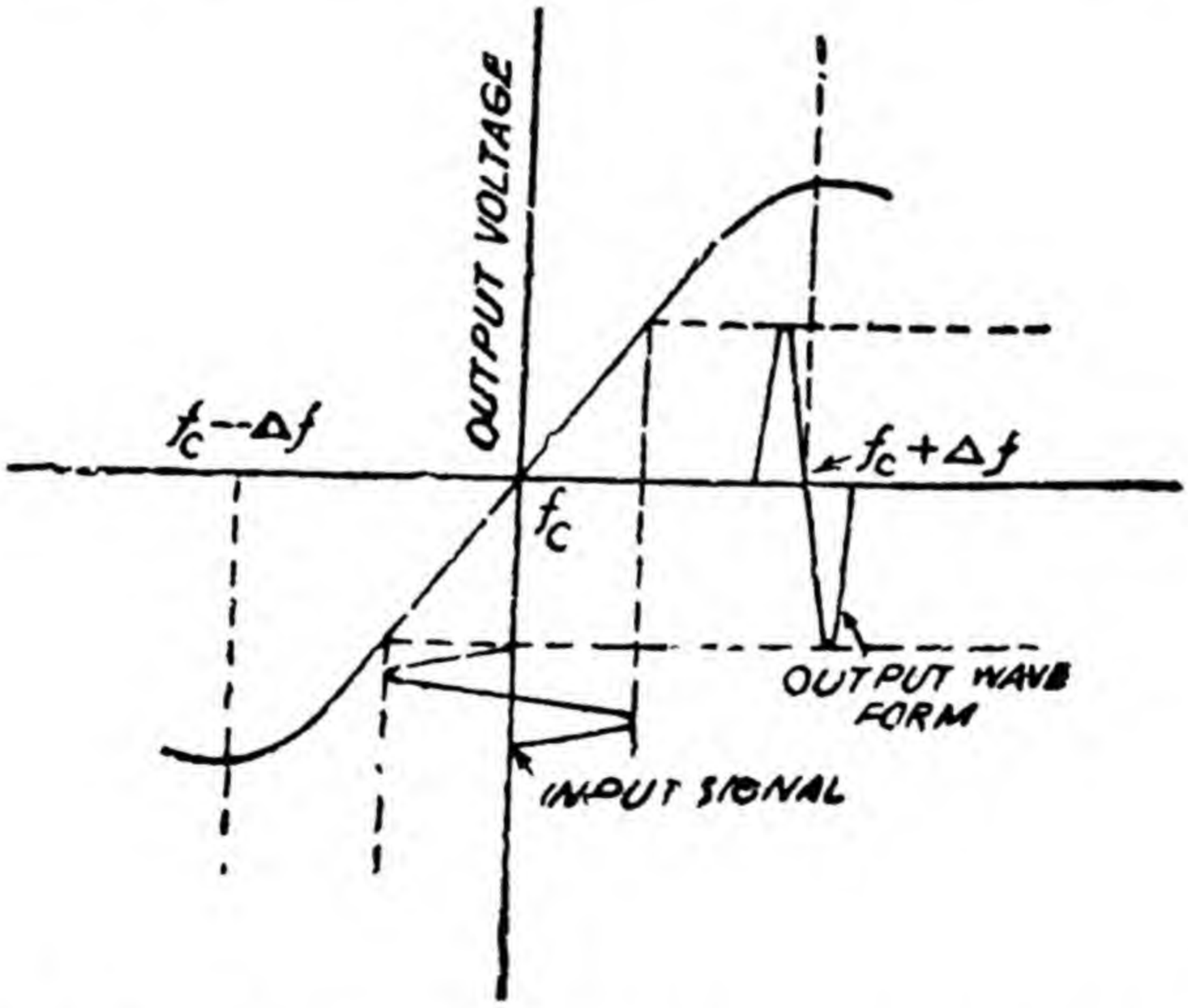


Fig 21. Relation between impressed frequency and output voltage.

noted is that capacitors C_A and C_B are large enough so as to ground carrier frequency voltage. They offer appreciable reactance at modulation frequency and therefore, voltage developed across C_A and C_B, *i.e.* E_A^0 and E_B^0 are audio frequency voltages.

Discriminator characteristic : A relation between impressed signal frequency and output voltage in discriminator circuit is shown in fig. 21.

We see from the curve that for not more than Δf frequcncy deviation, detection is linear. f_c denotes the centre frequency of modulated wave.

Disadvantages : We see from vector relations that the output of discriminator circuit is a function of input voltage *amplitude* and frequency variations. When the applied signal consists of a frequency modulated wave that also varies in amplitude, the detector output will contain undesired components corresponding to the amplitude variations as well as the output representing the frequency modulation. Thus interference is produced. To avoid this difficulty, limiter is used previous to discriminator stage to deliver it a constant amplitude input.

31·8. LIMITER :

The action of the limiter is to ensure that all variations in the carrier amplitude of modulated wave due to noise and interference are removed and a constant carrier amplitude is applied to the discriminator input.

Circuit for limiter is shown in fig. 22.

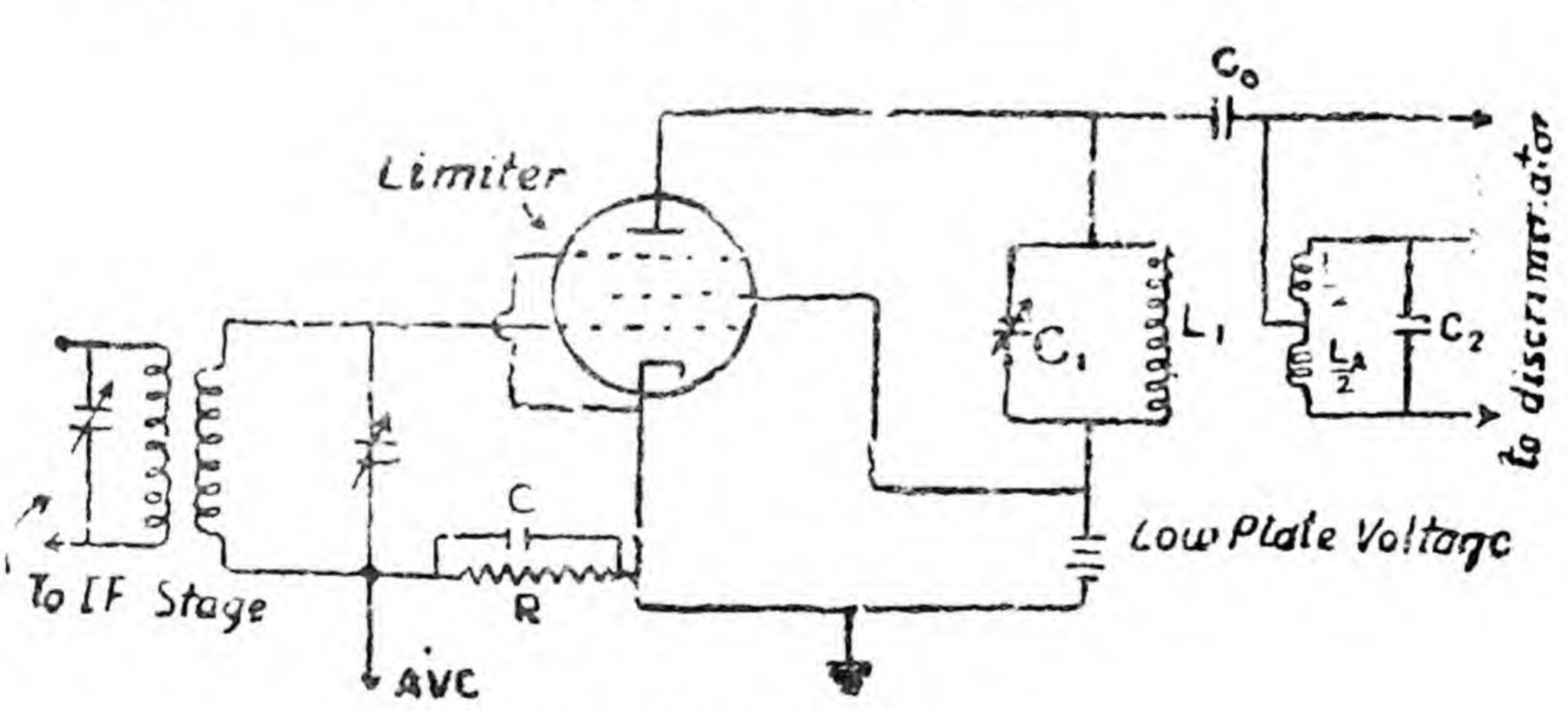


Fig. 22. Limiter circuit.

Operation : A sharp cut-off tube is used. Plate and screen potentials are kept low such that the tube saturates easily and cut off is produced Signal is limited on both positive and negative peaks. At positive swing effects of current flow in the grid circuit limit the amplitude while at negative swing, amplitude is limited due to plate current cut-off as shown in fig. 23.

For limiter action it is necessary that input signal must be sufficient to produce the desirec over-loading effect. It is, therefore, important that in frequency modulated receiver the preceding stages *i.e.* r.f. amplifier and audio-frequency amplifier, must produce sufficient output voltage to make the limiter operative.

It can be seen from the characteristic that if input signal amplitude increases beyond cut-off, it merely increases the flat portion in the output wave. Thus by limiter action a constant carrier amplitude of modulated wave is delivered to discriminator.

A combination of limiter and discriminator device, thus, is capable of satisfactory detection of frequency modulated wave. In next article, we shall describe another device self sufficient for detection process, without the need of limiter. The device is simply a modified form of discriminator circuit named as ratio detector.

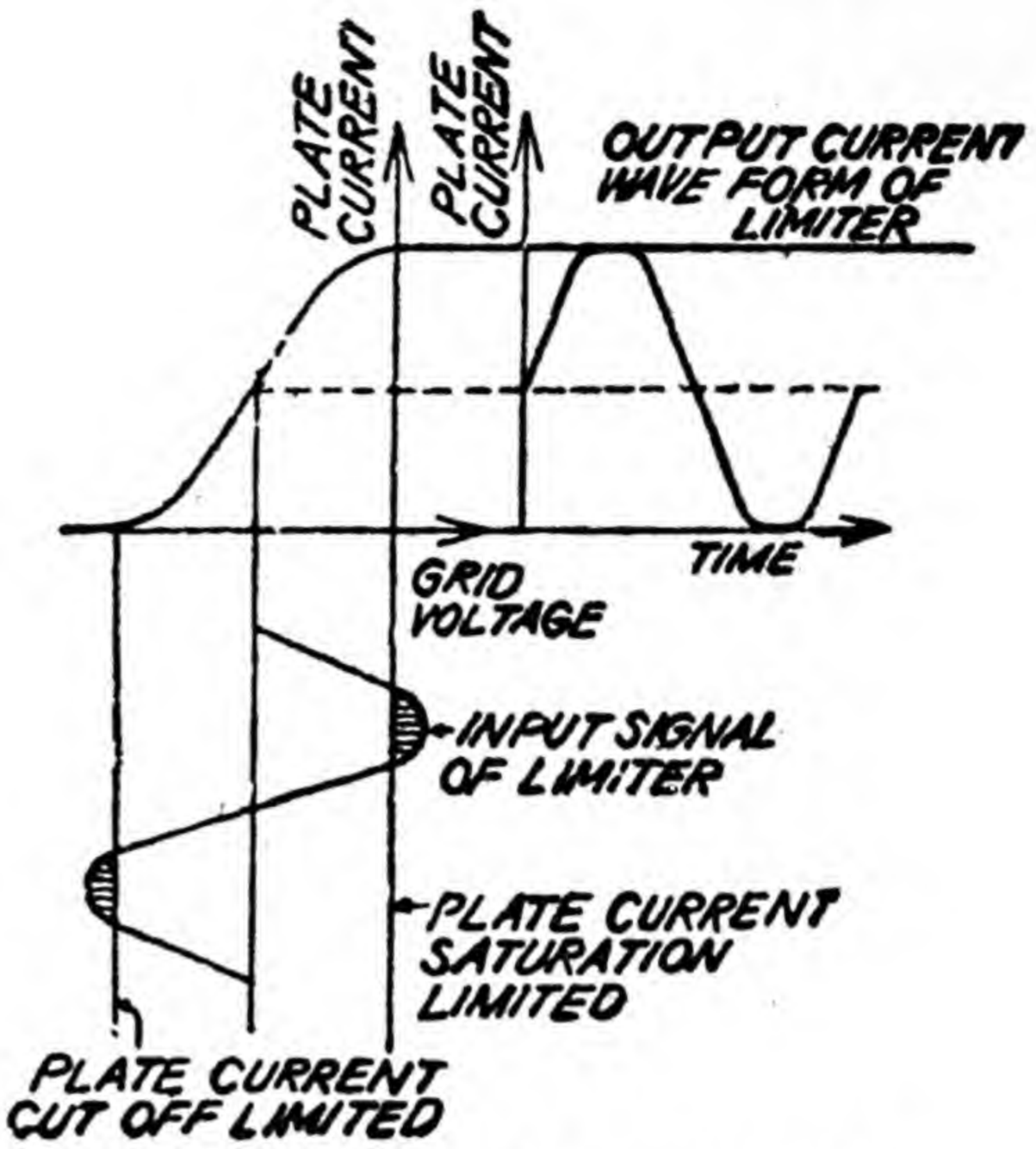


Fig. 23. Operational Characteristics of Limiter.

31·9. RATIO DETECTOR :

Ratio detector serves as an f.m. detector that responds only to the input frequency variations because its output depends upon the ratio of the voltage across the diodes rather than difference between them and hence the name.

A phase shift discriminator detector responds to the amplitude variation of the input signal which is not desirable. As a result, a limiter circuit is always associated with discriminator detector as its driving stage. The ratio detector is a modification of the discriminator which can be designed as to be unresponsive to amplitude modulation while behaving in the same way as the discriminator.

In discriminator, output is $E^\circ = E_A^\circ - E_B^\circ$. Suppose at centre frequency, signal input to the discriminator develops $E_A^0 = E_B^\circ = 4$ volts, say. When frequency modulated wave is input to the discriminator, E_A° will not be equal to E_B^0 ; instead, if E_A° increases by certain amount then E_B° will decrease by the same amount. Say $E_A^\circ = 6$ volts then $E_B^\circ = 2$ volts. The output will be $E^\circ = 6 - 2 = 4$ volts.

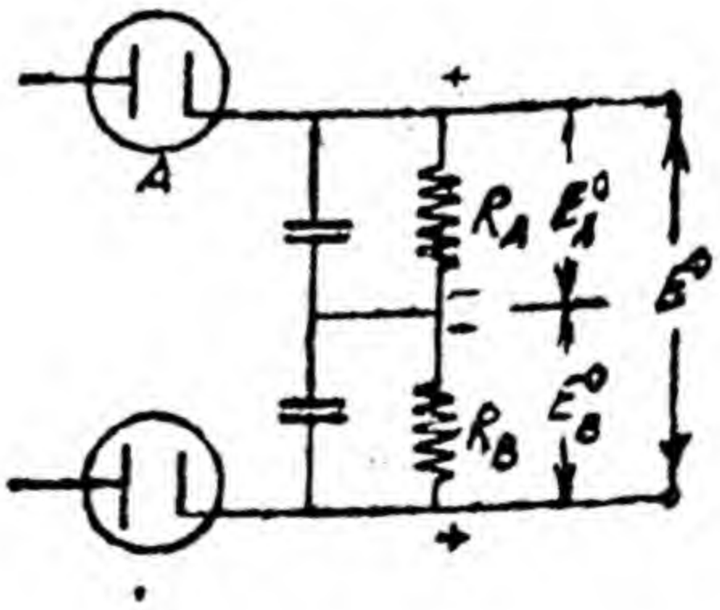


Fig. 24. (a)

Now suppose we increase strength of carrier (r.f) (which will happen if the amplitude modulation is also present) until we have $E_A^\circ = E_B^\circ = 8$ volts at the centre frequency. With the same frequency shifts as above, but with this stronger carrier, the voltage

$E_A°$ would rise to 12 volts and $E_B°$ decreases to 4 volts. The output, in this case, will be

$$E°=12-4$$
$$=8 \text{ volts}$$

It means, with same frequency shift, output of the discriminator in the two cases is different. It implies that this circuit responds to amplitude variations. We have observed in previous case that $E_A°+E_B°=8$ volts, in second case $E_A°+E_B°=16$ volts, which is possible in this circuit because there is no provision to keep $E_A°+E_B°$ as constant—the only condition is that $E°$ should increase by the same amount by which $E_B°$ decreases.

However we note that in both the cases the *ratio* of the voltages $E_A°$ and $E_B°$, that is

$$\frac{E_A°}{E_B°}=\frac{6}{2}=\frac{12}{4}$$

or $$E_A° : E_B°=3 : 1$$

is the same. So if a detector could be made to respond only to this ratio and not to the respective magnitudes, then it would be insensitive to amplitude variations. The ratio detector is such a circuit which responds only to this ratio and consequently, is unresponsive to carrier changes.

Preliminary requirement of the ratio Detector : We have mentioned above in the two cases that $E_A°+E_B°$ is different. Had it been possible to keep $E_A°+E_B°$ *constant* (=8 volts) in both the cases, effect of amplitude modulation would have been eliminated leaving only the effect of frequency shift (frequency modulated wave) at the output. Therefore, in ratio detector circuit, one of the diodes is reversed in polarity so that total voltage across R_A and R_B will be $(E_A°+E_B°)$ instead of $(E_A°-E_B°)$ and then this voltage is kept constant with the help of large capacitor C $(=8\mu f)$ and a resistor R combination. The function of such a CR circuit will be detailed later on. First we shall explain the function by taking a battery E_{bb} instead of CR combination.

Preliminary Form of Ratio Detector : A preliminary circuit is shown in fig. 24 (*b*). In this circuit polarity of diode T_A is reversed. Therefore, total voltage across R_A and R_B will be $(E_A°+E_B°)$.

Since C_A and C_B are in series directly across E_{bb}, the sum of their voltages must equal E_{bb}. That is $E_A°+E_B°=E_{bb}$. Also, due to the manner in which battery is connected to T_A and T_B, no current can flow around the circuit until a signal is applied. Now while $E_A°+E_B°$ can never exceed E_{bb}, $E_A°$ does not have to equal $E_B°$. In other words, ratio of $E_A°$ to $E_B°$ may vary. The output is obtained across R_B and ground.

When incoming signal is at centre frequency, $E_A°=E_B°$. When incoming signal has frequency greater or lower than centre frequency, then due to frequency shift, discriminator action starts and

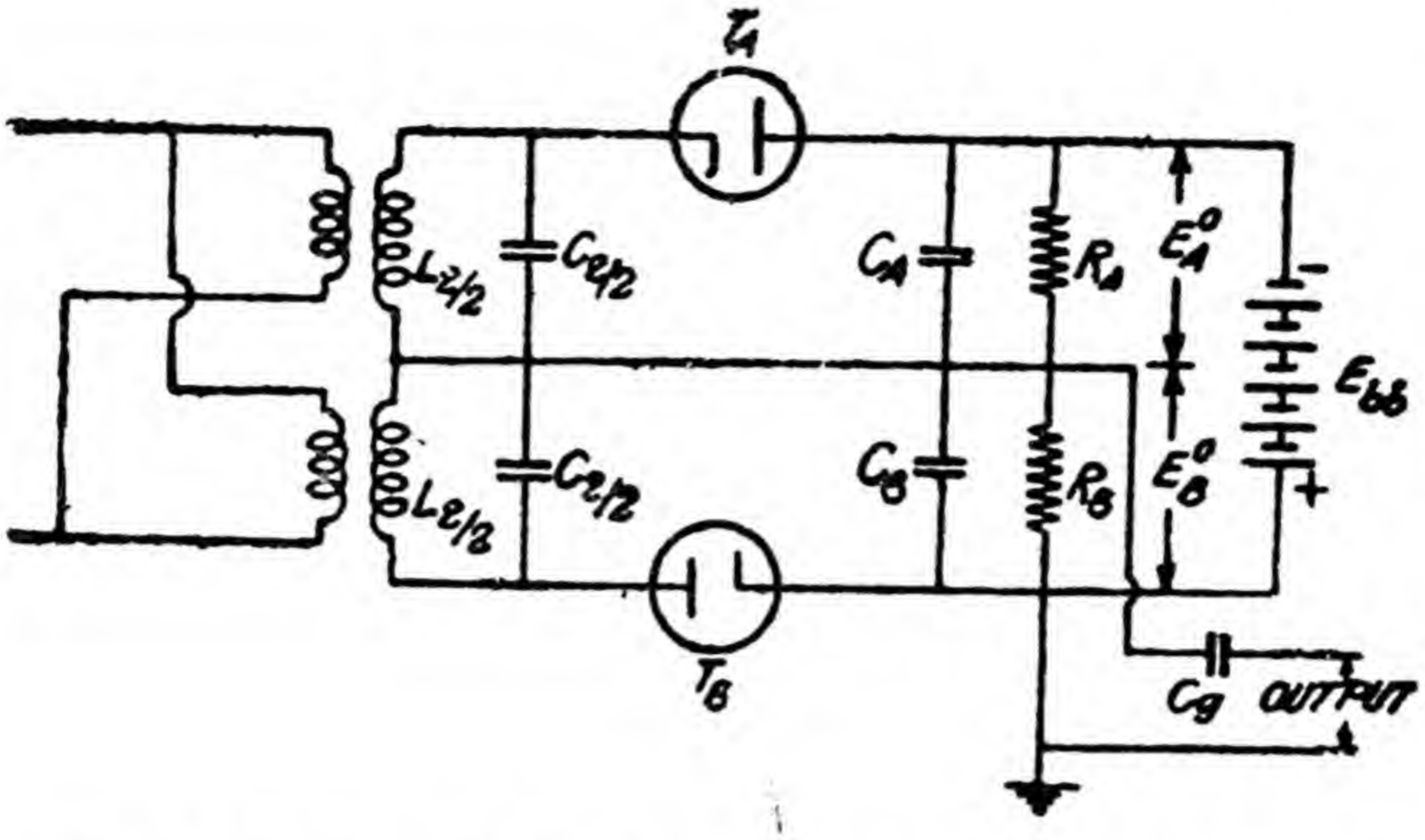


Fig 24 (b) Preliminary form of ratio detector.

$E_A^\circ \neq E_B^\circ$, but $E_A^\circ + E_B^\circ$ is still equal to E_{bb}. Carrier frequency voltage is grounded (due to the large values of C_A and C_B which offer reactance at audio frequency) and audio variations are obtained across R_B. Capacitor C_g prevents the rectified d.c. voltage in the detector circuit from reaching the grid of audio amplifier.

The purpose of E_{bb} in this preliminary circuit is to maintain audio output voltage which is purely a result of frequency modulated signal. E_{bb} keeps the total voltage $(E_A^\circ + E_B^\circ)$ constant, while it permits the ratio of E_A° to E_B° to vary This condition as we have already stated, eliminates the effect of all amplitude variations.

But selection of the value of E_{bb} is difficult. If E_{bb} is chosen high, weak signals will be lost ; on the other hand, if E_{bb} is chosen quite low, only small audio outpu voltages are obtainable. This difficulty is overcome on replacing E_{bb} by a high time constant circuit, RC, in practical ratio detector circuits.

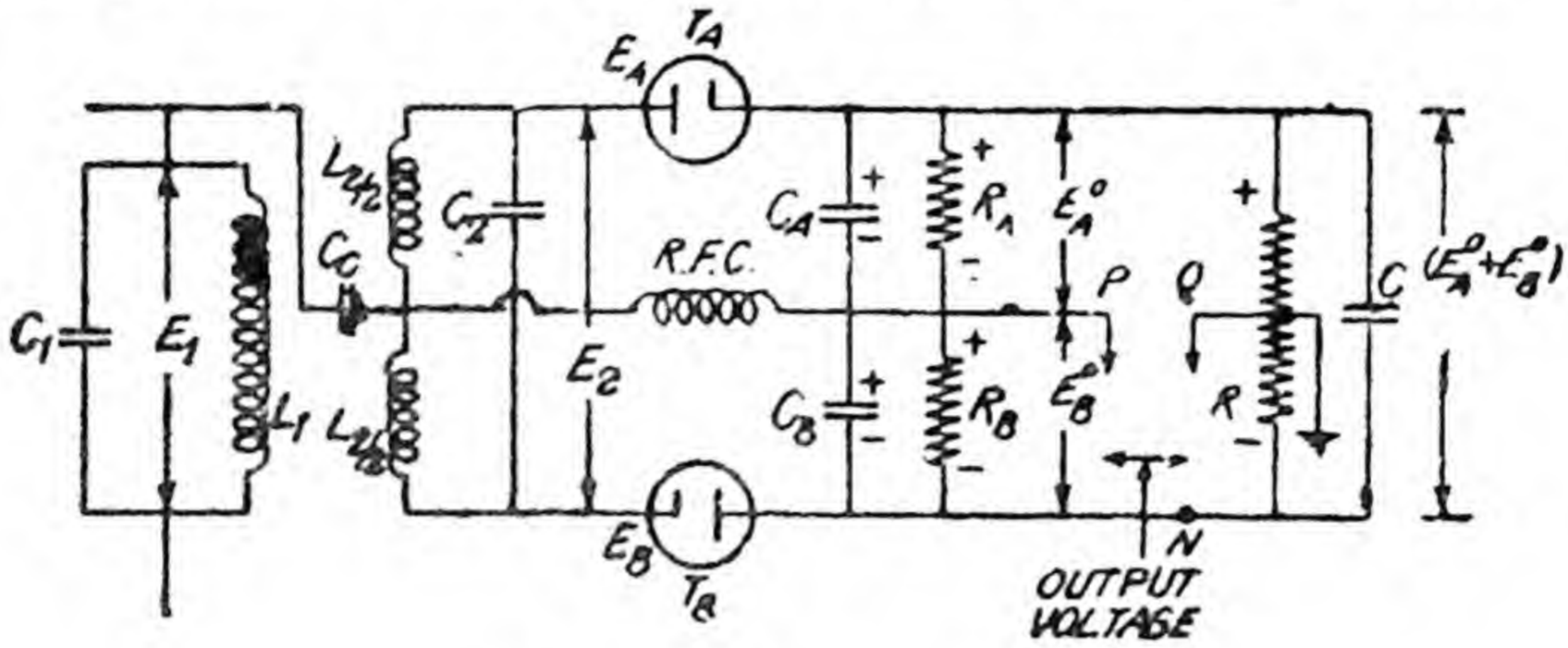
Practical Circuit : The circuit of ratio detector is shown in fig. 24 (*c*). There are two major changes as compared to discriminator circuit.

(*i*) diode T_B has been reversed in polarity so that output across R_A and R_B will now $E^\circ = E_A^\circ + E_B^\circ$.

(*ii*) the output voltage is obtained between points P and Q. Q is the centre tap on the high resistance R that shunts the load impedance of the two diodes. The time constant, CR, is large enough—of the order of 0·2 sec.

Operation : Since diode T_B has been reversed in polarity, current can now flow in the overall circuit—from bottom of L_2 through T_B, through long time constant circuit, RC, through T_A and back to the top of L_2. After several cycles of radio frequency, C charges approximately to peak value of voltage across L_2. Amplitude variations due to noise or other interference cause very little effect on the charge of capacitor C and voltage across

RC circuit remains constant.* This means voltage across C_A and C_B or R_A and R_B will remain constant *i.e.* E_A^0 and E_B^0 will be constant. Thus *RC* circuit serves the purpose of E_{bb}. The changes in carrier amplitude are merely absorbed by the capacitor.



$$E_A = E_1 + E_2/2$$
$$E_B = E_1 - E_2/2$$

Fig. 24. (*c*) Circuit of ratio detector.

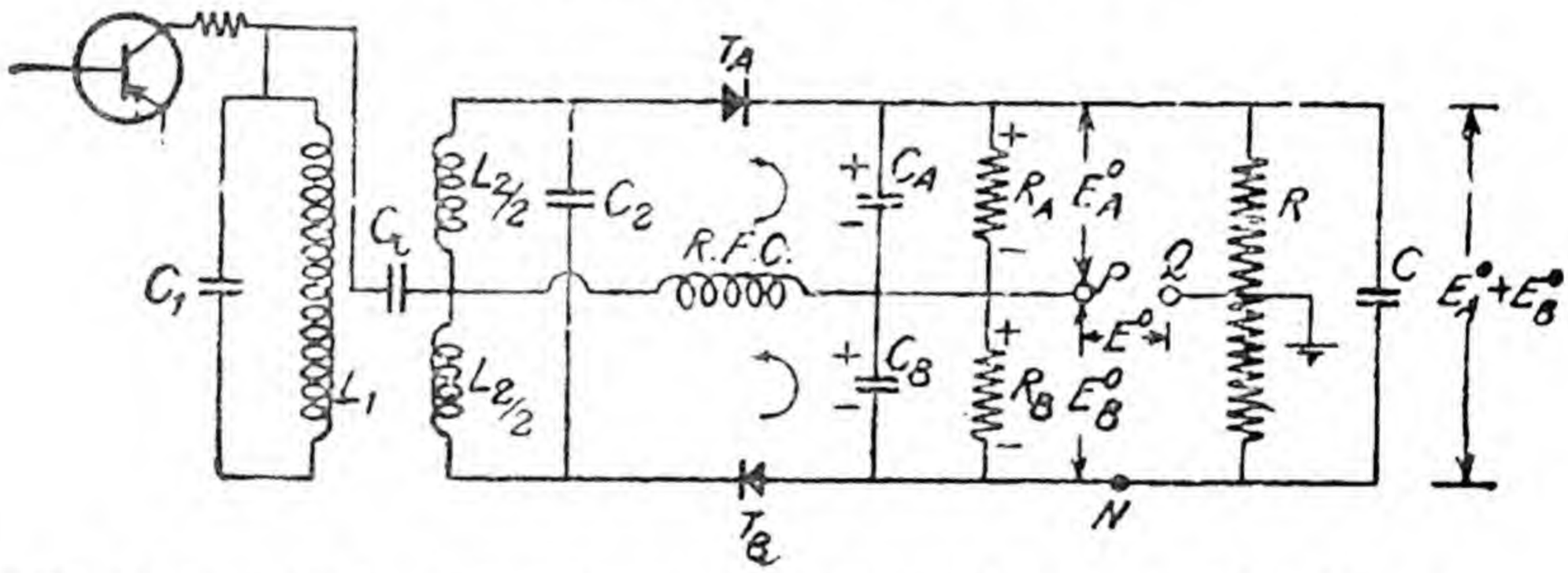


Fig. 24 (*d*) Ratio detector circuit quite equivalent to that shown in fig. 24 (*c*) except the use of semiconductor diodes.

At centre frequency, say $E_A^0 = E_B^0 = 6$ volts. At frequency modulated signal, say $E_B^0 = 10$ volts, then $E_A^\circ = 2$ volts (since $E_A^0 + E_B^0 = 12$ is constant).

The voltage of point *P*, *compared to point N*, is then $+10V$, and voltage of point *Q*, compared to point *N*, is $+6$ volts. Thus output will be

$$(+10)-(+6)=+4 \text{ volts.}$$

Output voltage : Output voltage in this case will be

$$E^0 = E_B^0 - \frac{E_A^0 + E_B^\circ}{2}$$

which is exactly half the magnitude of the output obtained in the discriminator circuit.

How the circuit is unresponsive to amplitude modulation : The

*If receiver is tuned to some other station of higher (or lower) signal level then after a number of r.f cycles the time constant circuit *RC* will gradually charge (or discharge) to the new voltage level. Note that due to tuning of new station, r.f. frequency will change.

additional RC circuit employed at the output terminals helps to remove these amplitude variations appearing at the output. The time constant RC is kept sufficiently large (of the order of 0·25 sec) C is large enough to have a reactance at the lowest modulation frequency of importance that is small compared with the resistance R in parallel with R_A and R_B.

Effect of C is to reduce greatly the amplitude fluctuations in the voltage $(E_A^\circ + E_B^\circ = E^\circ)$ appearing across R. Suppose E momentarily increases in amplitude; then it will cause a large charging current to flow through the diode into C. This represents the power absorbed by the diodes from the resonant circuits and so causes the voltages that these circuits apply to the diodes to be reduced in magnitude. Conversely, if the amplitude of the incoming signal attempts momentarily to drop below the average amplitude, then C attempts to prevent the voltage E° from dropping by supplying current that flows from C to R, R_A and R_B. This relieves the diode tubes of the necessity of supplying as much rectified current as before, thereby increasing their input impedance and reducing the loading on the resonant circuits with corresponding increase in the voltages they apply to the diodes. Thus presence of C reduces the amplitude variations that would otherwise occur in the voltage E° and likewise in the voltages E_1 and E_2.

A vector diagram is drawn in the figure 25, which shows the invariance of voltage $(E_A - E_B)$, the difference of voltages input to diodes A and B, with amplitude variations. Vectors E_1 and E_2 refer to the case when no amplitude modulation is present. E_2 is shifted in phase ϕ by secondary L_2C_2 circuit. The dotted vectors illustrate what happens when the amplitude of the incoming wave is momentarily increased, while maintaining the same frequency deviation. The corresponding vectors E_1' and E_2' are now longer than before but the phase shift ϕ' produced by the same frequency deviation is less. This change in phase angle is just enough to compensate for the increased length of the vectors, and results in the difference $(E_A' - E_B')$ being the same for the dotted system as difference $(E_A - E_B)$.

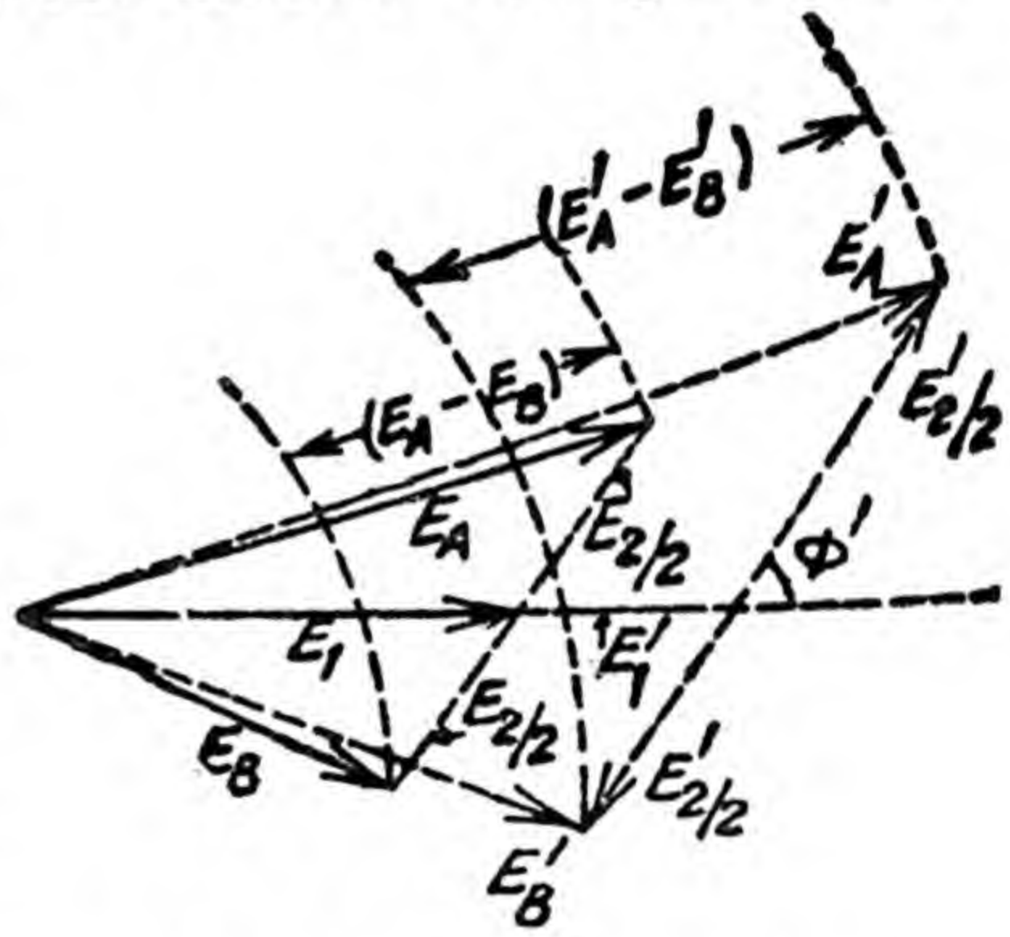


Fig. 25. Vector diagram.

EXPERIMENTS

Expt. No. 1. To draw the detection and dynamic characteristics of diode and to determine its efficiency of detection :

Procedure : Circuit arrangement for obtaining the detection characteristics of a diode is shown in fig. 26. A triode 6C5 is used as a diode by connecting its grid with anode.

The procedure is as follows :

(*i*) Using load resistance *R* of 5 kilohms, vary signal voltage from 1 to 10 volts and note rectified current each time.

(*ii*) Replace 4 kilohms resistor by 50 kilohms resistor and repeat the above process.

(*iii*) Change the capacitor *C* and repeat the whole operation.

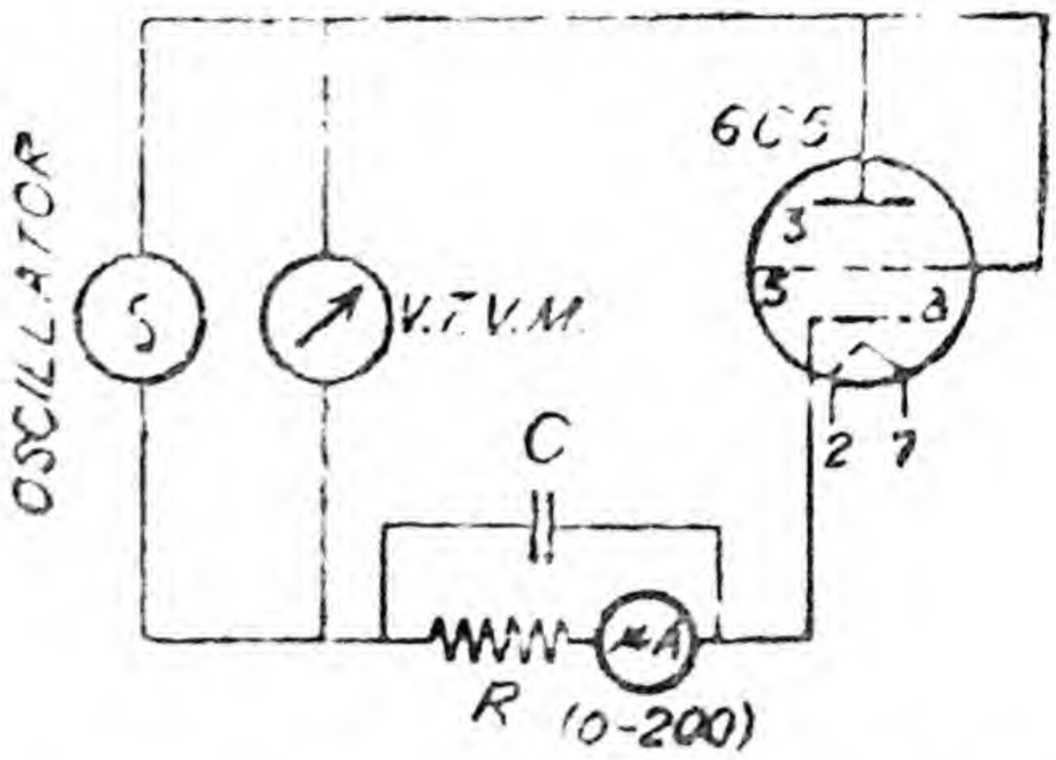


Fig. 26. Circuit arrangement for detection characteristics of diode.

In figure 27, curves of output current to a base of signal voltage for different values of load but with same capacitor value are plotted. Detection characteristics show that as the load is increased the current decreases for fixed signal voltage while the voltage developed across the load increases as it can be seen by calculating the product *RI*. Therefore, for square law detection, we should have low load and high plate resistance, and for linear detection we must have high load and low impedance values.

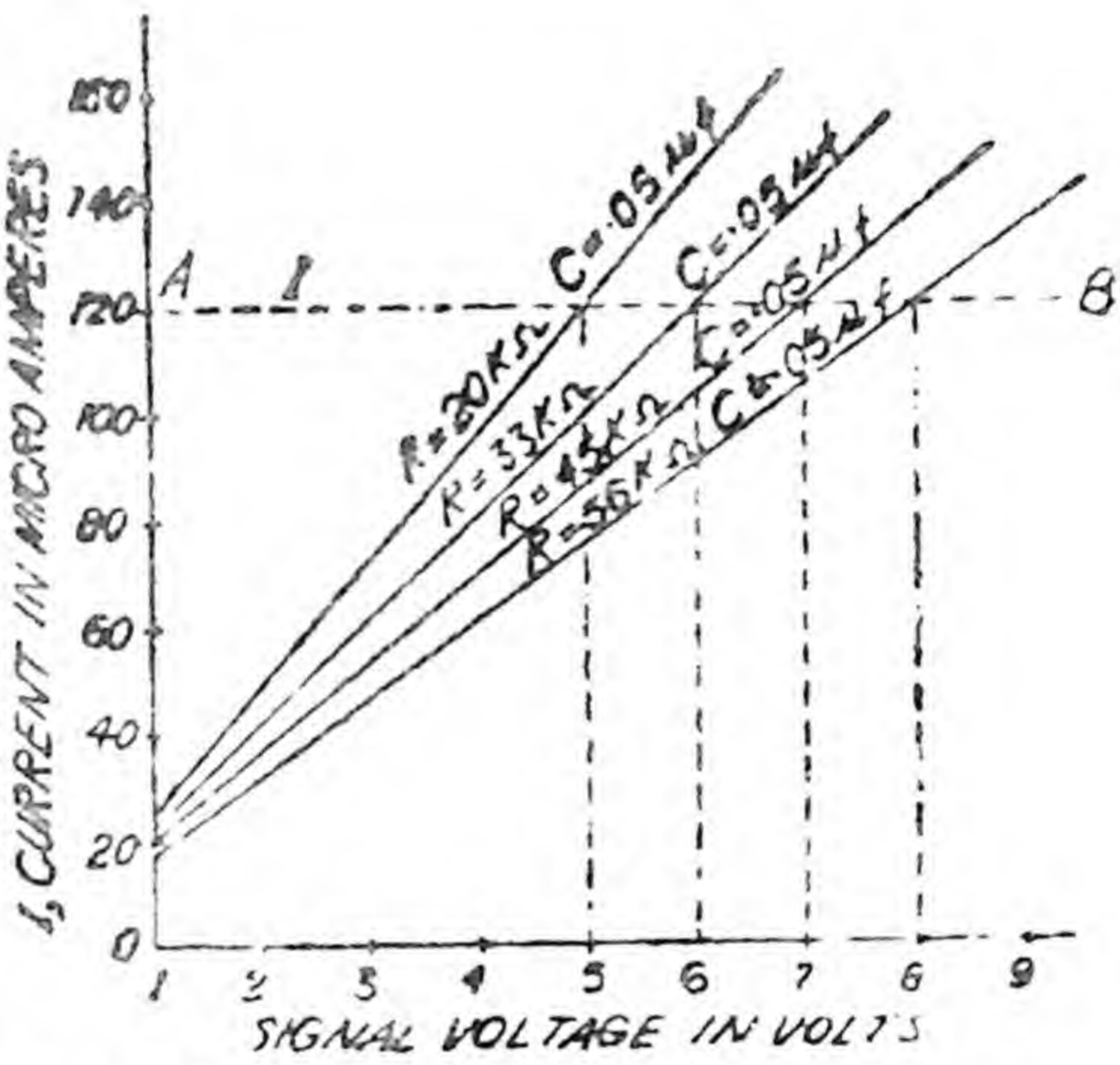


Fig. 27. Detection characteristics of diode.

For the determination of detection efficiency, a line *AB* cutting

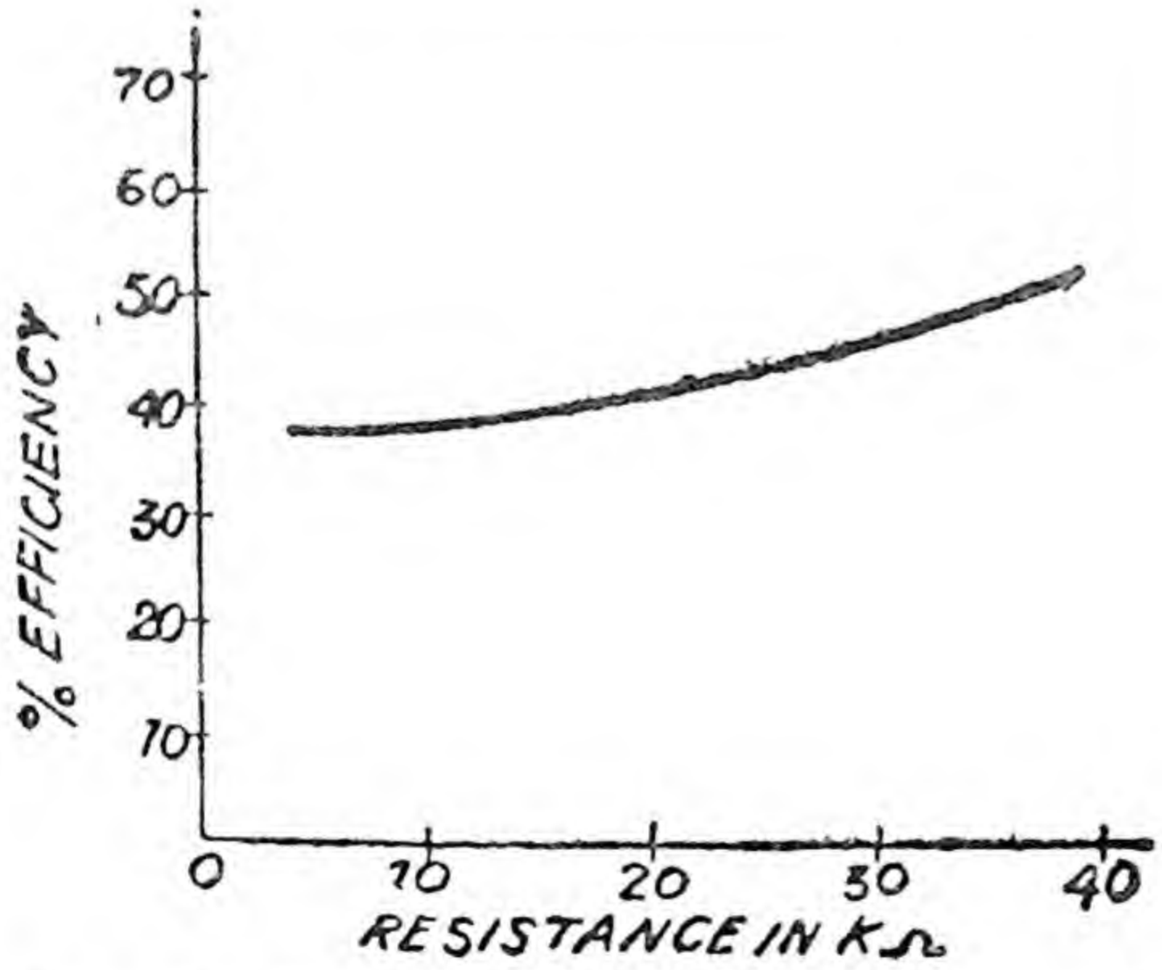


Fig. 28. Variation of percentage efficiency with load.

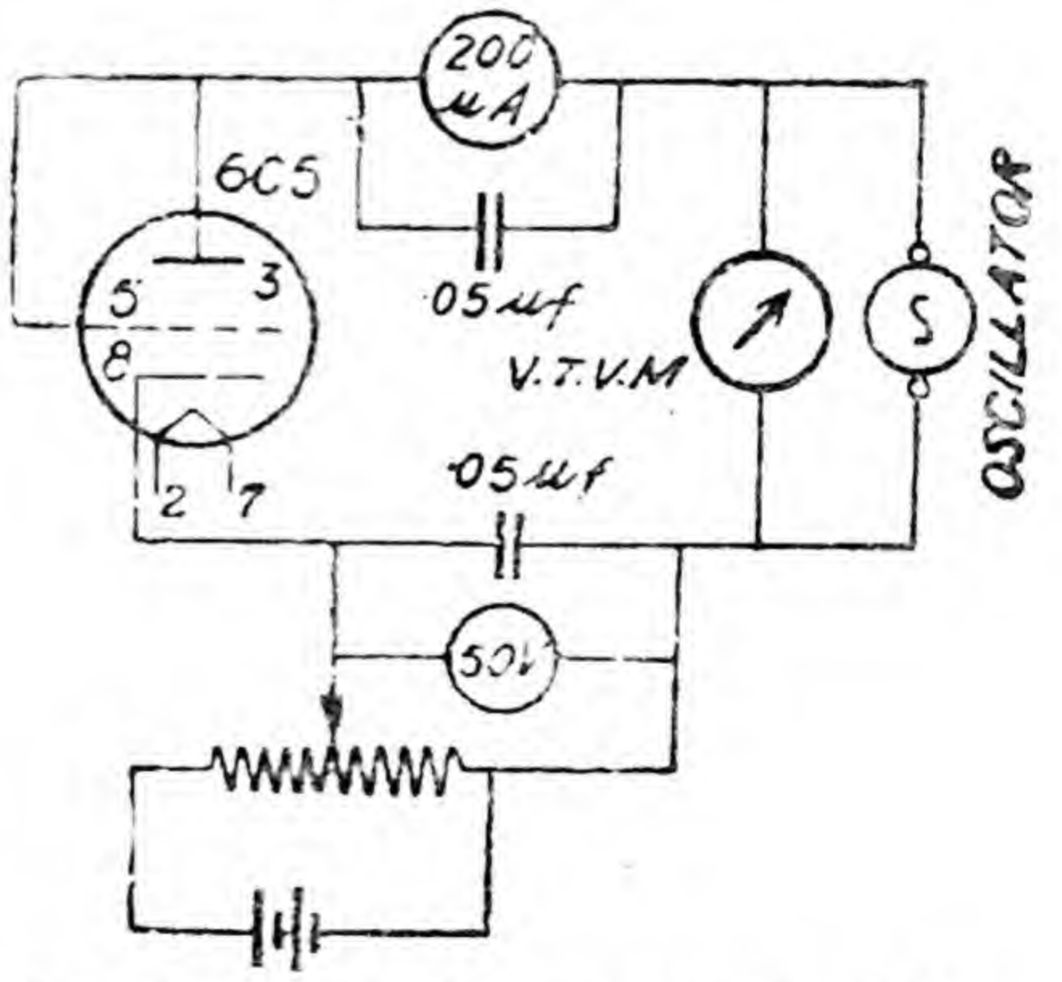


Fig. 29. Circuit for obtaining the dynamic characteristics of diode.

all curves at certain value of plate current I is drawn. Signal voltage E corresponding to different points of intersection can be calculated by the relation :

$$\text{percentage efficiency of detection} = \frac{RI}{E\sqrt{2}} \times 100.$$

A graph of percentage efficiency to a base of load resistance is drawn in figure 28. Curve shows that efficiency of detection slightly increases with load.

Dynamic Characteristics of Diode : Circuit arrangement is shown in fig. 29.

We proceed as follows :

(*i*) The voltage (signal) of the oscillator is set to a known value and rectified current with zero anode bias is measured.

(ii) Increase anode bias in steps and take the corresponding readings of microammeter till current reduces to zero.

(*ii*) Repeat the whole test for other values of signal voltages (oscillator voltage).

Curves of rectified current *versus* anode bias are plotted as shown in figure 30.

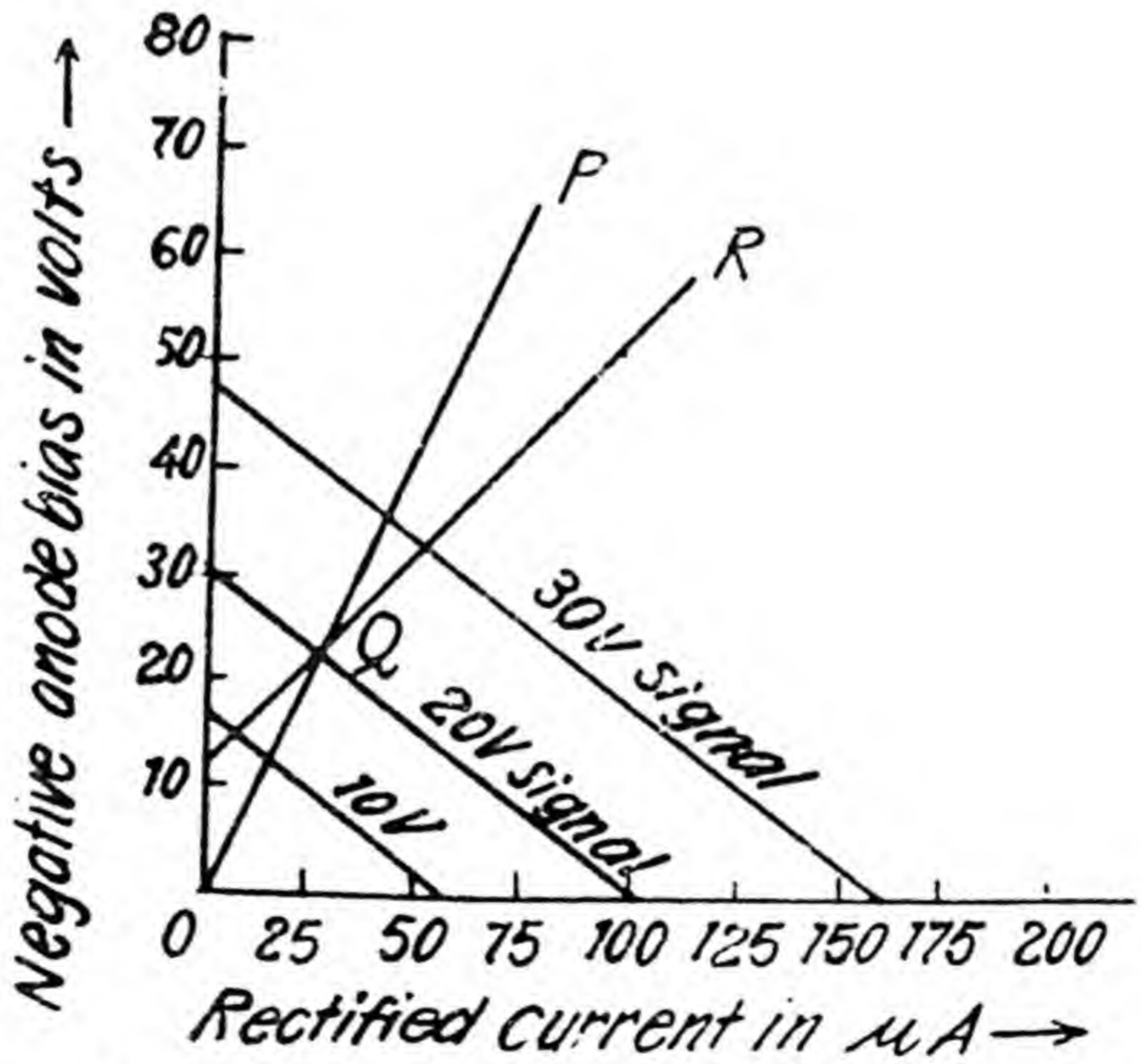


Fig. 30. Dynamic characteristics of a diode detector.

From origin a line OP, called d.c. load line, is drawn having a slope corresponding to 1 megohm. The point of intersection Q of line OP with curve for signal voltage 20 volts gives the operating point. Through Q a line RQ is drawn, the slope of which corresponds to the impedance, at a given audio frequency, of the load resistance in parallel with the audio frequency coupling circuit.

This line represents a.c. load line. The permissible percentage of modulation can be estimated from the points on this line corresponding to zero value of rectified current.

Expt. No. 2. (*A*) To determine the detection characteristic of an anode bend detector.

(*B*) To study the effects of grid condenser and grid leak on sensitivity of a cumulative grid detector.

Procedure : (*A*) Circuit for plotting the detection characteristics is shown in figure 31.

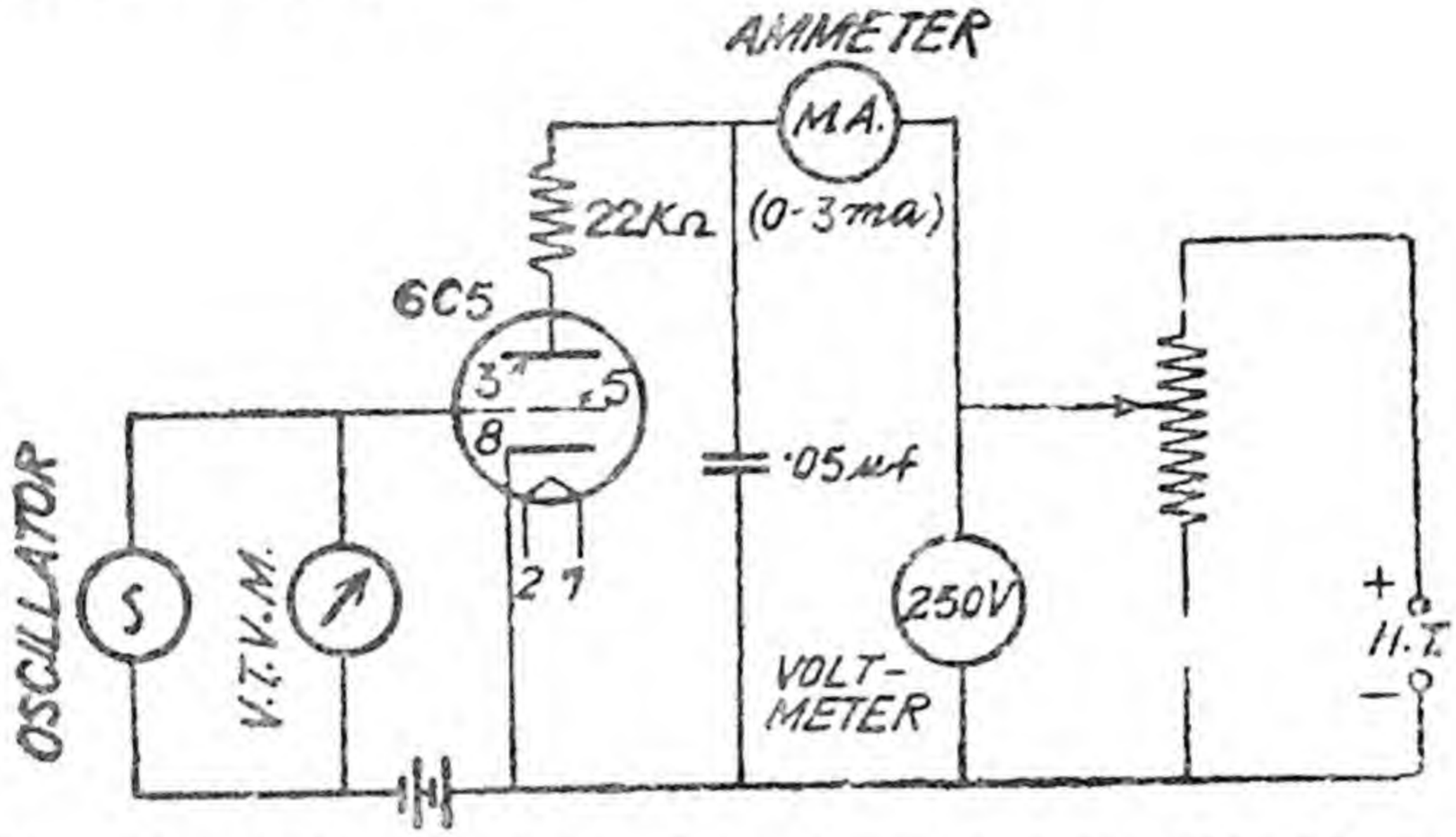


Fig. 31. Circuit for obtaining detection characteristics of an anode bend detector

We proceed as follows :

(*i*) Adjust the grid bias to anode cut off point and apply first low anode voltage.

(*ii*) Apply the signal and take the readings of plate current corresponding to different signal voltages.

(*iii*) The whole test is repeated for one or two higher values of anode vol'ages.

Curves of rectified current to a base of signal voltage for each anode potential are drawn as shown In fig. 32.

From curves we see that for weak signals, the anode bend

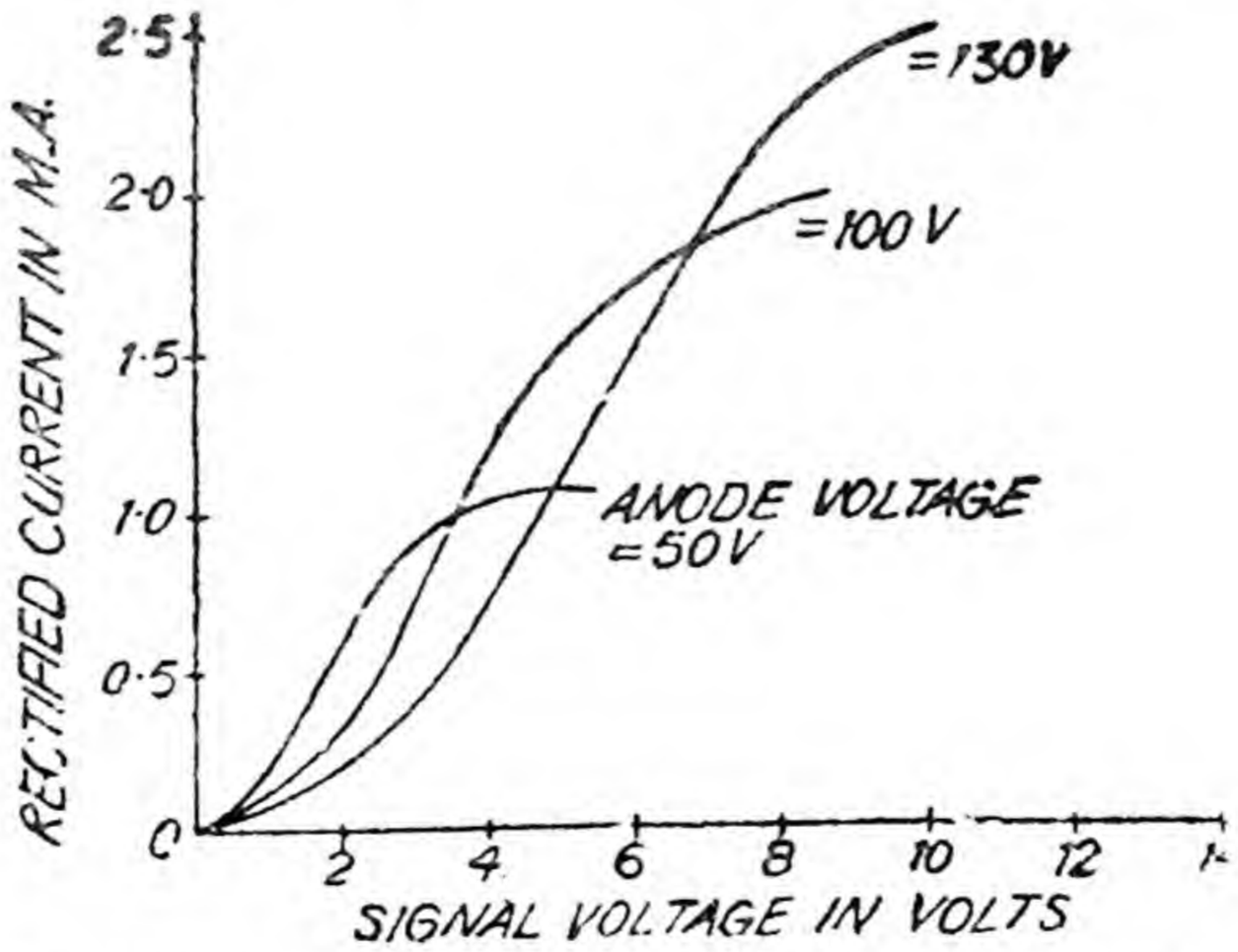


Fig. 32. Anode bend detector characteristics.

characteristic are more sensitive at low anode voltages than at higher values of the latter. The important effect to be noted is that the linear portion of the curve increases in length as the anode voltage is increased. This shows that for distortionless detection of high signal, high anode voltages are necessary.

(B) Procedure : The circuit is shown in figure 33, where *C* and *R* represent the grid capacitor and grid leak respectively. Microammeter in anode circuit is shunted by a bypass condenser and an auxiliary d.c. circuit to enable the steady anode current to be 'backed-off'.

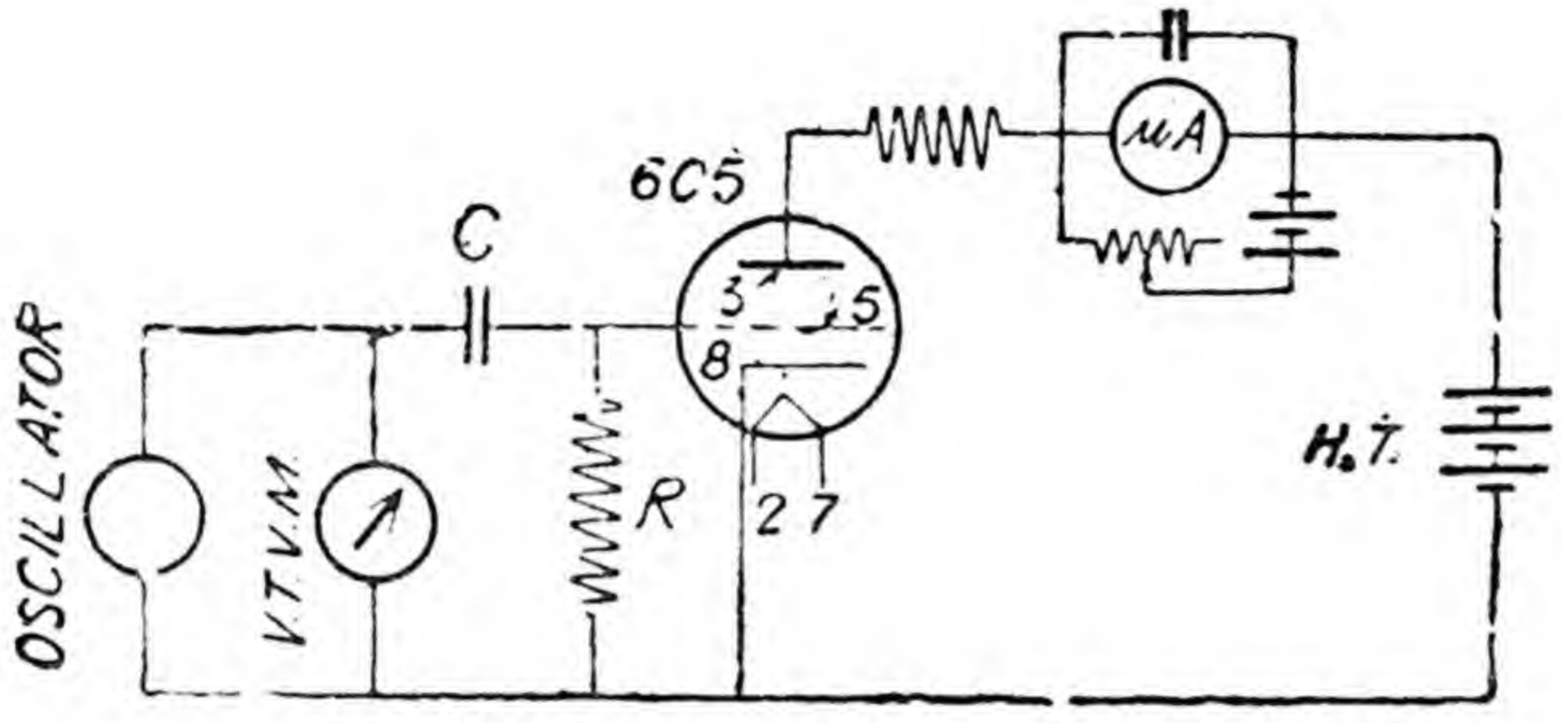


Fig. 33. Cumulative grid detector.

We proceed in the following manner :

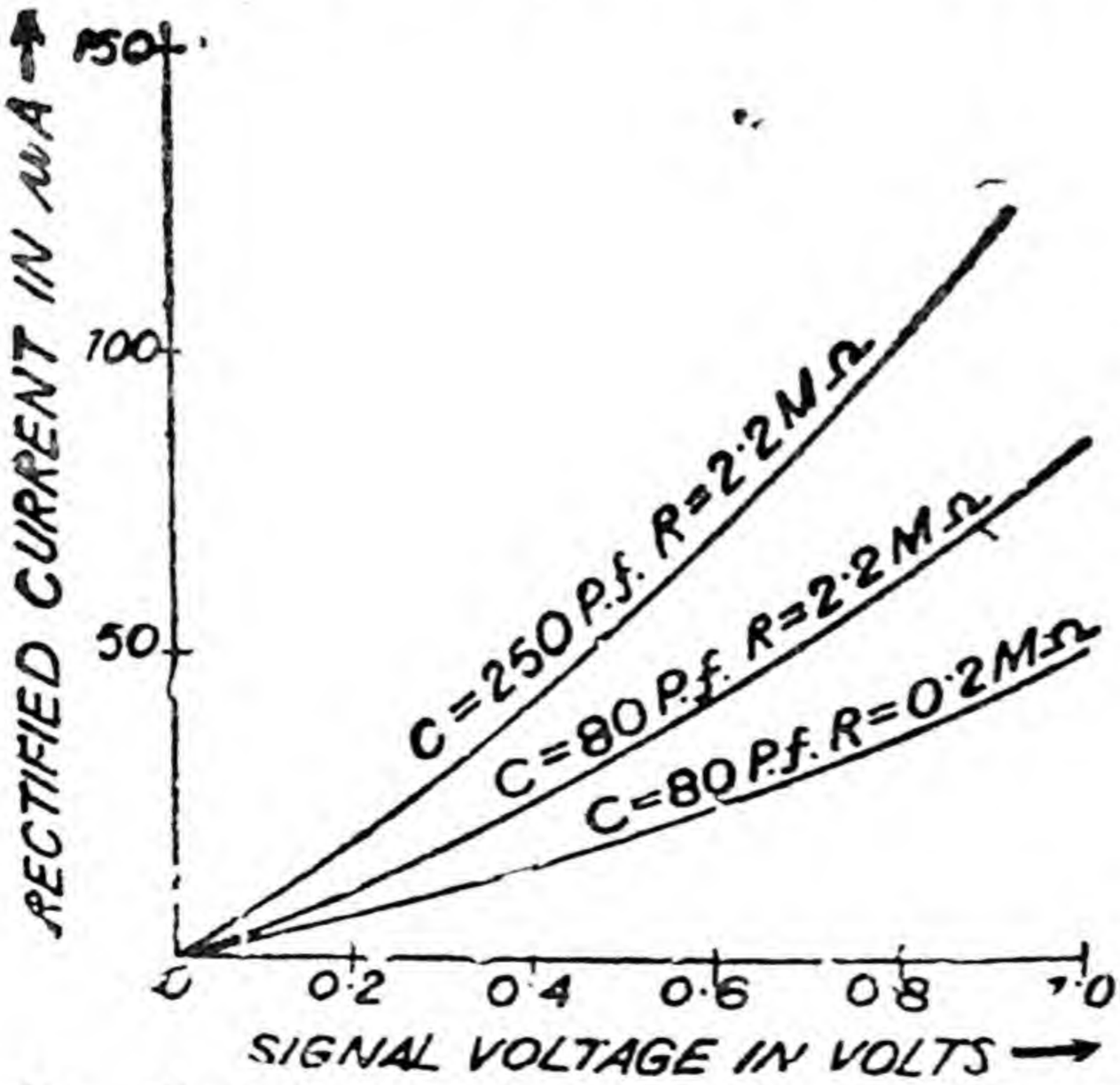
(*i*) Starting with low values of *C* and *R*, the ammeter in anode circuit is backed off to read zero with no applied signal.

(*ii*) Now apply the signal and increase it from zero to one volt approximately. Rectified current is noted for each setting of the signal voltage.

(*iii*) Repeat the test with higher value of grid leak resistance *R*.

(*iv*) Again repeat the test with a higher value of capacitance *C* in conjunction with the high value of *R*.

Curves of rectified current to a base of signal voltage for the different values of *R* and *C* are plotted as shown in fig. 34.



Fig, 34. Cumulative grid detector characteristics.

It can be seen that senstivity is greater for low signal voltages. For the same signal voltage, distortion is less when the load R is increased. The effect of increasing capacitance is, however, reverse *i.e.*, with same signal voltage distortion is increased if capacity of condenser, C, is increased.

EXERCISES AND PROBLEMS

1. Explain the action of a linear diode detector in the reception of an amptitude modulated wave. How would you obtain the dynamic characteristic of a diode detector ? What information can be obtained from them ? Draw the conventional circuit of a diode detector with A V.C.
2. Explain the square law and linear methods of detection, giving circuit diagrams. State their relative disadvantages and advantages.
3. A diode detector is used to detect the amplitude modulated signal. Find an expression for the time constant of the anode load for which there will be faithful reproduction.
4. Describe the action of the Foster Seeley discriminator circuit for the following signal conditions :
   (*a*) Centre frequency signals.
   (*b*) Signals above the centre frequency.
   (*c*) Signals below the centre frequency.
5. Describe the principle of operation of the ratio detector.
   In an envelope detector circuit, with $R=250$ kilohms and $C=260$ pf plot as a function of modulation frequency, the highest value of modulation index that can be employed without introduction of distortion due to inability to follow the modulation envelope.
6. A carrier voltage of 13 volts r. m. s. is applied to the linear diode circuit The load R is 250,000 ohms and the average current over a carrier frequency cycle is 30μ amp.
   (*a*) Find the efficiency of detection.
   (*b*) Find the voltage available for A. V C. use
7. Tick the correct answer :
   In a ratio detector—
   (*i*) the linearity is worse than in a phase discriminator:
   (*ii*) stabilization against si nal strength variation is provided;
   (*iii*) the output is twice that obtainable from a similar phase discriminator;
   (*iv*) the circuit is the same as in discriminator except that the diodes are reversed.

(Ans. *i*)

32 TRANSMITTERS AND RECEIVERS

BROADCAST TRANSMITTING CIRCUITS

In this chapter we shall discuss various parts of a transmitting circuit. The design of radio circuit largely depends upon the frequency to be used. In many applications of radio transmission, the carrier frequencies utilised cover a wide range. A classification of radio waves into frequency groups is given below :

Table-1 : CLASSIFICATION OF RADIO WAVES

Abbreviation	Class	Frequenncy range	Wavelength ran e
VLF	Very low frequency	10—30 kc/s	300,000—10,000 metres
LF	Low freq.	30—300 kc/s	10,000—1,000 m.
MF	Medium frequency	300—3,000 kc/s.	1,000—100 metres
HF	High freq.	3—30 mc/s.	100—10 metres
VHF	Very high frequency	30—300 mc/s	10—1 metres
UHF	Ultra high frequency	300—3000 mc/s.	100—10 cm.
SHF	Super high frequency	3,000—30,000 mc/s.	10—1 cm.

For interference free reception in a given area, it is necessary that there should be no sideband overlapping in the transmitted intelligence. To avoid this overlapping. transmitting sources that simultaneously cover the same area must use quite different carrier frequencies. The following Table-2 shows the radio frequency band widths covered by different types of signals and the carrier frequencies at which the signals are commonly transmitted.

According to use, for the transmission of information, there are two types of transmitters :

Table 2. R. F. TRANSMISSION BAND WIDTHS AND COMMON CARRIER FREQUENCIES OF DIEFERENT SIGNALS.

Signal Type	Radio frequency transmission band width	Typical carrier frequency
Telegraphy signals	80 c/s.-2 kc/s.	18 kc/s.-30 mc/s.
Telephony signals	10 kc/s.	500 kc/s-30 mc/s.
Telephony signals	150 kc/s.	88 mc/s-108 mc/s
Fascimile signals	6 kc/s.	500 kc/s-30 mc/s.
Television signals	6 mc/s.	54 mc/s-216 mc/s
Radar signals	2-10 mc/s.	200 mc/s-30,000 mc/s.

(*i*) Radio Telegraph Transmitter.

(*ii*) Radio Telephone Transmitter (Broadcast).

The difference between radio telegraph transmitter and radio telephone transmitter is that the latter is fitted with modulation equipment whereas in former, modulation system is replaced by keying system whereby the radiation of energy is periodically turned off and on inaccordance with the characters of telegraph code.

We shall discuss only radio telephone transmitter. A transmitter is essentially a device for producing radio frequency energy that is controlled by the intelligence to be transmitted.

According to methods of modulation, radio telephone transmitters can be classified as follows :

(*A*) Amplitude modulated transmitter.

(*B*) Frequency modulated transmitter.

(*C*) Phase-modulated transmitter.

Class (C') type of transmitter is not of practical importance.

32·1. CARRIER FREQUENCY REQUIREMENTS OF A RADIO TRANSMITTER :

Few main requirements of a radio transmitter regarding the carrier frequency are being described as under :

(1) Carrier frequency generated must remain stable at the specified value : To every transmitting station, specified carrier frequency is assigned for operation. Obviously, to avoid the interference with the signals of other transmitting stations alloted carrier frequency should remain stable. For this purpose master

oscillator is used to generate the sub-harmonic of the assigned carrier frequency ; the stage of harmonic generators is then employed to achieve that final value of the carrier frequency.

(2) **Carrier frequency should be readily adjustable :** In radio-transmitters employing crystal controlled master oscillator, carrier frequency can not readily be adjusted because this would require a change of crystal and hence a subsequent change in the tuning of circuits of next stages. If instead, *LC* circuits are used in master oscillator, variation in *L* or *C* can serve the purpose of frequency adjustment easily.

(3) **Frequency drift should be extremely small :** The maximum frequency drift allowed in radio transmitters is ± 20 c/s for medium wave transmitters and about $\pm \cdot 002\%$ for short wave and ultra high frequency transmitters.

If *LC* circuit is used as tank circuit in master oscillator, then to avoid such frequency drifts, it is essential to keep the tank circuit in a constant temperature chamber because inductance and capacitance of tank circuit are both subject to variation due to temperature changes.

Further, interelectrode capacitance of the valve used also varies with temperature. To reduce the resulting drift in frequency due to the variation in tube capacity, the oscillator current should be arranged such that the amount of electrostatic energy stored in interelectrode capacitances be very small compared to the total electrostatic energy stored in the resonant circuit. This requires a low resonant frequency value and a high circuit *Q*. For the resonant frequency to be low, master oscillator produce low carrier frequency which with the amplification of harmonic generators is raised to the final value of carrier frequency that is assigned to the particular transmitter.

Coupling of load with tank circuit also drifts the frequency. For this drift to be small, it is essential that coupling should be small and the load should be purely resistive. To eliminate this effect, a buffer or isolating amplifier stage is used immediately after the master oscillator stage. This buffer stage is operated as low voltage amplifier and consequently its grid circuit and plate circuit of master oscillator remain open. Thus this stage serves to eliminate the load effect on plate circuit of master oscillator.

32·2. AMPLITUDE MODULATED TRANSMITTER :

In amplitude-modulated transmitter, the energy is controlled by amplitude modulation. The block diagram is shown in fig. 1.

Description of parts : There are two channels in A M. Transmitter.

(1) **R.F Chanhel :** The radio frequency channel of the transmitter induces the circuits wherein the carrier wave is generated, amplified and modulated.

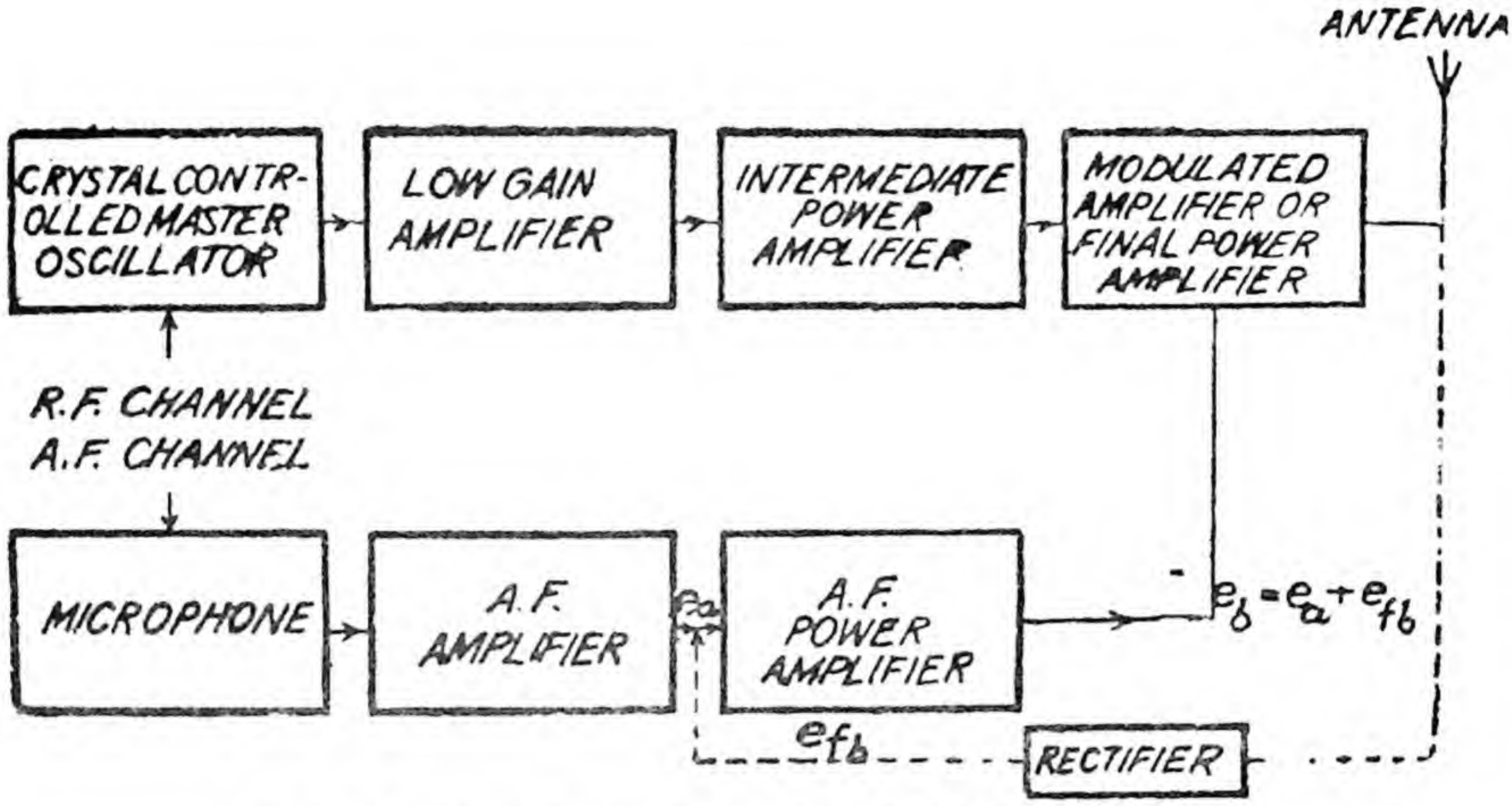


Fig. 1. Block diagram of A. M Transmitter.

(*a*) *Master Oscillator* : Figure 2 shows the circuit of a crystal-controlled master oscillator. The temperature of the crystal is kept constant by means of constant-temperature oven. The variable capacitor in the grid circuit is used for minor adjustment in the operating frequency. The variable plate capacitor is adjusted to a value corresponding to cross point on the sloped portion of plate current-capacitance curve [fig. 2 (*b*)] which gives a stable plate

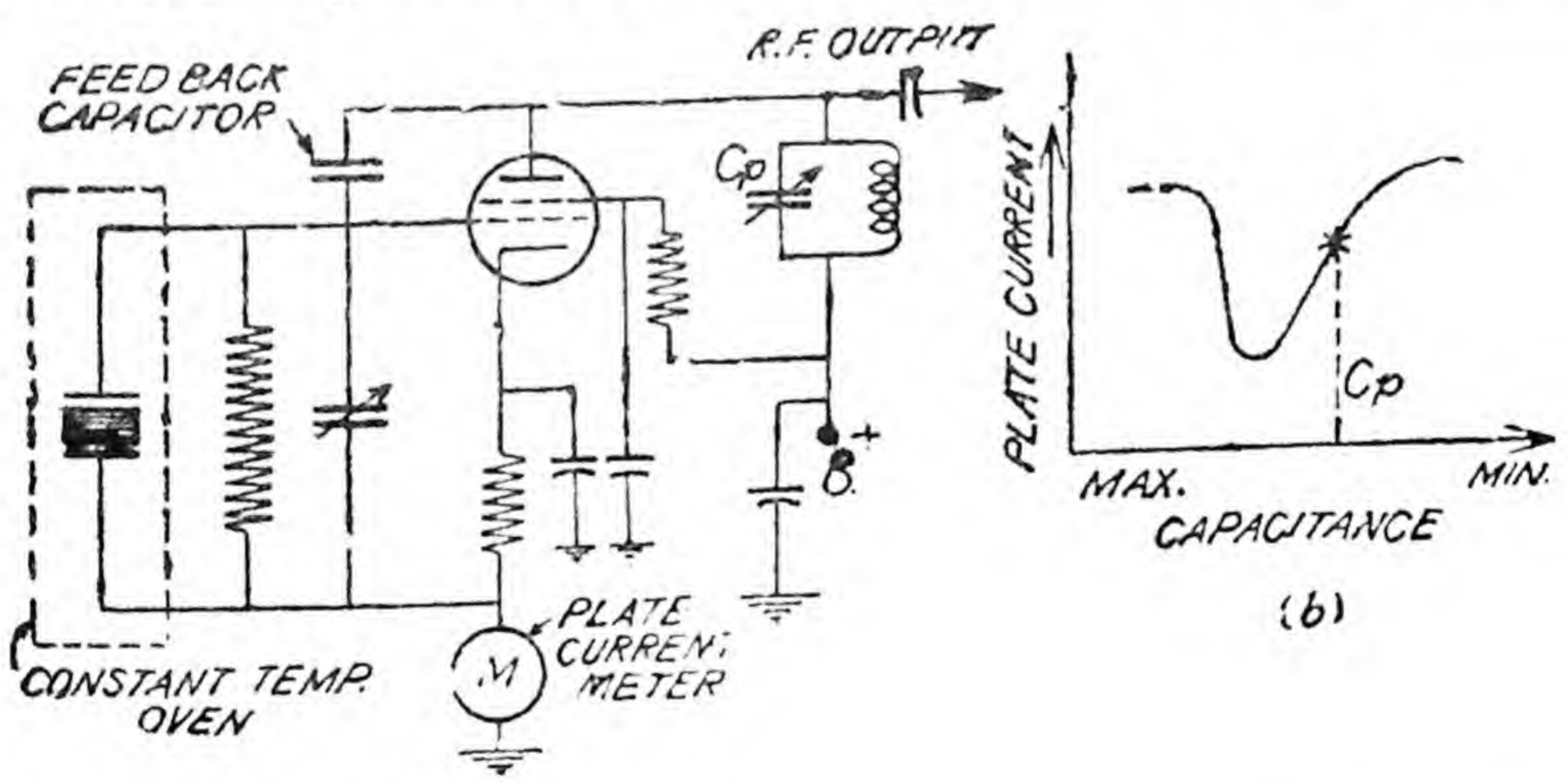


Fig. 2. (*a*) Circuit of crystal-controlled master oscillator. (*b*) Tuning curve.

operation. For plotting this curve, the plate current is measured by plate current milliammeter with the variation of capacitor in plate circuit. The dip in the curve signifies the position where the oscillations start.

(*b*) *Buffer and Intermediate Power Amplifier Stages* : Low gain (buffer) amplifier serves to isolate the oscillator from the succeeding amplifier stages. This amplifier usually employs a tetrode. The tube is biased approximately to cut off by a fixed bias source and the amplitude of the sinusoidal input voltage is adjusted so that it does not exceed the bias voltage. Under these conditions,

the load on the oscillator may be kept essentially constant because the tube does not draw any grid current.

Buffer amplifier is followed by an intermediate power amplifier which operates as a class *C* amplifier and designed to give sufficient amount of power to the succeeding stages.

The circuit diagram of buffer amplifier alongwith intermediate power amplifier is shown in fig. 3.

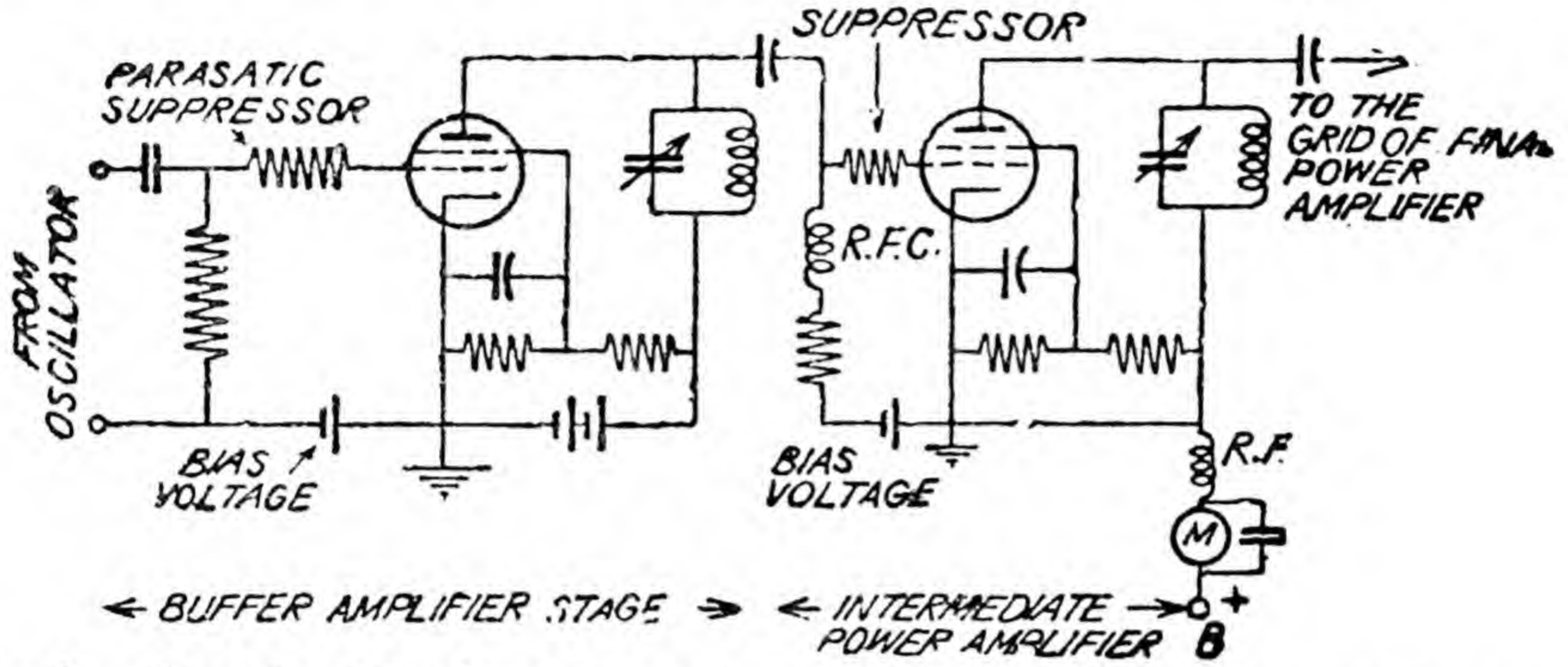


Fig. 3. Circuit of Buffer Amplifier and Intermediate Power Amplifier.

Undesired or parasitic oscillations are suppressed by inserting a small resistance in series with the grid of each tube which serves as a damping resistor in an undesired oscillatory circuit formed by the grid cathode capacitance and grid cathode leak inductance.

(*c*) *Freyuency Multipliers* : System which are rich in the generation of harmonics can be used as frequency multipliers. All the single tube amplifiers possess this characteristic due to the non-linearity existing in the plate current curve when the signal e.m.fs. are applied in the grid. Therefore one or more intermediate amplifier employed in r.f. channel will act as a frequency multiplier.

There are two distinct types of frequen y multipliers depending upon whether the distortion takes place in the plate circuit or in the grid circuit. The former is known as a plate distortion harmonic generator and the latter as a grid distortion harmonic generator.

In the operation of class *C* amplifier, we note that p.ate current contains several harmonics of the frequency of the sinusoidal input voltage. Thus an oscillatory current will be sustained in the plate circuit at a harmonic corresponding to which the circuit is tuned. Thus the output will be at a frequency equal to multiple of the input voltage frequency

(*d*) *Final Power Amplifier Stage* : The total power delivered to the antenna is determined by the final stage of R.F. channel. The final power amplifier operates at highest power level to increase the power output of the transmitting system.

If the modulating voltage is applied to the final r f. stage, then it is said to be *High Level Modulation* as shown in block diagram and if modulating voltage is applied to preceding stage (at which power is low) then it is said to be *Low Level Modulation*.

The advantage of high level modulation is that the final stage can be operated as a class *C* amplifier and thus maximum efficiency can be obtained.

The circuit of final power amplifier is drawn in A.F. channel with audio amplifier unit.

(2) A. F. Channel : This section generates the modulating signal and amplifies it

(*a*) *Microphone* : It converts sound energy into electrical energy. Various type as carbon microphone, capacitor microphone, velocity microphone can be used.

(*b*) *Audio Amplifier Unit* : The amplifier stage is usually a pentode voltage amplifier. This develops appropriate amount of excitation voltage applied to the grid of power amplifier. The audio frequency power is then increased by power amplifier to modulate carrier. The following figure 4 shows the complete circuit of modulation and modulated amplifiers.

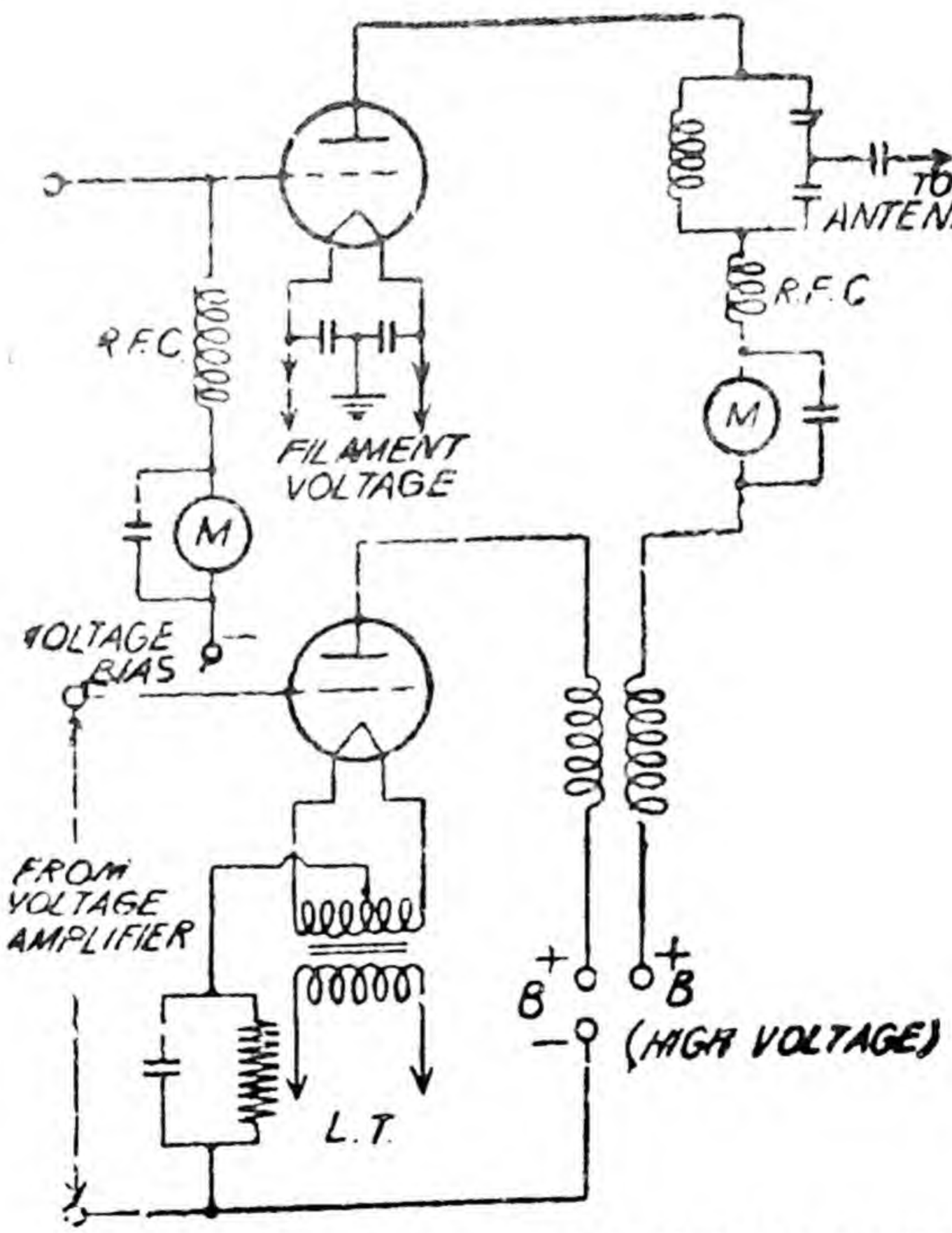


Fig. 4. Audio power amplifier and modulated amplifier circuits.

(3) Circuits for Improving Transmitter Performace :

(i) Negative feed back-circuits are used to improve the performance of transmitter. A portion of modulated output is rectified and introduced in opposite phase to the original input *i.e.* in the input of modulating amplifier as shown by doted lines in the block

diagram. This process of negative feed-back avoids the amplitude, frequency and phase distortion which may otherwise occur due to preceding stages of amplifiers in R.F. channel.

(*ii*) In order to minimise the possibility of over modulation, limiting amplifier may be included in A.F. channel which automatically reduces the channel gain when the signal amplitude becomes excessive.

(*iii*) Volume compressors are used to raise the average degree of modulation.

(4) Transmitter Power Supply. Radio transmitters are classified in regard to operating power as follows :

(*a*) *Direct current operated.* The transmitter employs direct current throughout filament, bias and plate supply.

(*b*) *Alternating current operated.* In this operation, the filament supply is an alternating current, while plate and grid bias voltages are supplied by means of rectifier.

A d.c. voltage power supply for transmitters may be obtained from electromechanical d.c. generators, and high vacuum tube rectifiers.

(5) Transmitter Monitoring Equipment Various meters are inserted in the system to check up necessary details and operations such as plate and grid current meters, field strength meters to determine radiation pattern of the antenna system and modulation monitors to check up the degree of modulation.

32 3. FREQUENCY MODULATED TRANSMITTER :

Frequency modulated transmitters are widely used at frequencies above 40 mc/s for radio relay system and other purposes.

F.M Transmitters for very high and ultra high frequency bands : Generally two types of f.m. transmitters are used :

(*a*) F.M. Transmitter employing phase modulation,

(*b*) F.M. Transmitter using reactance tube.

We shall discuss first the common features of both the types as follows :

(*i*) In both cases, modulation is performed at low power level. The required transmitter power and frequency are, then, generated by employing harmonic generators and a chain of class *C* amplifiers.

(*ii*) Class *C* amplifiers in f.m. transmitters can be employed after modulation stage because frequency deviations are not affected due to the distortion in amplitude caused by class *C* amplifiers.

(*iii*) Harmonic generators are largely employed in transmitter circuit to produce reasonable frequency deviation. In most f.m transmitters, by some means, frequency deviation produced by a given modulating signal can be made greater than at the

medium or low modulating frequencies. This is known as pre-emphasis which reduces background noise.

We shall first describe f.m. transmitter employing phase modulation.

32·3-1. PHASE MODULATED TYPE F.M. TRANSMITTER :

As discussed in the chapter of modulation, if modulating voltage amplitude, applied to the modulator, is made to vary inversely as the modulating frequency, the phase modulator output will have the characteristics of a f.m. wave.

The block diogram of phase modulated type f m. transmitter is shown in fig. 5.

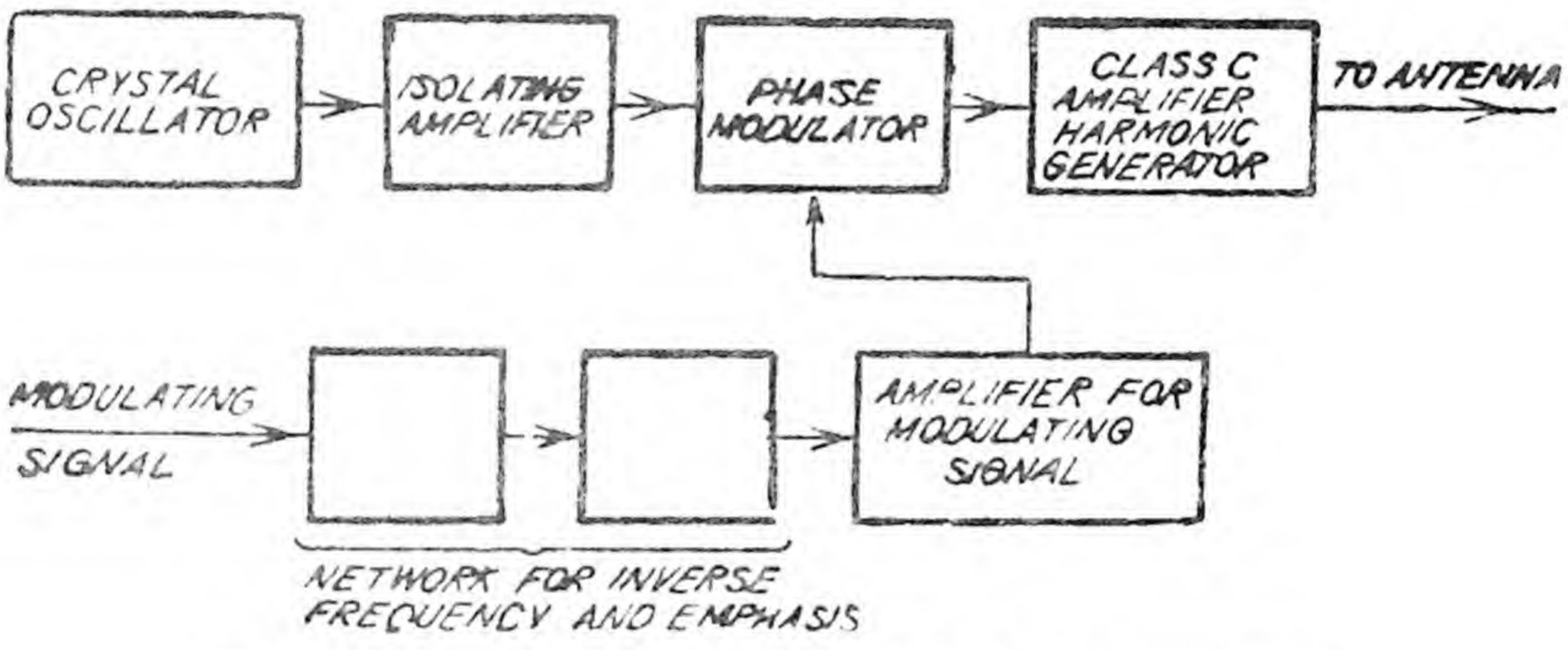


Fig. 5. Phase modulated type of f.m. transmitter.

Description of Parts :

(*i*) Crystal oscillator,

(*ii*) Isolating amplifier, and

(*iii*) Phase modulator and network for inverse frequency.

The following fig. 6 shows the basic circuit of Armstrong phase modulator alongwith the network for inverse frequency.

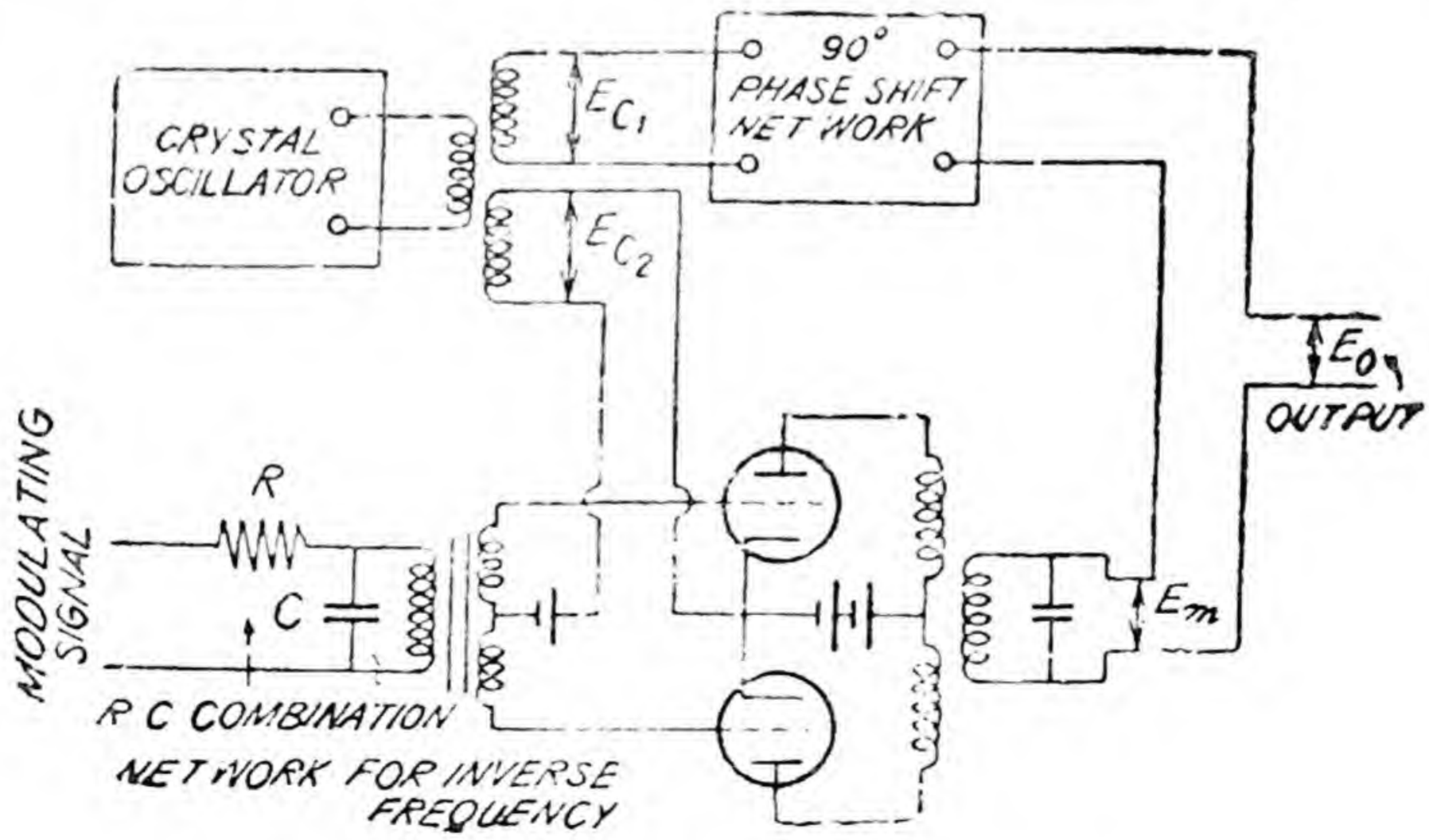


Fig. 6. Armstrong phase modulator circuit.

Process of Phase Modulation : In Armstrong modulation method, phase modulation is produced by special amplitude modulation method. The modulating voltage amplitude modulates the carrier in a balanced modulator to produce the two sidebands alone. Both these sidebands are shifted in phase by 90° and are added to the carrier. This gives rise to pure phase modulation of the carrier provided that the resultiug phase deviation is less than 0·5.

Equation for phase modulated wave is

$$e = E_c \sin (\omega_c t + \triangle\phi \sin \omega_m t)$$
$$= E_c J_0 (\triangle\phi) \sin \omega_c t + E_c J_1 (\triangle\phi) [\sin (\omega_c + \omega_m) t - \sin (\omega_c - \omega_m) t]$$
$$+ E_c J_2 (\triangle\phi) [\sin (\omega_c + 2\omega_m) t + \sin (\omega_c - \omega_m) t] + \ldots$$

For the particular case of $\triangle\phi = 0{\cdot}5$, the values of $J_n (\triangle\phi)$ can be determined from the curves of fig. 19 (art. 30·5). These are

$$J_0 (0{\cdot}5) = 0{\cdot}938$$
$$J_1 (0{\cdot}5) = 0{\cdot}242$$
$$J_2 (0\ 5) = 0{\cdot}030$$

and $$J_n (0\ 5) \approx 0 \text{ for } n > 2.$$

The equation for phase modulated wave then becomes

$$e = 0{\cdot}938\ E_c \sin \omega_c t + 0{\cdot}242\ E_c [\sin (\omega_c + \omega_m) t - \sin (\omega_c - \omega_m) t]$$
$$+ 0{\cdot}030\ E_c [\sin (\omega_c + 2\omega_m) t + \sin (\omega_c - 2\omega_m) t] + \ldots$$
$$\approx 0{\cdot}938\ E_c \sin \omega_c t + 0{\cdot}242\ E_c [\sin (\omega_c + \omega_m) t - \sin (\omega_c - \omega_m) t]$$

where the second side band components which have amplitude of $0{\cdot}03 E_c$ have been neglected. Approximation is more valid when $\triangle\phi < 0{\cdot}5$. Equation can be further put in the form

$$e = 0{\cdot}938\ E_c \left[\sin \omega_c t + \frac{\triangle\phi}{2} \sin (\omega_c + \omega_m) t - \frac{\triangle\phi}{2} \sin (\omega_c - \omega_m)\ t \right] \quad \ldots(1)$$

The corresponding equation for a.m. wave is

$$e = E_c \left[\sin \omega_c t - \frac{m_a}{2} \cos (\omega_c + \omega_m) t + \frac{m_a}{2} \cos (\omega_c - \omega_m) t \right]. \quad \ldots(2)$$

From eqs. (1) and (2), it is obvious that two equations are identical provided :

(*i*) $\triangle\phi$ is made equal to m_a, and

(ii) side bands of amplitude modulated wave are shifted in phase by 90° with respect to the carrier.

This principle of conversion of amplitude modulated wave into phase modulated wave is illustrated in figs. 7 (*a*) and 7(*b*). Vector *AB* represents carrier voltage E_c, and two vectors *BC* and *BD* represent the sideband components. *BC* and *BD* vectors rotate in opposite directions at the rate ω_m rad./sec. In amplitude modulated

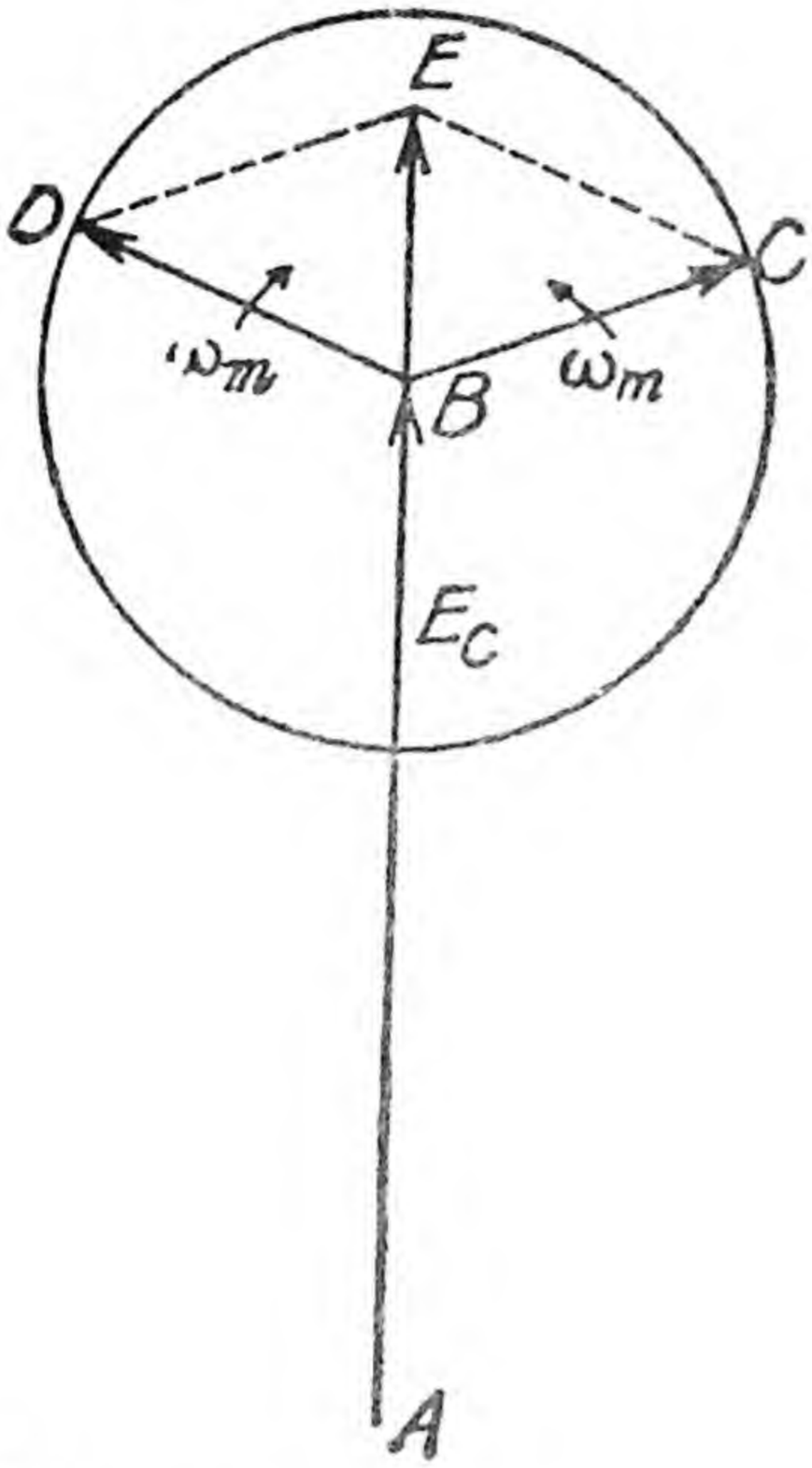


Fig. 7. (*a*) Vector representation of an amplitude modulated wave

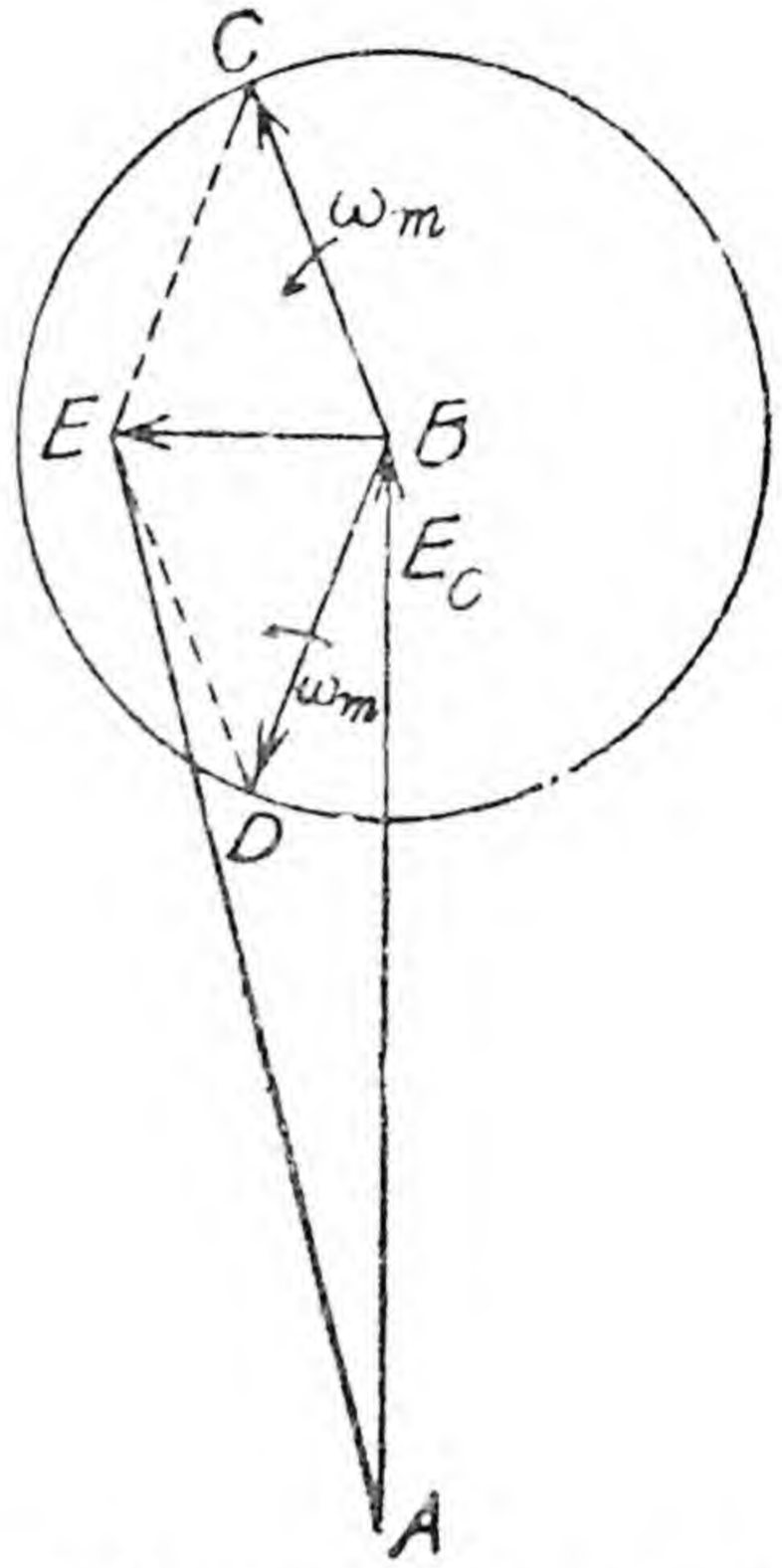


Fig. 7. (*b*) Vector representation of phase modulated wave obtained by rotating side bands in fig. 7(*a*) through 90°.

wave resultant BE of sideband components BC and BD is always along the carrier AB. Therefore there will be no phase or frequency modulation, *i.e.*, purely amplitude modulation represented by vector AE.

In fig. 7 (*b*), keeping the carrier voltage unaltered, sideband vectors are advanced in phase by 90°. Consequently, resultant vector AE now differs in phase with unmodulated carrier voltage vector AB. Thus a phase deviation ($\triangle\phi$) is produced which is maximum when two sidebands are in phase (*i.e.*, along the same direction). It may be noted that AE is slightly greater than AB and therefore along with phase modulation, some amplitude modulation is also present which is small if phase deviation $\triangle\phi$ is small.

A circuit utilising Armstrong principle for the generation of frequency modulated wave is shown in fig. 6.

Network for Inverse Frequency :

We shall first explain how an RC combination, shown in fig. 6. canbe used as an inverse frequency network. It has been shown that if time constant RC (say $R=200\ K\Omega$ and $C=2\mu F$) is very

large compared to the interval of input signal, then output voltage, e_0 across capacitor will be

$$e_0 = \frac{1}{RC}\int e_i \, dt$$

The input signal voltage is a sinusoidally varying modulating voltage, *i.e.*,

$$e_i = E_m e^{j\omega_m t}$$

so that

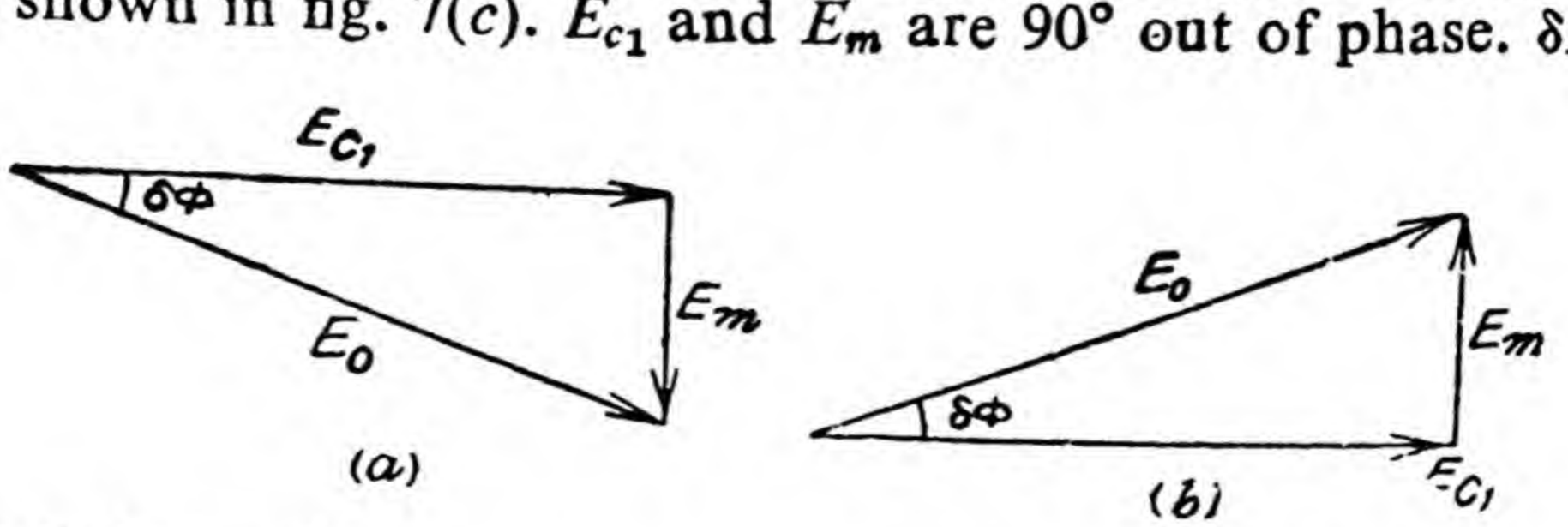
$$e_0 = \frac{1}{RC}\,\frac{E_m e^{j\omega_m t}}{j\omega_m}$$

or $$| e_0 | = \frac{e_i}{\omega_m RC},$$

that is, if modulating frequency increases, output across capacitor decreases. This output voltage is applied to grids of the two tubes shown in fig. 6 through the transformer and phase modulates the carrier to produce frequency modulation.

Function of the circuit : The outputs E_{c1} and E_{c2}, derived from crystal-controlled oscillator are made to have a phase difference of 90° by means of phase shift network. E_{c2} is applied to the grids of the two tubes connected in symmetrical circuit arrangements as shown in fig. 6. When the instantaneous potentials of the two grids are equal and vary in synchronism, the r.f. voltage across the common load impedance of both tubes will be zero because the circuit is quite symmetrical. When the modulating signal is applied then the potentials of the grids vary at an audio frequency rate with respect to each other. Thus the r.f. voltage, E_m will vary cyclically in amplitude. The instantaneous phase of E_m with respect to E_{o_1} reverses ea ch time the modulating signal goes through zero. The amplitude of E_m depends upon the strength of modulating signal. This voltage then combines with E_{o_1} to form the modulator output E_0.

The phase relations among E_m, E_{c_1} and the modulator output E^0 are shown in fig. 7(*c*). E_{c_1} and E_m are 90° out of phase. δE is the



) E_0 *lagg ng* E_{c_1}.

(*b*) E_0 *leading* E_{c_1}

Fig. 7. (*c*) Phase relation in E_m, E_0 and E_{c_1}.

s tantaneous phase difference between E_0 and E_{c_1} which is directly

proportional to the instantaneous amplitude E_m for small phase deviation angles. As a result, the instantaneous phase deviation will be proportional to the instantaneous value of modulating signal. Since instantaneous value of modulating signal varies inversely as the modulating frequency, a frequency modulated wave is produced by the process of phase modulation.

(iv) **Pre-emphasis Network** : A relatively large amount of energy in speech or music is contained in the low frequency portion of audio spectrum. Since the energy in high frequency portion is small, the high frequency components ordinarily produce frequency deviation far below the maximum allowable value of 75 kc/s. This causes interference of *high energy* modulating frequencies to be more effective. In order to overcome this unequal distribution of energy in audio frequency spectrum, pre-emphasis network is used in f.m. transmitters. Pre-emphasis device accentuates the amplitude of high frequency components with respect to the low frequency components. In figure 8, a simple pre-emphasis circuit is shown.

In the plate circuit of audio frequency amplifier, an inductance and resistance are connected in series.

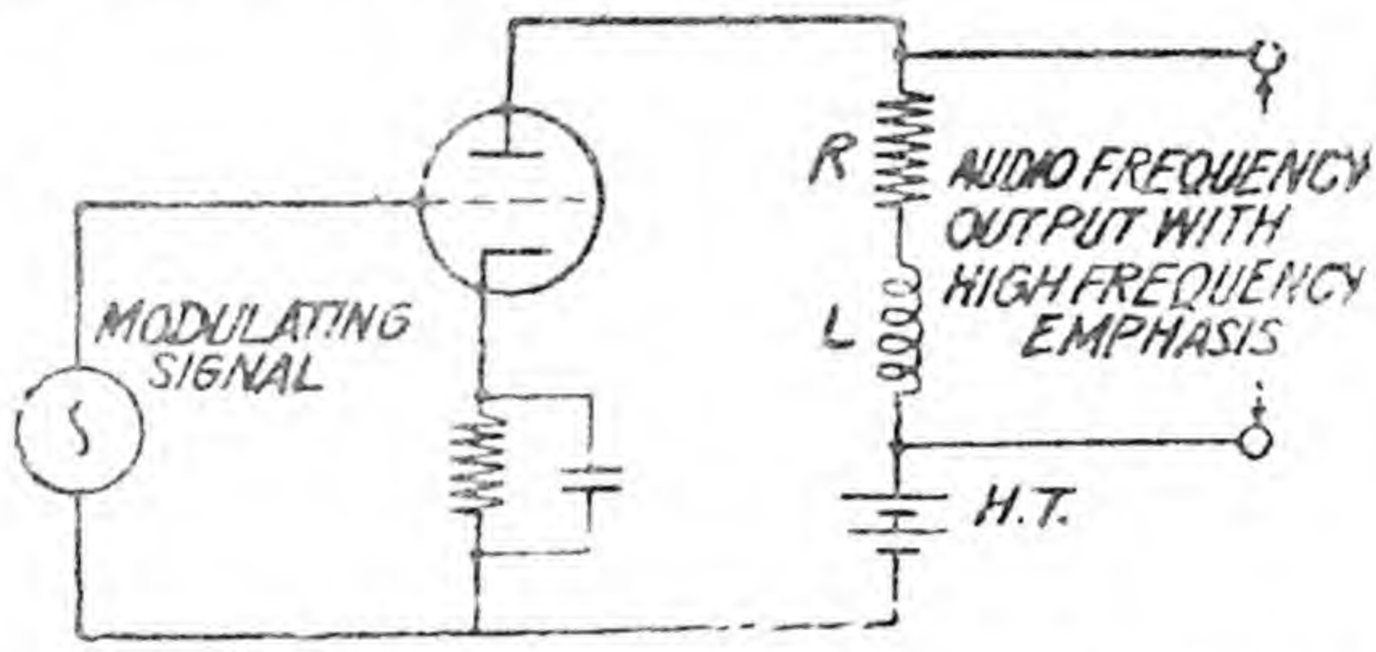


Fig. 8(*a*) Audio-amplifier with frequency emphasis.

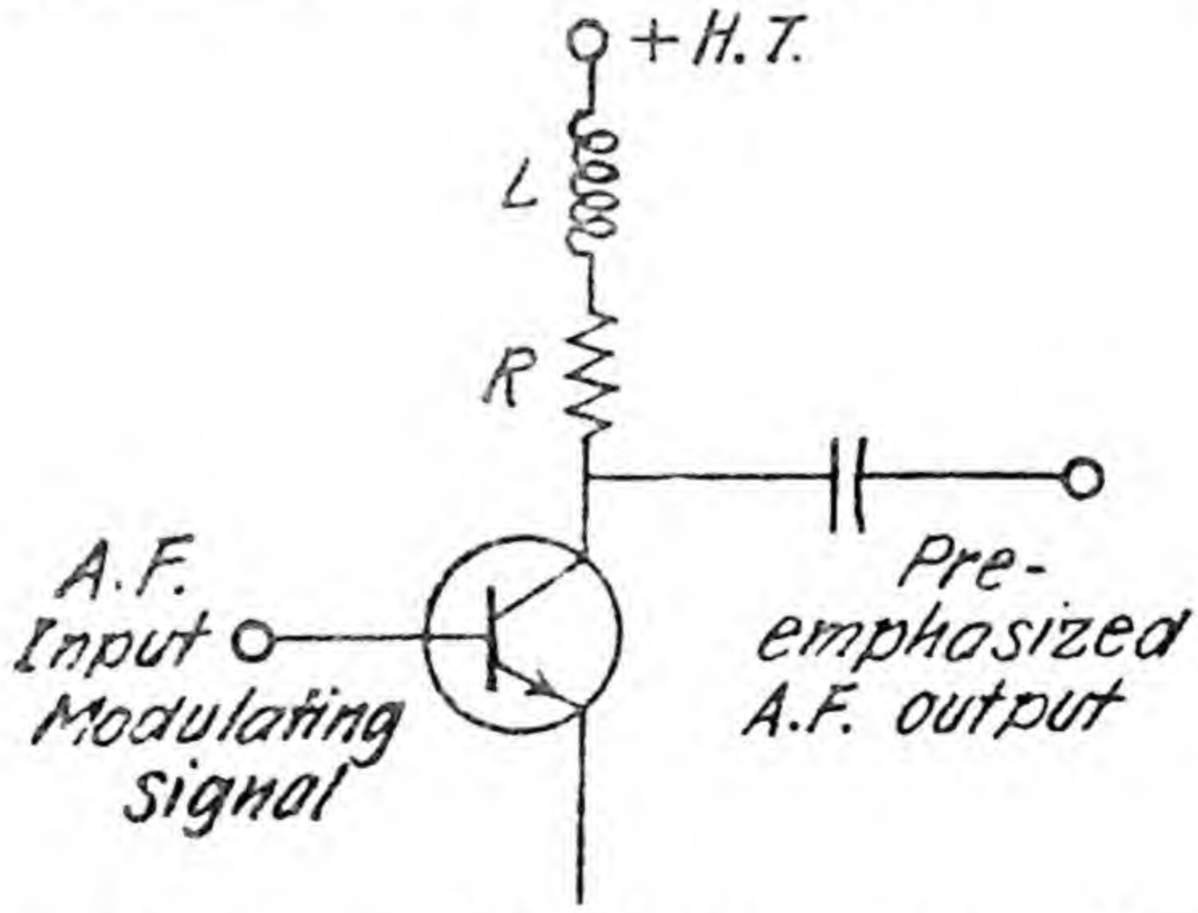


Fig. 8 (*b*) Transistor amplifier with frequency emphasis.

Since the inductance offers high reactance to high frequency components in comparison to low frequency components. the output voltage developed across the series combination of L and R will be relatively high for high frequency components.

Advantage of phase modulated type F.M. Transmitter : Main advantage of phase modulated type f.m. transmitter is that the carrier frequency is derived directly from a crystal oscillator and, is therefore, independent of tube characteristics and circuit adjustment. The carrier frequency employed in the process of modulation is thus very stable.

Disadvantage : The disadvantage of phase modulation is that at low modulation frequencies, the maximum frequency deviation produced by the modulator is small. Therefore, in order to produce reasonable deviation various stages of harmonic generators are employed in the transmitter.

26·3·2. F.M. TRANSMITER USING REACTANCE TUBE :

The block diagram of reactance tube type f.m. transmitter is given in fig. 9.

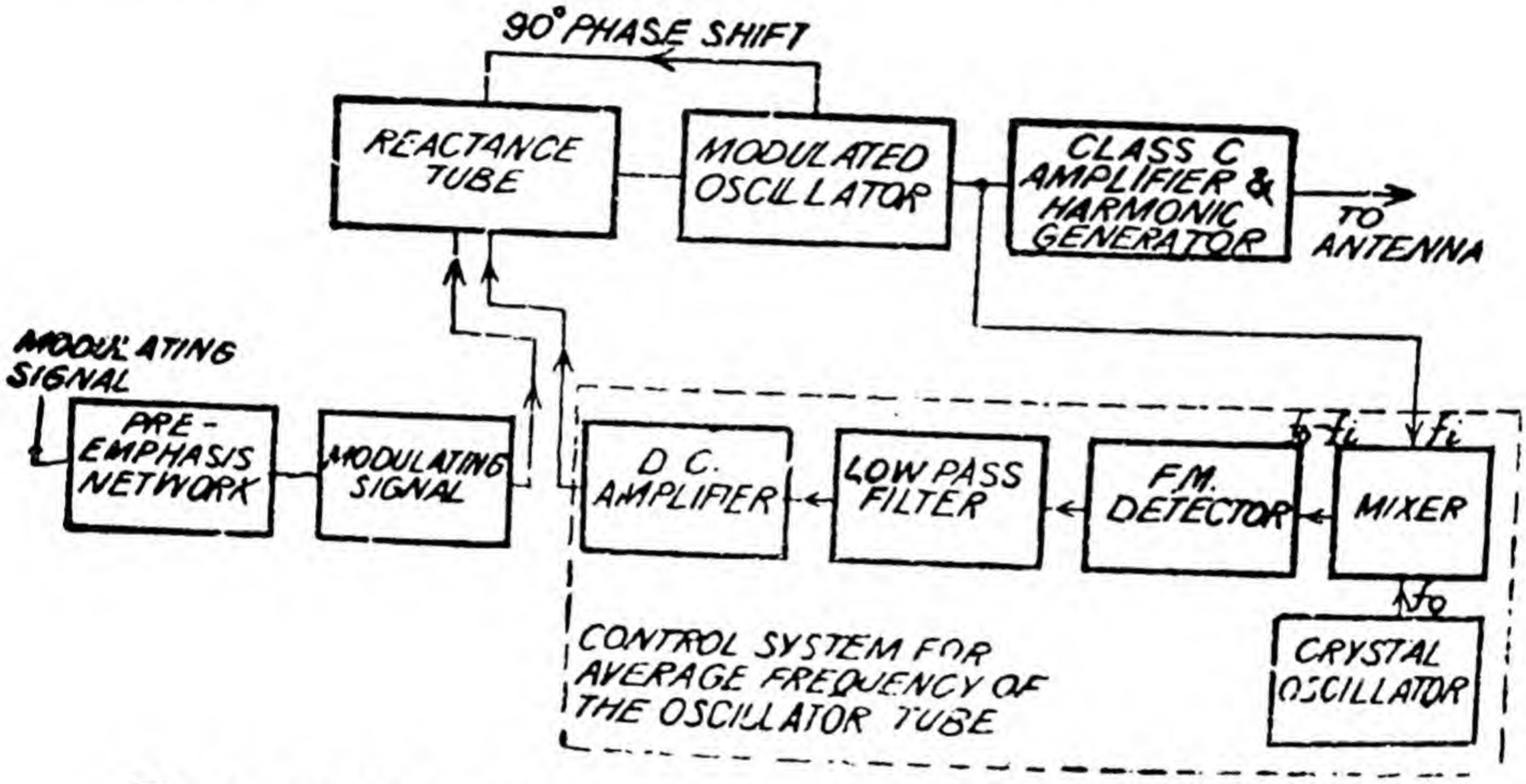


Fig. 9. Block diagram of f.m. transmitter using reactance tube.

Description of Main Parts : We have described some parts of above transmitter in previous articles, *e.g.* pre-emphasis network, audio frequency amplifier. Here three systems are being discussed in brief :

(i) Pre-emphasis circuit : Though we explained this network in previous article but here we shall explain it in detail.

As stated earlier, a relatively small amount of energy is contained in high frequency portion of audio spectrum. Since these high frequency components have small amplitudes, they produce frequency deviation far less than the maximum permitted value of 75 kc/s. Consequently, the signal strength at these high modulating frequencies is very low relative to the tube and circuit noise which are uniformly distributed over the entire spectrum occupied by the channel. Therefore to improve signal to noise ratio at these frequencies, pre-emphasis circuits are used to accentuate their

amplitudes. At the receiver this emphasis on amplitudes is removed to restore the original amplitude with the help of de-emphasis circuit.

The pre-empasis circuit may use either *RL* or *RC* circuit as shown in figs. 10 (*a*) and 10 (*b*).

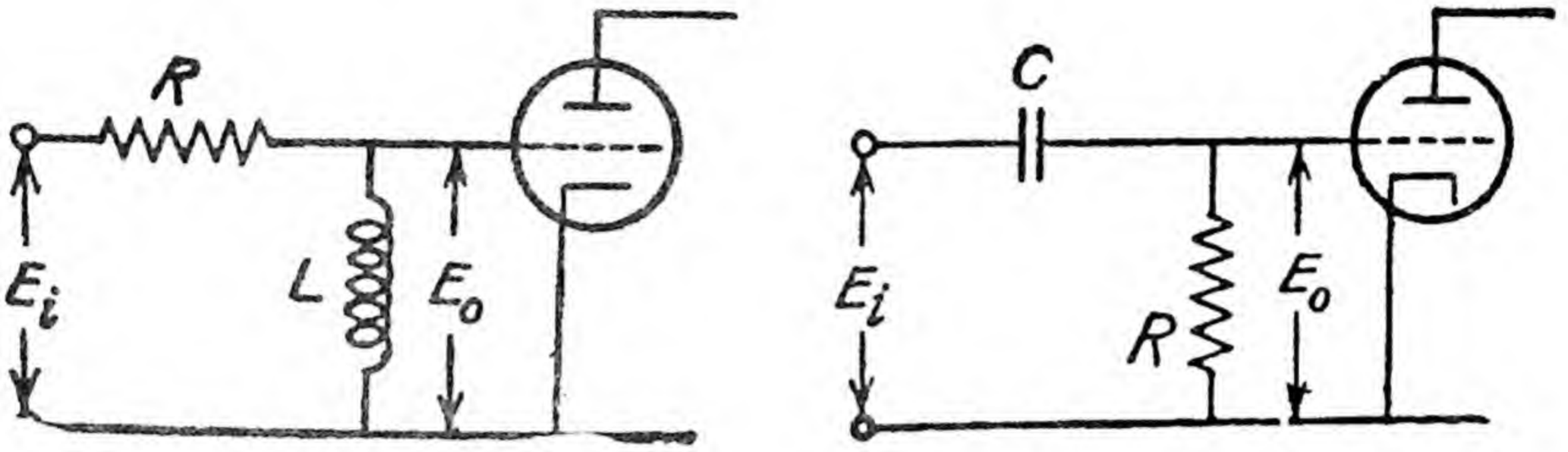


Fig. 10. (a) (b) Pre-emphasis circuits.

The transfer ratio in case of *RL* circuit is given by

$$\frac{E_0}{E_i} = \frac{j\omega L}{R+j\omega L} = \frac{1}{1+\frac{R}{j\omega L}}$$

$$= \frac{1}{\sqrt{\left[1+\left(\frac{R}{\omega L}\right)^2\right]}} \Big/ \underline{\tan^{-1}\left(\frac{R}{\omega L}\right)}$$

where ω is operating frequency. If reference frequency (designed value),

$$\omega_1 = \frac{R}{L},$$

then

$$\frac{E_0}{E_i} = \frac{1}{\sqrt{\left[1+\left(\frac{\omega_1}{\omega}\right)^2\right]}} \Big/ \underline{\left(\tan^{-1}\frac{\omega_1}{\omega}\right)} \qquad \ldots(1)$$

For *RC* pre emphasis circuit,

$$\frac{E_0}{E_i} = \frac{R}{1+\frac{1}{j\omega C}} = \frac{1}{1+\frac{1}{j\omega CR}}$$

$$= \frac{1}{\sqrt{\left[1+\left(\frac{1}{\omega CR}\right)^2\right]}} \Big/ \underline{\left(\tan^{-1}\frac{1}{\omega CR}\right)}$$

If reference frequency $\omega_1 = \frac{1}{CR}$, then

$$\frac{E_0}{E_i}=\frac{1}{\sqrt{\left[1+\left(\frac{\omega_1}{\omega}\right)^2\right]}}\Big/\left(\tan^{-1}\frac{\omega_1}{\omega}\right) \qquad \ldots(2)$$

The value of reference frequency is chosen as

$$\frac{1}{\omega_1}=75\ \mu\ \text{sec}$$

or $$f_1=2100\ \text{c/s},$$

Therefore for $f=\frac{\omega}{2\pi}=2100$ c/s, the transfer ratio in case of both the circuits is

$$\left|\frac{E_0}{E_i}\right|=\frac{1}{\sqrt{\left[1+\left(\frac{2100}{2100}\right)^2\right]}}=\frac{1}{\sqrt{2}}=70{\cdot}7\%.$$

But if $f=21{,}000$ c/s, then

$$\left|\frac{E_0}{E_i}\right|\approx 1,$$

that is, for frequencies equal to 21 kc/s or above, the output voltage is almost equal to input voltage, whereas for 2·1 *kc/s* output was only 70·7% of input voltage. Similarly it can be seen that at still lower operating frequency, say 0·21 *kc/s*, output voltage E_0 is only 10% of input voltage. In other words only high frequency components are relatively emphasised.

(ii) **Reactance Tube and Modulated Oscillator :** The method is to control the frequency of an oscillator by the modulating voltage with the aid of reactance tube.

We know that the frequency of a tank circuit depends upon the reactance of its components. So if we couple some other reactance which varies with modulating voltage with this tank circuit, then the frequency generated will vary with coupled reactance and hence with modulating voltage. The output wave will be then frequency modulated.

The reactance tube is a pentode which shunts the tank circuit of an ordinary r.f. oscillator. The pentode draws a reactive current that is varied inaccordance with the modulating voltage. This reactive current has an effect equivalent to shunting a reactance across the oscillator tank circuit and so affects the generated frequency.

Refer to circuit 10 (*c*). Resistance R is much smaller than the reactance of C. By means of RC phase splitter, the voltage between grid and cathode, E_{gk}, of reactance tube is made 90° out of phase with E_{pk}, the plate cathode voltage. Consequently, plate current I_{pk} will lead E_{pk} by 90° which happens in the case of capacitive reactance. Thus the action of reactance tube is equivalent to

shunting a capacitive reactance between the plate and ground of oscillator tank circuit. Now in order to obtain the frequency

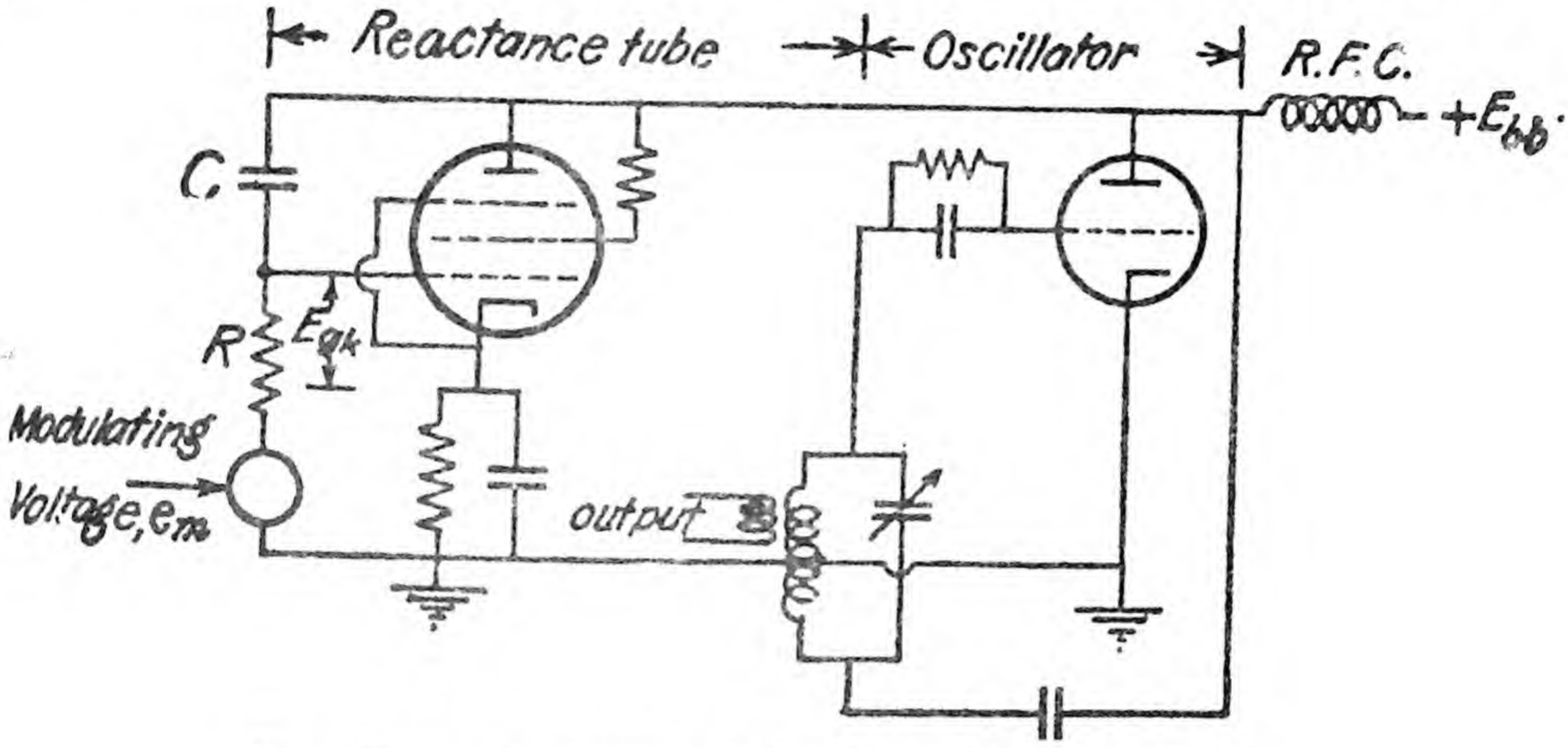


Fig. 10 (c) Reactance tube and modulated oscillator.

modulation, modulating voltage, $e_m = E_m \cos \omega_m t$, superimposed on r.f. voltage, is applied to the control grid of the tube. The oscillator frequency voltage E_{gk} between grid and cathode is given by

$$E_{gk} = \frac{E_{pk}}{R - \frac{j}{\omega C}} R.$$

The resulting plate current is

$$I_{pk} = g_m . F_{gk} = g_m . \frac{E_{pk}}{R - \frac{j}{\omega C}} . R.$$

Since $R \ll \frac{1}{\omega C}$, we get

$$I_{pk} = jR . g_m . \omega C \; E_{pk}$$

Equivalent capacitive reactance is

$$= \frac{E_{pk}}{I_{pk}} = - \frac{j}{\omega . CK \; g_m}$$

or equivalent capacitance is

$$C_{eq} = CR \; g_m .$$

Thus the capacitive reactance varies with g_m, the transconductance of the tube. When a modulating voltage is applied to the control grid of the tube, this will vary the g_m of the tube which, in turn, varies the equivalent capacitive reactance and hence the instantaneous frequency of the oscillator.

(iii) **Control System :** (*a*) *R.C.A. Method :* **Refer to fig. 9.** This system comprises a large number of devices as mixer, crystal oscillator, discriminator, low pass filter, and d.c. amplifier.

The details of each device can be read in their respective chapters. Here we shall explain the principle of operation of the system as a whole.

We require that the transmitter must have correct carrier frequency assigned to it. The control system is employed to correct for any deviation from its assigned frequency value f_i.

The output of a crystal oscillator which generates the stable frequency f_0 is applied to the mixer alongwith the output drawn from frequency-modulated oscillator. The mixer then combines both to give a resultant deviation $f_d=(f_0-f_1)$. This output of mixer is then applied to a phase shift discriminator, the centre frequency of which is f_d. Therefore, if f_i is correct then $f_d=(f_0-f_i)$ will not develop any d.c. output across discriminator. But if transmitter is not operating at the assigned value of frequency, then (f_0-f_i) will have different value than f_d and d.c. component will be developed by the discriminator at the output. This d.c. output with polarity depending upon (f_0-f_i), after suitable amplification, is applied to the control grid of reactance tube in such a way as to modify the frequency of the oscillator to attain the assigned value of frequency. Any variation of modulating frequency is prevented by low pass filter from being transmitted to the reaetance tube.

(*b*) *F. C. C. Method :* Refer to fig. 10 (*d*).

In this method, crystal oscillator operates at the frequency of master oscillator (say 4 *mc/s*). The same subharmonic (say 8 *th. i.e.* 500 *kc/s*) of frequency master oscillator as well as that of crystal oscillator are fed to a balanced phase detector which compares the phases of these two voltages of almost the same frequency. The output d.c, voltage of phase detector is proportional to phase difference of these two subharmonics. If there is no phase difference between them, no d.c. voltage develops. A slight change in the frequency of transmitter changes the phase of its subharmonic, resulting in a phase difference with crystal oscillator frequency subharmonic. Consequently, d.c. voltage is developed at the output of phase detector. This d.c. voltage is fed to the grid of the reactance tube and it changes the master oscillator frequency in such a way as to make it coincide with crystal oscillator frequency.

(*c*) *Bell Laboratory Method :* Refer to fig. 10 (*e*).

In Bell Laboratory Method crystal oscillator operates at a subharmonic of the master oscillator frequency. The output of crystal oscillator and subharmonic of master oscillator are then fed to stabilizing modulator. If frequency drift of master oscillator (carrier frequency) is zero, modulator output is zero but if carrier

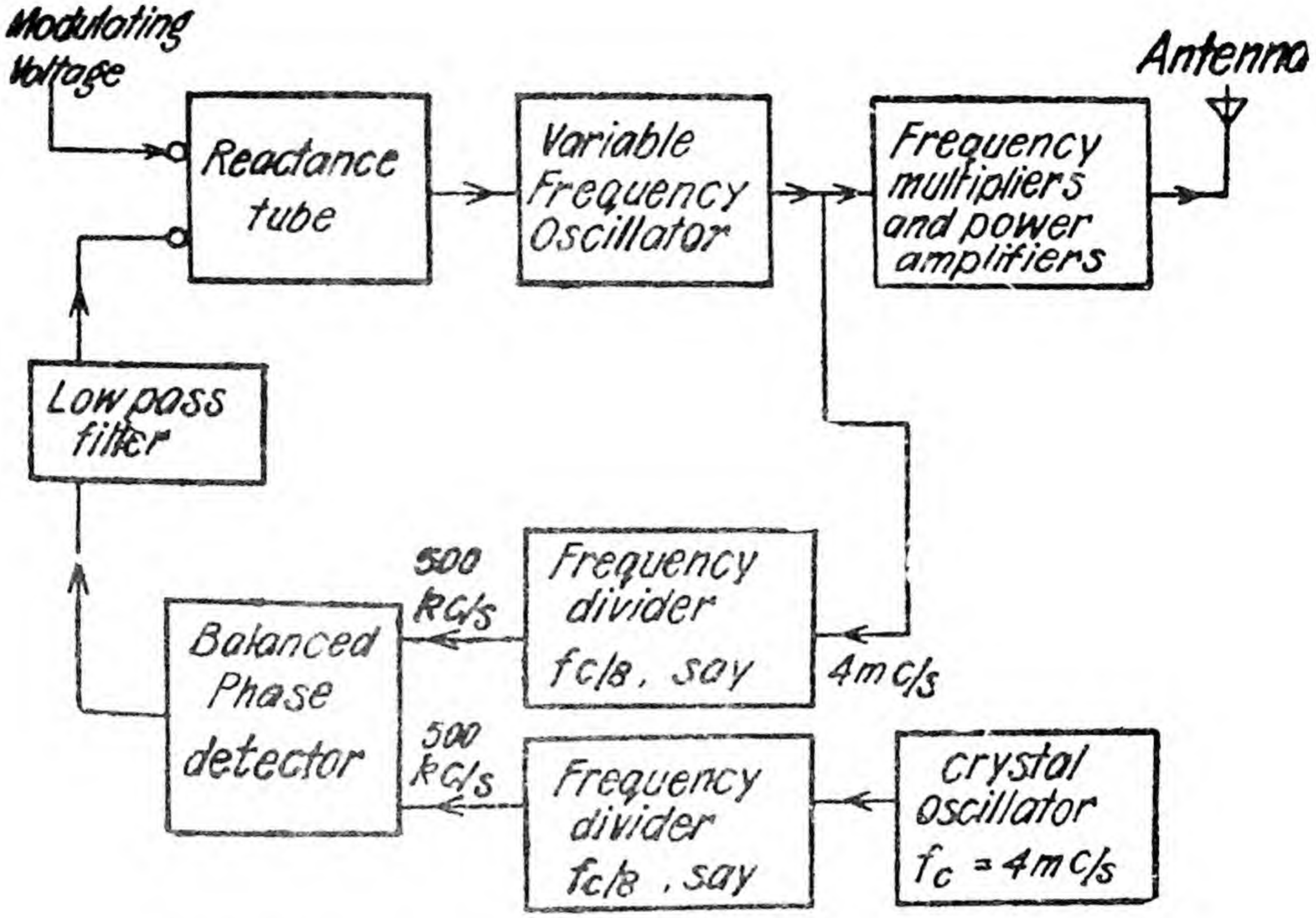


Fig. 10 (d). Reactance tube F.M. transmitter using FCC method for frequency stabilisation.

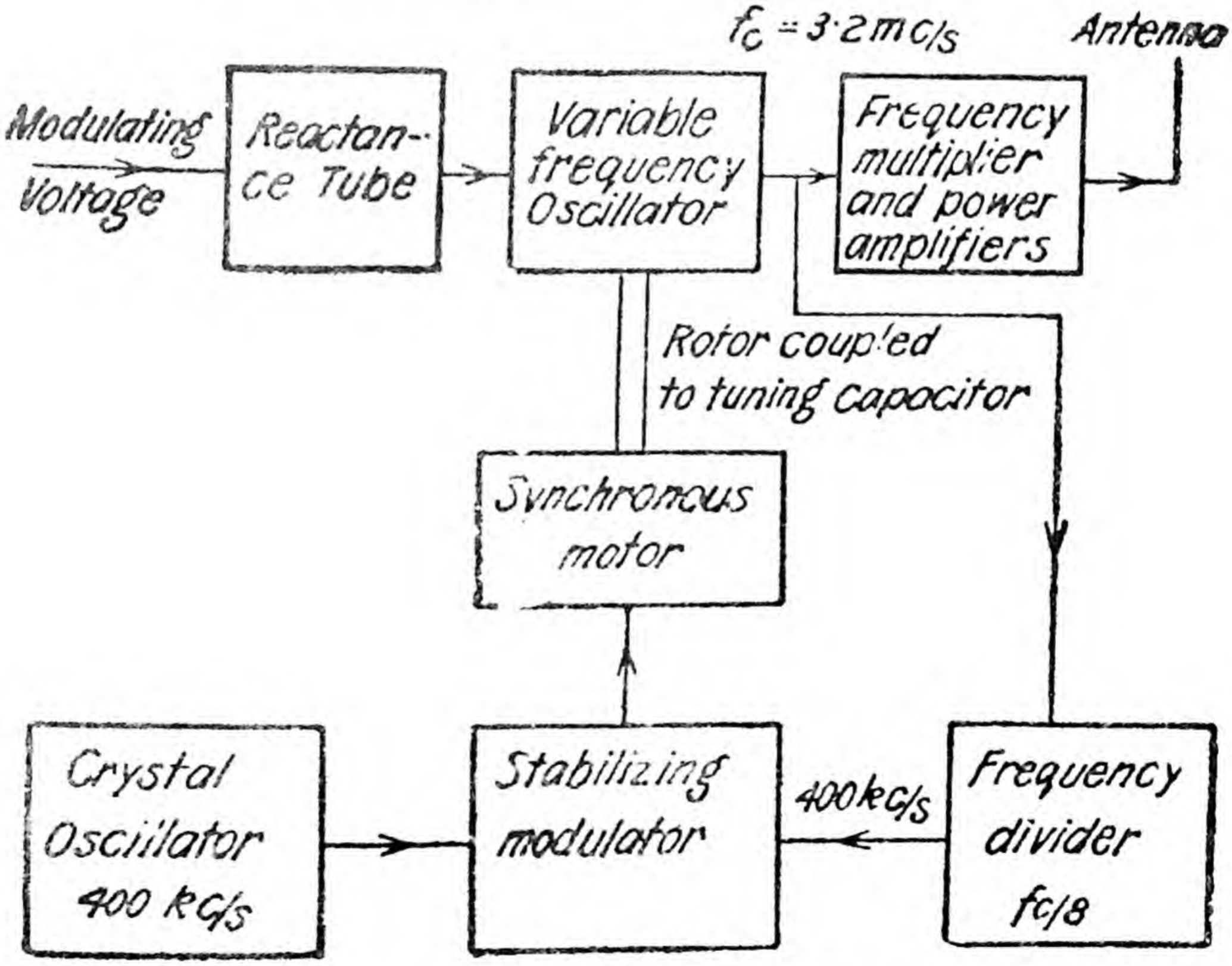


Fig. 10 (e). Reactance tube F.M. Transmitter using Bell Laboratory method for frequency stabilization

frequency is changed (drift from its assigned value) then modulator produces two voltages, relative phase of which depends upon the direction of frequency drift and the amplitude upon the magnitude of frequency drift. The two voltages are amplified and fed to two phase stator windings of an induction motor creating a rotating magnetic field. The sense of rotation depends upon whether the transmitter frequency is above or below the desired value. The rotor which is coupled to the shaft of tuning capacitor of master oscillator will accordingly rotate clockwise or anticlockwise and bring back the transmitter frequency to its assigned value.

32·4. COMPARISON OF F. M. AND A. M. SYSTEMS :

The difference in the two types of systems is obviously due to basic differences between amplitude modulation and frequency modulation as briefed below :

(*i*) *Sidebands* : In contrast to amplitude modulation wherein modulation of carrier gives rise to only one pair of sidebands, frequency modulation produces a large number of pairs of sidebands depending upon deviation ratio and are centered about the carrier frequency.

(*ii*) *Bandwidth and carrier frequency* : The bandwith in *FM* system is much wider on account of a number of sideband pairs *e.g.* with a maximum frequency deviation of 75 *kc/s* a channel bandwidth of 200 *kc/s* is required. On the other hand, in *AM* broadcasting where audio signal band is from 16 to 2000 *c/s*, the bandwidth is only 10 *kc/s*. Consequently, in *FM* system in order to accommodate sufficient number of channels in a given frequency spectrum and to have proper selectivity, carrier frequency are opted between 88 and 108 *mc/s*.

(*iii*) *Power requirements of modulating source* : In *AM* wave intelligence is contained in sidebands. The amplitude and power level of sideband frequencies increases with percent modulation and consequently an increase in modulation level also increases the power content of modulated wave. This extra power is derived from the modulating signal source. In frequency modulating wave, as the amplitude is not affected by the modulation, power content of the wave remains the same. Therefore practically no energy is supplied from the modulating or any other source for modulating the carrier.

Further, since *FM* wave can be amplified by class *C* amplifiers without distortion, their modulation is affected in a lower power stage of *FM* transmitter which requires low modulating signal power. The modulated wave is then amplified by last stage class *C* amplifiers of the transmitter. But in amplitude modulation systems modulation is usually carried out at the last power stage which requires considerable power from the modulating source.

(*iv*) *Distortion* : In amplitude modulation system distortion may arise due to over modulation. Therefore modulating voltage

amplitude should not exceed carrier voltage amplitude. There is no such limitation in *FM* system and a distortion free modulation is obtained. However, in *FM* system frequency deviation Δf should not exceed carrier frequency which is usually the case. A maximum frequency deviation of 75 *kc/s* is arbitrarily fixed so as to have transmission free from noise and interference with adjacent channels.

(*v*) *Fidelity of Reception of Signals* : *FM* systems reduce the disturbance arising due to noise and interference of different signals Thus due to suppression of circuit and tube noise, a clear reception is obtained in case of frequency modulation.

Also in *AM* broadcasting the bandwith of a channel is directly proportional to highest frequency in the modulating signal. At present since a bandwidth of 10 *kc/s* is assigned to each *AM* broadcasting channel, the highest modulating frequency used is only 5 *kc/s*. On the other hand, in case of *FM* system, if deviation ratio is large, bandwidth is practically independent of modulating signal so that all the significant sidebands will lie within 200 *kc/s* bandwidth that is assigned to *FM* system.

(*vi*) *Noise and Interference reduction* : The most important advantage of *FM* over *AM* system lies in the substantial reduction of noise and interference effects. This has been discussed in the next article.

32·5. NOISE AND INTERFERENCE REDUCTION IN FREQUENCY MODULATION SYSTEMS :

The frequency modulation systems are preferred because of their ability to reduce the disturbance arising due to noise and interference of different signals For this purpose they are properly designed and operated in the suitable way.

Suppose the noise voltage E_n is smaller than signal voltage E_s. When E_n is superimposed on E_s, the amplitude of the resultant voltage varies between $(E_s + E_n)$ and $(E_s - E_n)$, while phase of resultant voltage relative to E_s oscillates between $\delta = \pm \sin^{-1} E_n/E_s$ as shown in figure (11). Thus the resultant wave is modulated by the noise both in amplitude and phase. Consequently, at the output of the receiver, the noise will appear. To prevent this noise from receiver appreciably at the receiver output, we adopt the following ways :

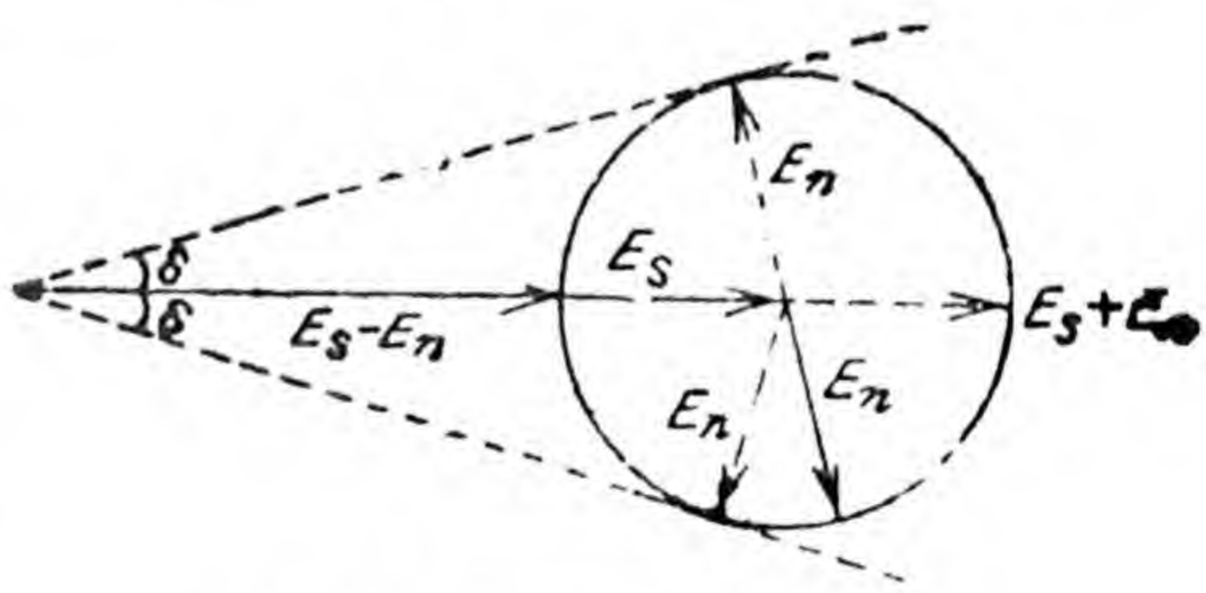


Fig. 11. Superimposition of noise over signal.

(*i*) In the frequency modulated receivers a limiter stage is employed prior to the discriminator stage. Limiter delivers a constant amplitude input to the discriminator and thus amplitude

fluctuations due to noise are rendered ineffective. Instead of limiter discriminator, we can also use ratio detector.

(*ii*) By employing limiter or ratio detector, appearance of noise in the receiver output through amplitude fluctuation is checked but it still can appear through its effect on phase. To check the noise to appear in the output due to the latter effect, a frequency modulated wave with appreciable frequency deviation (Δf) is transmitted so that, for full modulation, modulation index (m_s) due to signal is large compared with unity. For example. we consider a signal of 1 *kc* that produces a frequency deviation of 75 *kc* then modulation index *of signal* will be

$$(m_s)_f = \frac{\Delta f}{f_m} = \frac{75kc}{kc}$$
$$= 75.$$

While with $E_m < E_s$ a noise voltage half as large as the signal has modul tion index

$$(m_n) = \frac{E_n}{E_s} = 0{\cdot}5,$$

so that phase deviation produced by the noise will be only $\frac{0{\cdot}5}{75} = {\cdot}0066$ times as great as the deviation produced by the signal

That means noise weaker than signal will be suppressed completely. It is obvious that greater the frequency deviation Δf, more will be the value of $(m_s)_f$ and hence greater is the extent to which the noise is suppressed. While higher the modulation frequency f_m, less will be the value of $(m_s)_f$ and hence smaller is the extent to which the noise is suppressed In ordinary speech. higher frequencies are relatively weak and so frequency deviation produced by them is small. This causes the reduction in modulation index of the signal and thus noise interference become appreciable at high modulation frequencies. To avoid this noise interference and better reproduction of high modulation frequencies, their amplitude before modulation at the transmitter is increased by employing a pre emphasis network. When noise at some instant during transmission exceeds the signal voltage, the desired signal is almost suppressed.

BROADCAST RECEIVING CIRCUITS

Receiving devices are those which convert the transmitted radio waves into perceptible signals. Based on modulation process, the receiving circuits are of two different types :

(*A*) Amplitude-modulated (a m.) receivers.

(*B*) Frequency modulated (f m.) receivers.

The performance of a radio receiver is determined by means of certain characteristics. Before discussing a.m. and f.m. receivers, we shall first describe some important characteristics of radio receiver on which these receivers are rated.

32·6-1. CHARACTERISTICS OF RECEIVERS :

The important characteristics of receivers are sensitivity, selectivity, fidelity, noise figure, described as follows :

(a) Sensitivity : The sensitivity of radio receiver is that *characteristic which determines the minimum strength of signal input capable of causing a desired value of signal output.* Therefore, expressing in terms of voltage or power, sensitivity can be defined as the minimum voltage or power at the receiver input for causing a standard output.

In case of amplitude-modulation broadcast receivers, the definition of sensitivity has been standardised as "amplitude of carrier voltage modulated 30 per cent at 400 cycles, which when applied to the receiver input terminals through a standard dummy antenna will develop an output of 0 5 watt in a resistance load of appropriate value substituted for the loudspeaker".

(b) Selectivity : The selectivity of a radio receiver is that characteristic which *determines the extent to which it is capable of differentiating between the desired signal and signals of other frequencies.*

(c) Fidelity : This is defined as the *degree with which a system accurately reproduces at its output the essential characteristics of signals which i impressed upon its input.*

(d) Stability : It is a *measure of the ability of a radio receiver to deliver a constant amount of output for a given period of time when the receiver is supplied with a signal of constant amplitude and frequency.* Instability of receiver arises due to variations in output voltage of power supply unit, temperature variations and mechanical imperfections.

(e) Noise figure and signal to noise ratio : Noise figure is a measure of the extent to which the noise appearing in the receiver output in the absence of a signal is greater than the noise that would be present if the receiver were a perfect receiver from the point of view of generating the minimum possible noise. The noise

figure determines the smallest power that may be received without being drowned out by the noise. The noise figure is expressed in decibels or as a power ratio. The larger the noise figure, the noiser the receiver.

Signal and noise can be measured and compared in terms of r.m.s. voltage or current, or in terms of power, and the ratio of signal to noise can be expressed in decibels. For satisfactory reception of the typical broadcast programme, a signal to noise ratio of 40 decibels is necessary.

32·6 2. EXPERIMENTAL DETERMINATION OF RECEIVER CHARACTERISTICS :

A laboratory method for the measurement of receiver characteristics is shown in fig. 12. We use here an artifcial signal to represent the voltage that is induced in the receiving antenna. This artificial signal is applied through 'dummy' antenna, which in

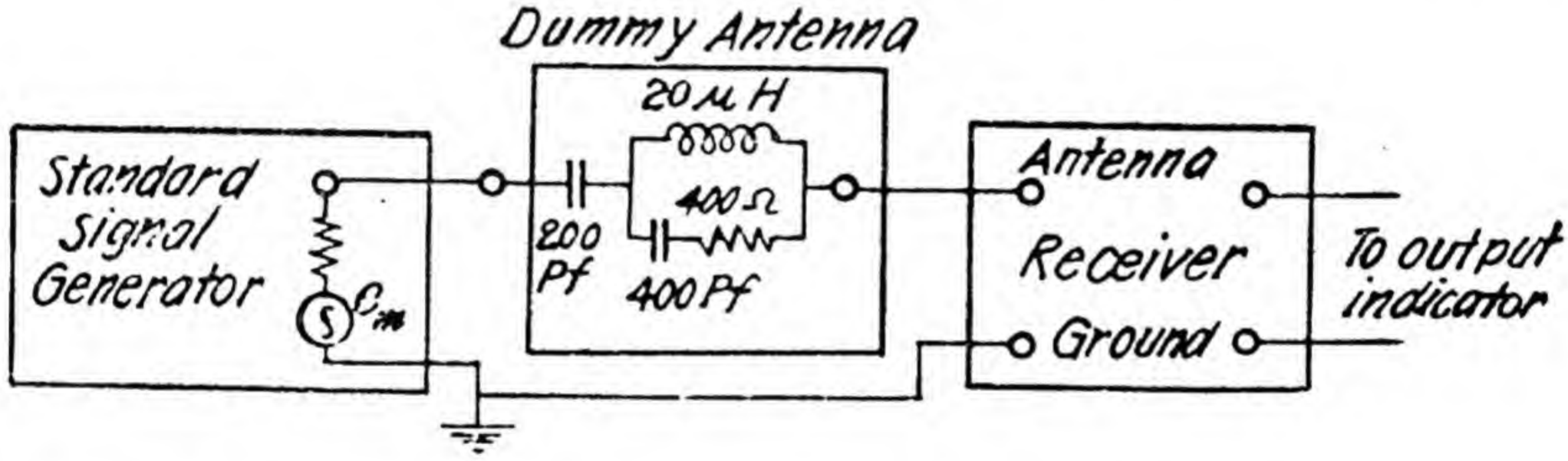


Fig. 12. Arrangement for the measurement of receiver characteristics.

association with internal impedance of the signal generator stimulates the impedance of the actual antenna with which the receiver is to be used. The receiver output is determined by substituting the resistance load of proper value for the loudspeaker and measuring the audio frequency power.

(a) Sensitivity : Sensitivity is determined by impressing different r.f. voltage in series with a standard dummy antenna and adjusting the intensity of input voltage until standard output is obtained at resonance for various carrier frequencies. Sensitivity is expressed in micro-volts. A sensitivity curve is shown in fig. 13.

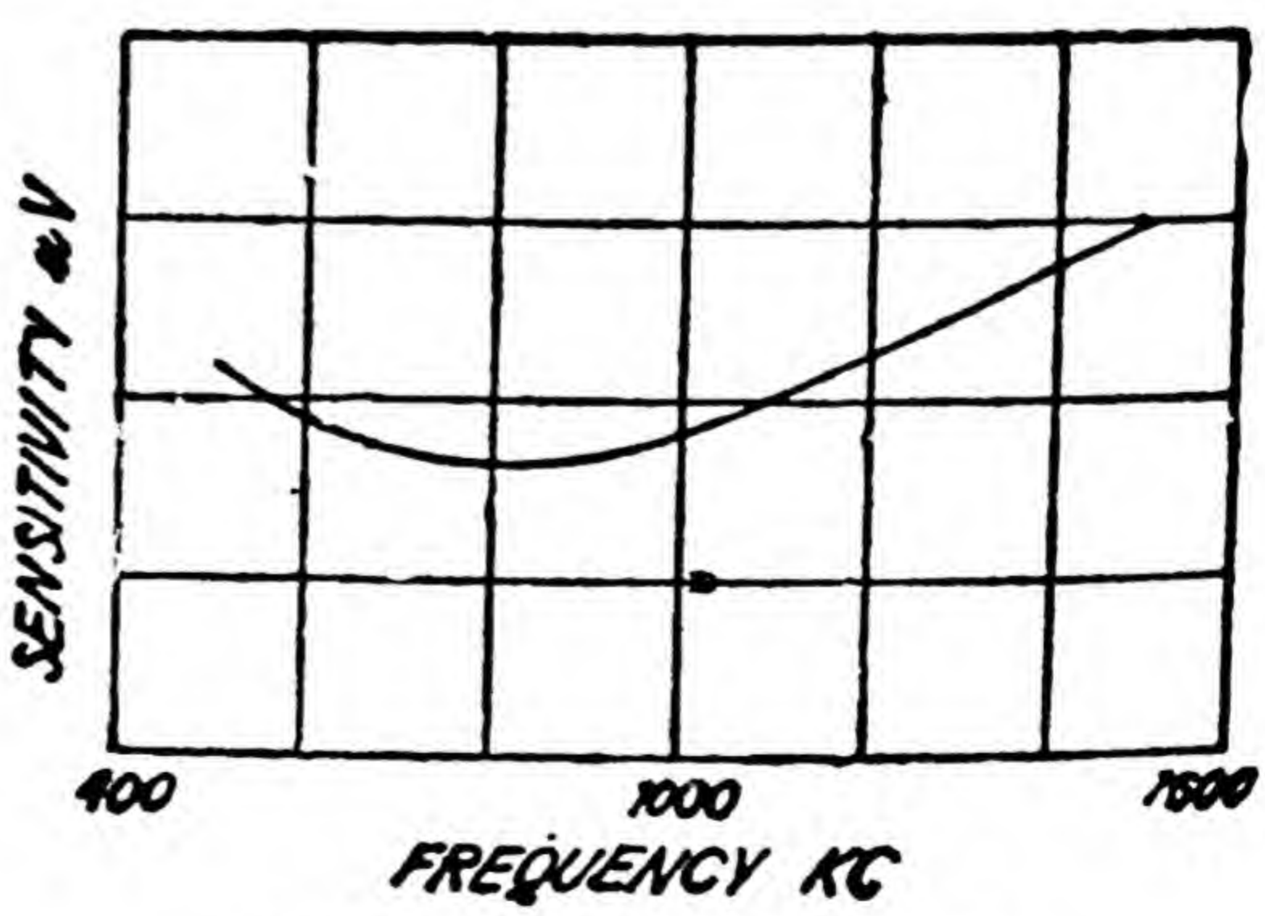


Fig. 13. Sensitivity curve.

(b) Selectivity : Selectivity is expressed in the form of a curve that gives the carrier signal strength with standard modulation that

is required to produce the standard test output plotted as a function of cycles off resonance of the test signal.

The method of obtaining the selectivity curve is as follows :

The receiver is tuned to the desired frequency and mutual volume control is set for maximum value. At standard modulation, the signal generator is set at the resonant frequency of the receiver. The carrier output of the signal generator is varied until the standard test output is obtained. At the same tuning of receiver, the frequency of signal generator is varied above and below the frequency to which the receiver is tuned. For every frequency, the signal generator voltage, applied to the receiver input, is adjusted to give the standard test output from the receiver. The data are plotted in fig. 14.

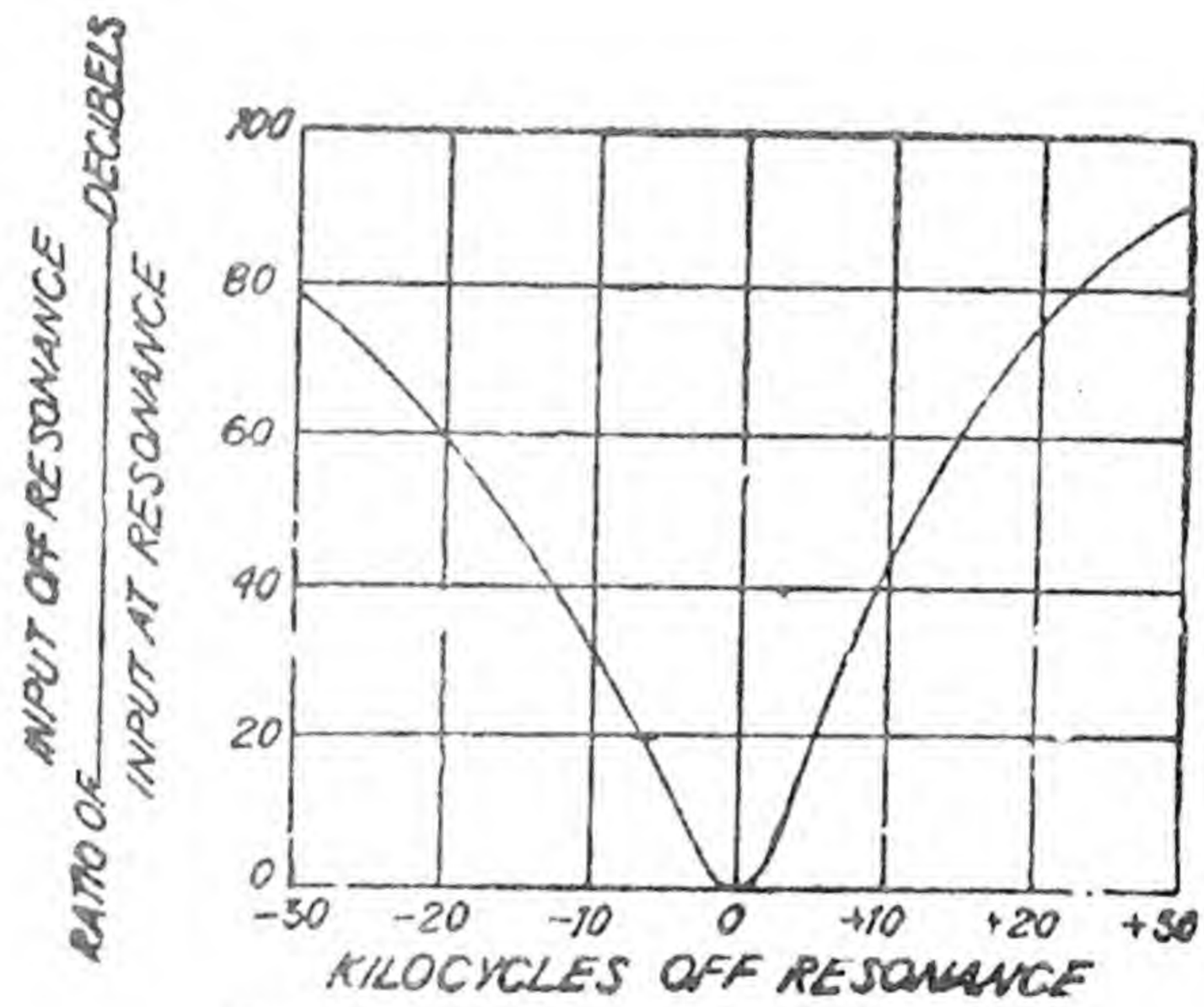


Fig 14. Selectivity curves.

(c) Fidelity : Fidelity is the term expressing the behaviour of receiver output with modulation frequency of input voltage. To

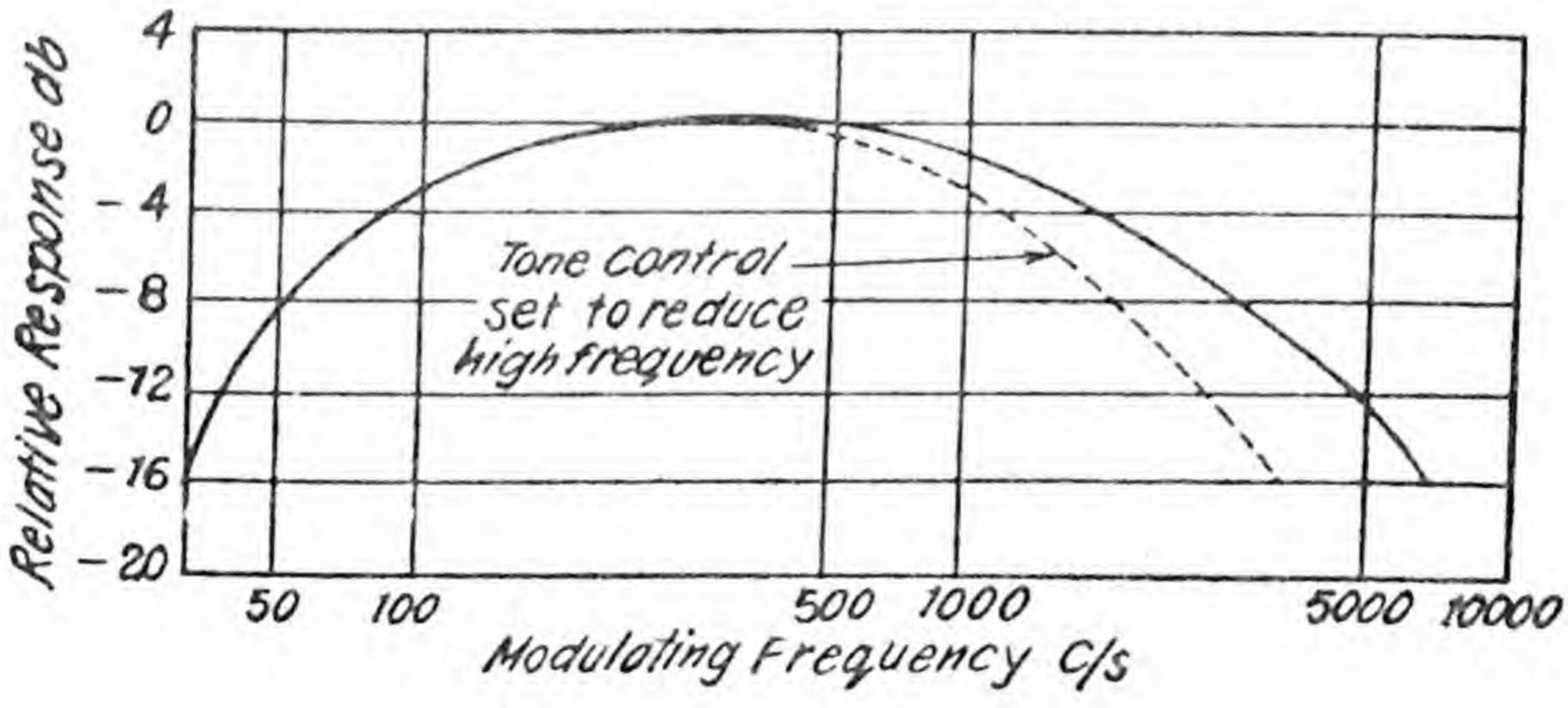


Fig. 15. Fidelity curve.

obtain a fidelity curve, the carrier frequency of the signal generator is adjusted to resonance with the receiver, standard 400 cycles modulation is applied, the signal generator carrier level is set at a convenient arbitrary level and the manual volume control of the receiver is adjusted to give the standard test output. The modulation frequency is then varied over the audio range, keeping degree of modulation constant. A graph is then plotted in the ratio of actual output in volts to the output at 400 c/s against modulation frequency as shown in fig. 15.

(d) Noise Figure : The noise figure of receiver is best measured with a source of a random noise such as temperature limited diode. The noise generator and dummy antenna are connected but the noise generator is not yet tuned on. The noise output power P_1 of i. f. amplifier at this stage is measured. Then tune on the noise generator and increase its output until the noise output of the i. f. amplifier has increased to $2P_1$. Note the space current I of the diode. It can be shown that noise figure of the receiver in decibels is

$$10 \log_{10} (20IR),$$

where R is the resistor through which the noise current flows.

(e) Maximum Undistored Power Output : This is determined by increasing the output in successive steps and measuring both the power output and the percentage of harmonics. The maximum undistorted output is the least power output which contains a total r.m.s. harmonic voltage content of 10 percent of fundamental voltage.

(f) Signal to noise ratio : Noise figure does not give signal to noise ratio of the receiver output after detection and audio amplification but it simply represents a merit of the actual receiver as compared to an ideal receiver having minimum possible noise. Signal to noise ratio is determined by equivalent noise band input (e.n.s.i.) test.

The procedure of measuring e.n.s.i. is as follows :

(*i*) Select a convenient carrier level of about 100 db, ($10\mu V$) and say it $(E_s)_{db}$.

(*ii*) With this carrier applied without modulation, measure the r.m.s. noise voltage delivered by the receiver and say this E_n'.

(*iii*) Switch on 400 cycles 30 percent modulation and inserting 400 cycle band pass filter in the output circuit of the receiver (which will remove all noise practically), measure the r.m.s 400 cycles output, say this E_s'.

The e.n.s.i. can be computed from the following formula :

$$(E_n)_{db} = (E_s)_{db} + \left(\frac{E_s'}{E_n'}\right)_{db} + 10.$$

In general the signal to noise voltage ratio at the receiver out-

put can be expressed in terms of E_n according to the relation

$$\text{signal to noise voltage ratio} = \frac{m_a E_c}{E_n},$$

where E_c and m_a a e the carrier amplitude and degree of modulation respectively of the signal applied to the receiver input for which the signal to noise ratio is desired. The relation for e.n.s.i. considering general case can be put in the form

$$E_n = m_a E_s \left(\frac{E_n'}{E_s'}\right) \text{ in } \mu V,$$

E_n, E_s being in microvolts while E_n' and E_s' in volts.

(g) Hum measurement : Alternating power frequen y currents produce hum which can be said as a tone of 10 cycles fundamental. Hum originating in a.f. circuits is called "residual hum" and is measured with no r.f. input to the detector. Hum produced by modulation o the carrier in the receiver is called "modulation hum".

The hum voltages present in the receiver output can be observed by employing square law measuring instrument. The result will usually depend on the setting of manual volume control.

32·7. FREQUENCY CHANGING OR CONVERSION :

In the reception of radio signals, better response is obtained if frequency of the incoming signal is changed to some lower radio frequency value (without affecting modulation envelope). This lower value of frequency is termed as intermediate frequency and process of changing the frequency is called heterodyne action or mixing or beating. There are two types of ways of achieving this conversion.

32·7-1. SINGLE ELECTRODE MIXING :

In this process, signal voltage as well as locally generated oscillator voltage are applied at the same grid of mixer tube as shown in fig. 16. When the two voltages are applied to the same grid, mixing action is achieved through utilizing the square law plate current-grid voltage property of the tube. In this process of mixing, the important property is conversion transconductance defined by

$$g_{cn} = \frac{I_{if}}{E_{rf}},$$

where I_{if} is the plate current at intermediate frequency and E_{rf} is the signal voltage at the signal frequency. An expression for it can be deduced as follows :

Square law conversion theory :

Let the plate current grid voltage characteristic of the detector follow square law so that a.c. plate current is given by

$$i_p = a_1 e_g + a_2 e_g^2.$$

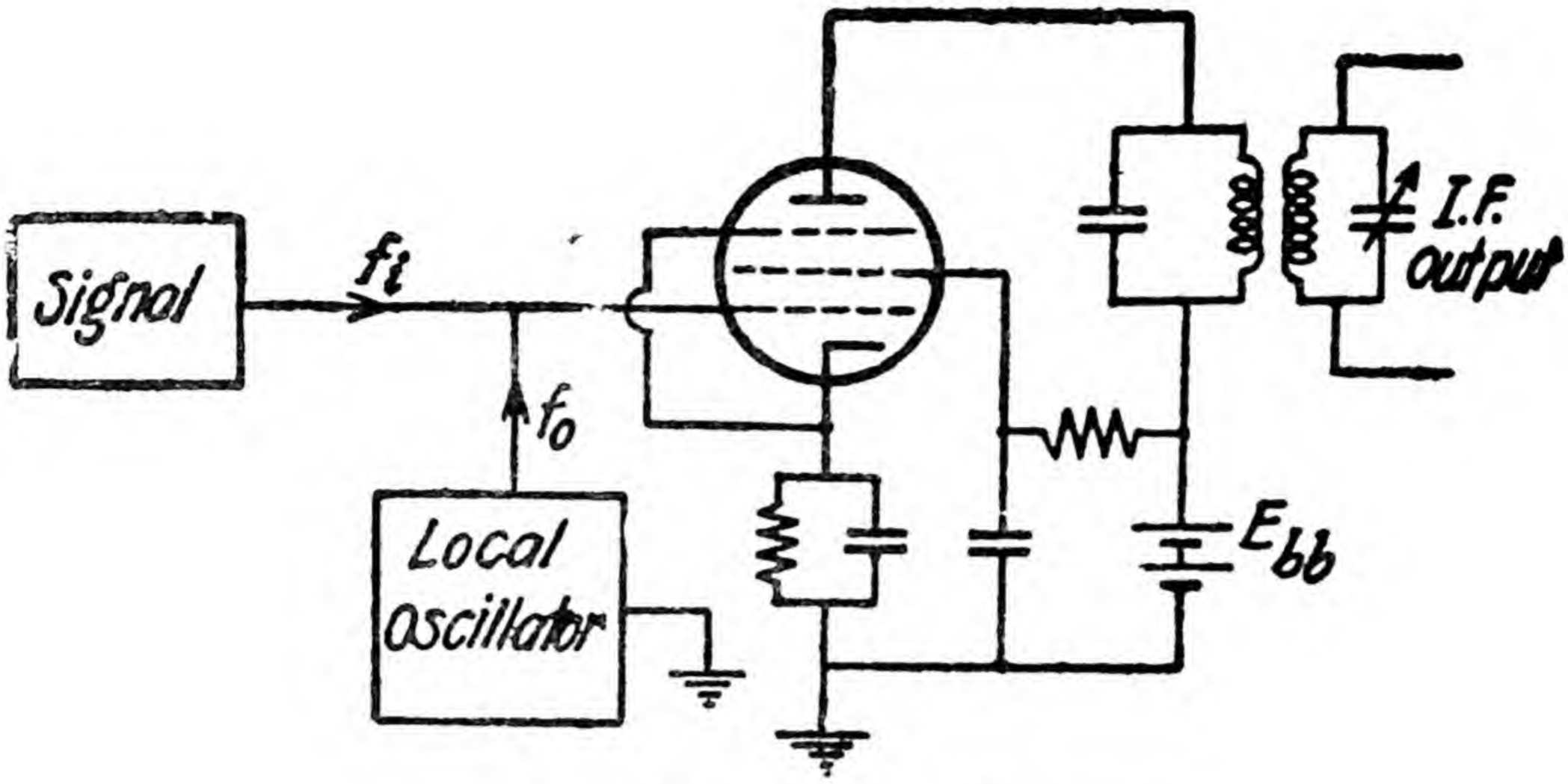


Fig 16. Single electrode mixing.

where a_1 and a_2 are constants and e_g is the total a.c. voltage applied at the grid and is given by

$$e_g = E_c\ (1 + m_a \cos \omega_m t) \cos \omega_c t + E_0 \cos \omega_0 t$$

where first term on right hand side refers to modulated signal input and second term with subscript 0 refers to the local oscillator voltage. Putting e_g in expression for i_p, we get

$$i_p = a_1\ [E_c\ (1 + m_a \cos \omega_m t) \cos \omega_c t + E_0 \cos \omega_0 t] + a_2\ [E_c\ (1 + m_a \cos \omega_m t) \cos \omega_c t + E_0 \cos \omega_0 t]^2$$

Terms with coefficient a_1 are of no use in mixing action as they refer to the linear region of characteristic, we can write

$$\begin{aligned} i_p &= a_2\ [E_c\ (1 + m_a \cos \omega_m t) \cos \omega_c t + E_0 \cos \omega_0 t]^2 \\ &= a_2\ [E_c^2 \cos^2 \omega_c t + E_0^2 \cos^2 \omega_0 t + E_c^2\, m_a^2 \cos^2 \omega_m t \cos^2 \omega_c t \\ &\quad + 2E_c^2 m_a \cos \omega_m t \cos^2 \omega_c t \\ &\quad + 2E_c\ (1 + m_a \cos \omega_m t) \cos \omega_c t\ E_0 \cos \omega_0 t] \\ &= a_2\ [E_c^2 \cos^2 \omega_c t + E_0^2 \cos^2 \omega_0 t \\ &\quad + \frac{m_a^2\, E_c^2}{4} \cos^2 (\omega_c + \omega_m)\, t \\ &\quad + \frac{m_a^2\, E_c^2}{4} \cos^2 (\omega_c - \omega_m)\, t \\ &\quad + m_a\, E_c^2 \cos \omega_c t \cos (\omega_c + \omega_m)\, t \\ &\quad + m_a\, E_c^2 \cos \omega_c t \cos (\omega_c - \omega_m)\, t \\ &\quad + 2E_c\, E_0 \cos \omega_c t \cos \omega_0 t \\ &\quad + m_a\, E_c E_0 \cos \omega_0 t \cos (\omega_c + \omega_m)\, t \\ &\quad + m_a\, E_c E_0 \cos \omega_0 t \cos (\omega_c - \omega_m)\, t] \end{aligned}$$

The resonant circuit in the plate of the mixer tube is tuned to the difference frequency $(\omega_0 - \omega_c)$, called intermediate frequency and it rejects all other frequency terms except those that are very close to $(\omega_0 - \omega_c)$. Thus neglecting all terms in above expression except those close to $(\omega_0 - \omega_c)$, we get

$$i_p = a_2 [E_c E_0 \cos (\omega_0 - \omega_c) t + \tfrac{1}{2} m_a E_c E_0 \cos (\omega_0 - \omega_c - \omega_m) t + \tfrac{1}{2} m_a E_c E_0 \cos (\omega_0 - \omega_c + \omega_m) t]$$

Putting $\omega_0 - \omega_c = \omega_i$, the intermediate frequency, we get

$$i_p = a_2 [E_c E_0 \cos \omega_i t + \tfrac{1}{2} m_a E_c E_0 \cos (\omega_i - \omega_m) t + \tfrac{1}{2} m_a E_c E_0 \cos (\omega_i + \omega_m) t]$$

$$= a_2 E_c E_0 [1 + m_a \cos \omega_m t] \cos \omega_i t$$

$$= I_{if} \cos \omega_i t \qquad \ldots(1)$$

where I_{if} is the amplitude of intermediate frequency component. The tuned circuit is adjusted to have a bandwidth of $2\omega_m$ centred about the intermediate frequency ω_i. Then conversion transconductance is given by

$$g_{cn} = \frac{I_{if}}{E_c (1 + m_a \cos \omega_m t)} = a_2 E_0 \qquad \ldots(2)$$

If the tuned circuit primary has impedance R_t at resonant frequency ω_i, then output voltage across tuned circuit is

$$E_{if} = I_{if} R_t$$

$$= a_2 E_c E_0 (1 + m_a \cos \omega_m t) R_t$$

and the instantaneous value is given by

$$e_{if} = E_{if} \cos \omega_i t$$

$$= a_2 E_c E_0 (1 + m_a \cos \omega_m t) R_t \cos \omega_i t. \qquad \ldots(3)$$

Obviously, as represented by eq. 3, carrier frequency is changed from ω_c to ω_i but the modulation characteristics remain the same, (same modulation index and modulation frequency). However, it may be noted that as amplitude of output wave, E_{if}, now includes E_c as well as E_0, square law heterodyne action provides an inherent amplification.

32·7·2. TWO ELECTRODE MIXING :

In this process signal voltage and the oscillator voltage are fed to two separate electrodes of the tube and mixing action takes place

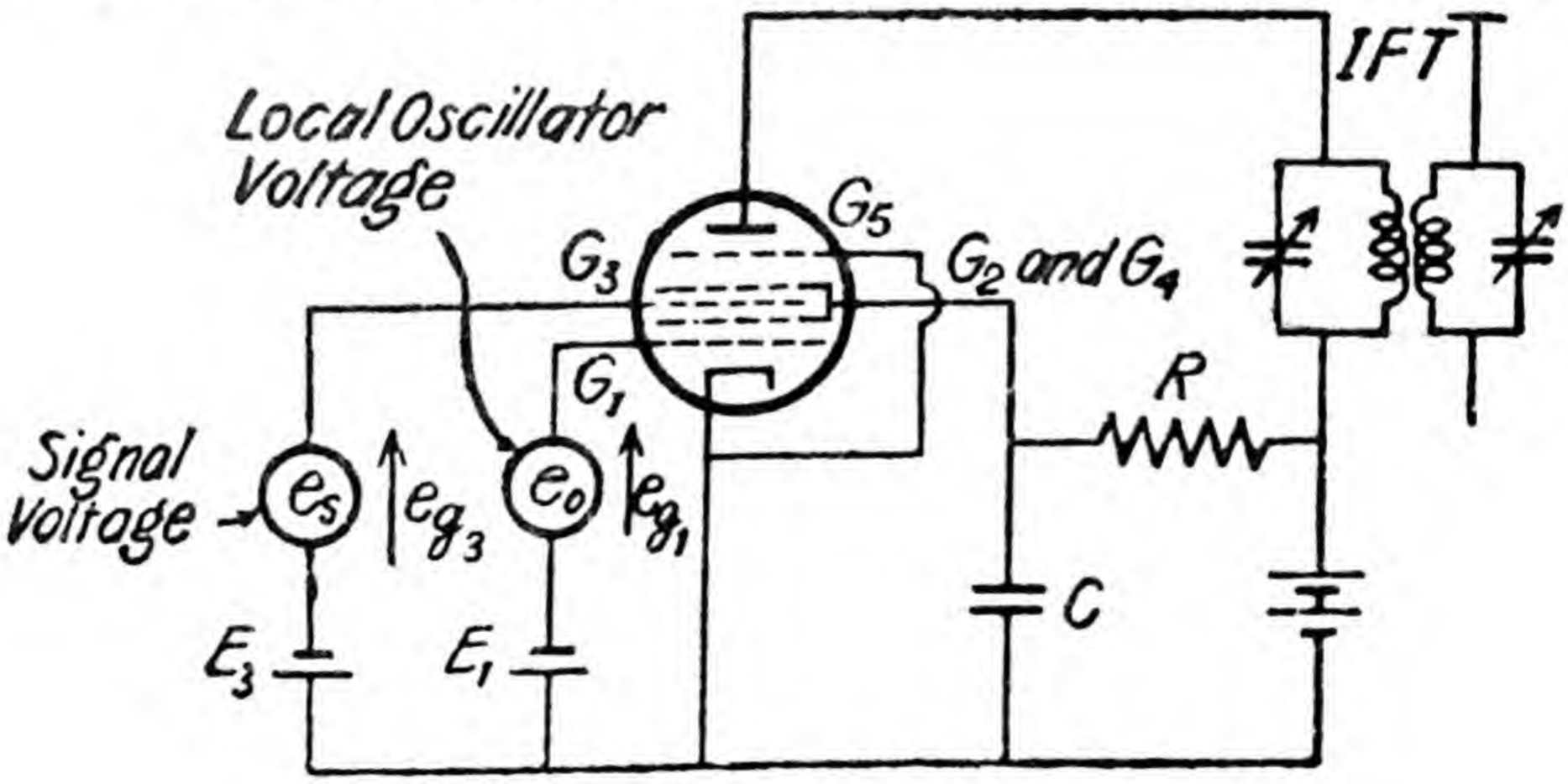


Fig. 17. A Pentagrid Converter.

through electron coupling *i.e.* coupling through electron stream. Pentagrid converter, triode-hexode converter etc. are the examples of this type of mixing. A pentagrid converter is shown in fig. 17. As in pentode, grids G_1 and G_2 control the total cathode (space) current i_c. Grid G_3 is completely shielded from the cathode so that its influence upon the cathode current is practically zero. A part of electron stream from cathode stops at grid G_2 but most of it reaches grid G_3 through G_2. The potential of grid G_3 controls the passage of the fraction of this electron stream towards plate and the fraction returned back to grid G_2, as is evident from the curves in fig. 18 which depict plate current i_p as function of eg_3 for various values of eg_1. Since for all curves eg_2 is constant and eg_1 is constant

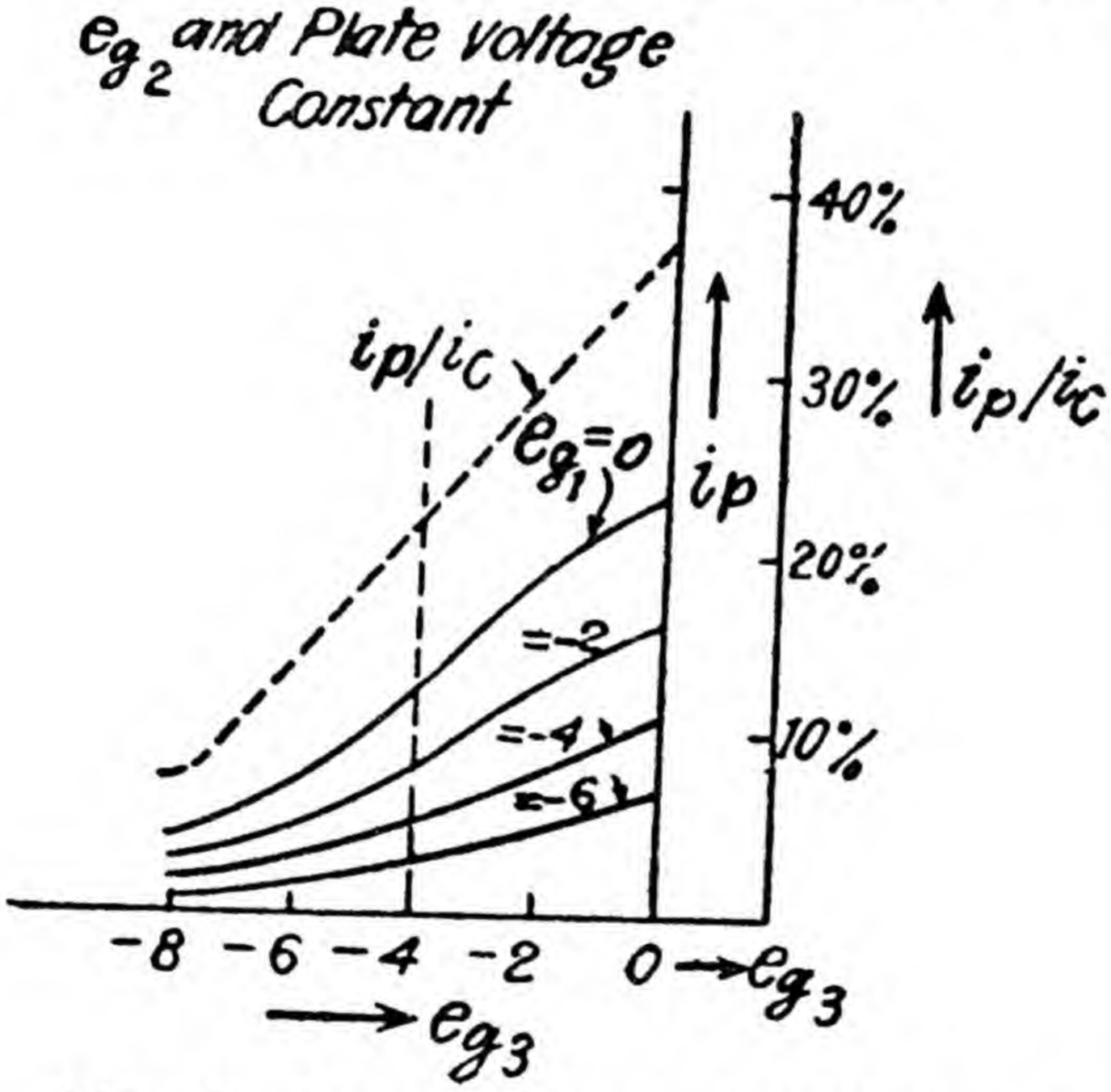


Fig 18. Characteristics of Pentagrid Tube *6SA7*

for a single curve, cathode current i_c, is constant for each curve. If i_p/i_c is plotted as a function of eg_3 then for all values of eg_1, a single line (dashed in figure) is obtained which shows that i_p/i_c is a function of eg_3 and independent of eg_1. Mathematically,

$$\frac{i_p}{i_c} = a_1 + b_1\, eg_3 \qquad \ldots(1)$$

where a_1 and b_1 are constants.

Suppose i_c–eg_1 characteristics, for constant screen grid voltage, are straight lines over a considerable range of values. Then we can write

$$i_c = a_0 + b_0\, eg_1 \qquad \ldots(2)$$

where a_0 and b_0 are constants.

From eqs. (1) and (2), we get

$$i_p = (a_1 + b_1 eg_3)\,(a_0 + b_0 eg_1).$$

Further from fig. 17,

$$eg_1 = -E_1 + e_0 = -E_1 + E_0 \cos \omega_0 t$$

and $$eg_3 = -E_3 + e_3 = -E_3 + E_c (1 + m_a \cos \omega_m t) \cos \omega_c t$$
$$= -E_3 + E_c' \cos \omega_c t.$$

Therefore

$$i_p = [a_1 + b_1 (-E_3 + E_c' \cos \omega_c t)] [a_0 + b_0 (-E_1 + E_0 \cos \omega_0 t)]$$
$$= [A_1 + b_1 E_c' \cos \omega_0 t] [A_0 + b_0 E_0 \cos \omega_0 t].$$

where $A_1 = a_1 - b_1 E_3$ and $A_0 = a_0 - b_0 E_1$.

The above expression can be further simplified to

$$i_p = A_0 A_1 + b_1 A_0 E_c' \cos \omega_0 t + b_0 A_1 E_0 \cos \omega_0 t$$
$$+ b_1 b_0 E_c' E_0 \cos \omega_0 t \cos \omega_0 t$$
$$= A_0 A_1 + b_1 A_0 E_c' \cos \omega_c t + b_0 A_1 E_0 \cos \omega_0 t$$
$$+ \tfrac{1}{2} b_0 b_1 E_0 E_c' \cos (\omega_0 - \omega_c) t$$
$$+ \tfrac{1}{2} b_0 b_1 E_0 E_c' \cos (\omega_0 + \omega_c) t.$$

Therefore, plate current consists of frequencies ω_c, ω_0, $(\omega_0 - \omega_c)$, and $(\omega_0 + \omega_c)$. Since resonant circuit in the plate circuit is tuned to the difference frequency $(\omega_0 - \omega_c)$, currents of all other frequencies will be rejected by the circuit. Therefore the term of interest will be

$$i_p = \tfrac{1}{2} b_0 b_1 E_0 E_c' \cos (\omega_0 - \omega_c) t$$
$$= I_{if} \cos \omega_i t \qquad \ldots(3)$$

which is same as eq. (2) of art. 32·7-1 obtained for single electrode mixing. The conversion transconductance in this case will be

$$g_{cn} = \frac{\text{change in plate current at difference frequency}}{\text{change in signal grid voltage}},$$

for constant plate voltage.

If signal voltage changes from zero to E_c', the change in plate current at difference frequency is I_{if}. Therefore conversion transconductance is

$$g_{cn} = \frac{I_{if}}{E_c'}$$
$$= \frac{\tfrac{1}{2} b_0 b_1 E_0 E_c'}{E_c'}$$
$$= \tfrac{1}{2} b_0 b_1 E_0 \qquad \ldots(4)$$

which predicts that conversion transconductance is proportional to oscillator grid voltage and increases linearly with it. However, there is a definite upper limit to g_{cn}. As E_0 is increased, a point reaches when the negative swing of voltage draws the plate current to cut off so that any further increase in E_0 is inoperative in increasing g_{cn}. Moreover large positive oscillator voltage also gives rise to damping to the source of oscillation.

32·8. AMPLITUDE MODULATED RECEIVERS :

(A) TUNED RADIO FREQUENCY RECEIVERS

(Straight receiver) :

In r.f. receiver, selectivity and r.f. amplification are due to the use of selective circuits and tubes functioning at the frequency of incoming signal. The t. r. f. receiver uses two or more tuned circuits which are usually adjusted simultaneously by a single gang capacitor tuning control.

The circuit diagram of a fi e tube a.c. operated t. r. f. receiver is shown in figure 19. The circuit uses two stages of tuned r. f. amplification, a biased detector, one stage of a.f. voltage amplication and one stage of a.f. power amplification. It is designed to provide :

(*a*) sufficient selectivity for the reception of all local stations.

(*b*) sufficient volume.

(*c*) good fidelity,

(*d*) high signal to noise ratio.

(*a*) *A. F. Amplication Stage* : To obtain sufficient signal strength one or more a f. amplification stages are employed. The a.f amplification stage may be either of the transformer coupled or RC coupled type.

(*b*) *R. F. Amplification Stage* : Selectivity of receiver increases by the use of r.f. amplification stages. For every stage, additional tube is required. At each stage, amplification and tuning takes place. If too many tuned stages are added, the circuit may become too selective and thus there will be an adverse effect on the fidelity of the receiver.

(*c*) *Volume Control* : The strength of the signal in the primary of the antenna coil T_1 is varied by varying resistance R_1. The action can be understood as : when x moves towards A, a greater portion of signal input is shunted to ground and thus signal strength is decreased in the primary of antenna coil T_1. The second effect is that the cathode bias resistance of 6SK7 is increased. Both these actions will decrease the gain of 6SK7 and thus volume is decreased. A reverse effect will be produced, *i.e.*, volume increases when x is moved towards B.

(*d*) *Tone Control* : Potentiometer R_{10} and capacitor C_{14} act as tone control and provide a path to ground for a portion of a. f. current. The value of C_{14} is such as to ground only higher value of audio frequency. The percentage of a. f. current shunted to ground depends upon the value of R_{10}. When R_{10} is low, then only low and medium audio frequencies will pass to output transformer.

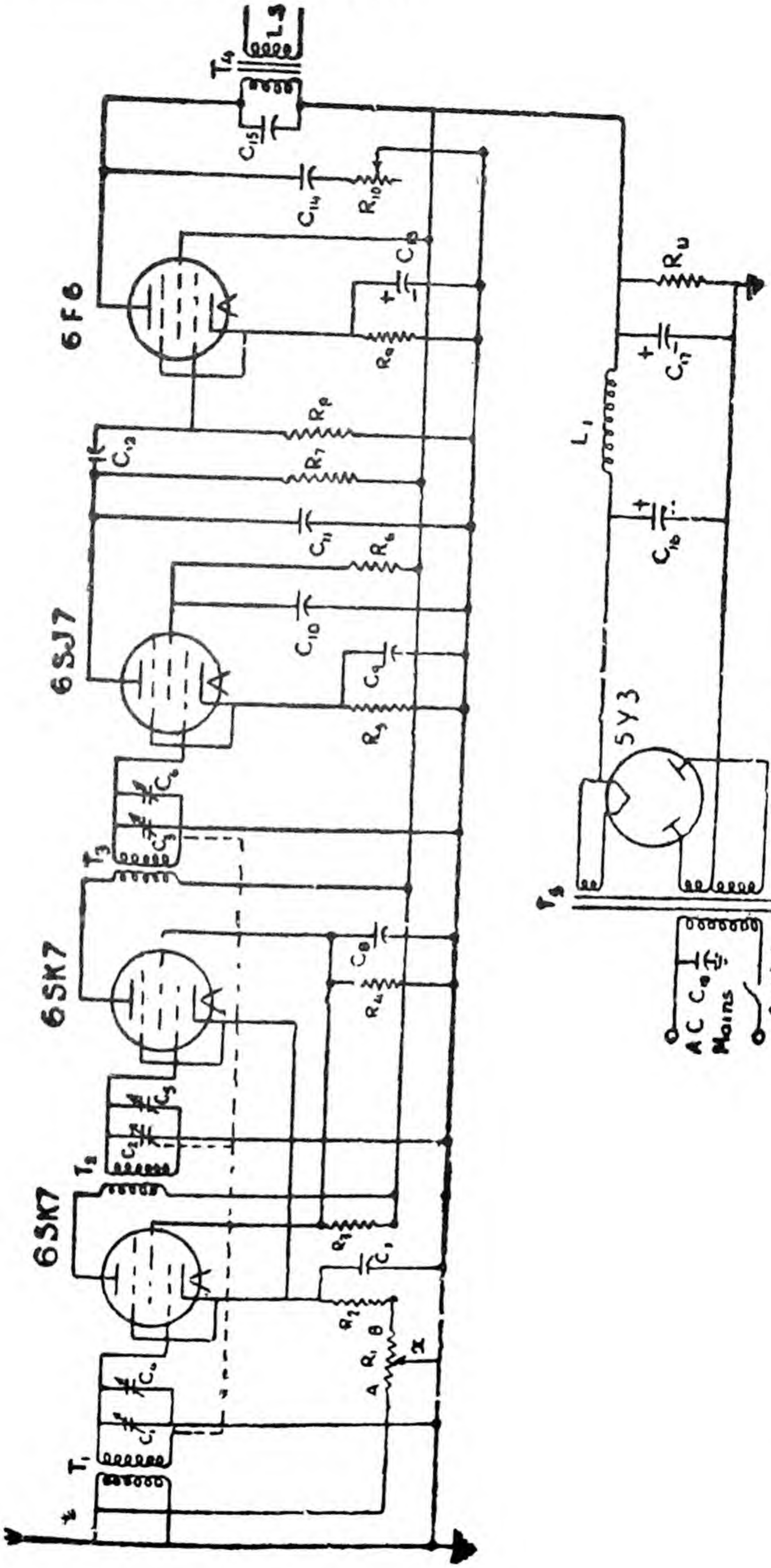


Fig. 19. Circuit diagram of a five tube a.c. operated t r.f receiver.

Values of circuit components : In the following table, values of circuit elements (resistance and capacitance) are given :

Resistors	Capacitors	Other Elements
R_1=10 K ohms.	$C_1=C_2=C_3$ =10 to 3.5 p.f.	T_1=Antenna coil
R_2=250 ohms.	$C_4=C_5=C_6$ =4 to 30 p.f.	T_2, T_3 }=r.f. transformer
R_3=50 K ohms.	$C_7=C_8=C_9=C_{10}$ =0·5μF	T_4=output transformer T_5=power transformer
R_4=6: K ohms.	C_{11}=2·0 p.f.	$L.S.$=Loudspeaker
R_5=33 K ohms.	C_{12}=0·05 μF	L_1=15 H.
R_6=1 megaohms.	C_{13}=0·5 μF	
R_7=220 K ohms.	C_{14}=0·1 μF	
R_8=470 K ohms.	C_{15}=0·01 μF	
R_9=400 ohms.	$C_{16}=C_{17}$=200 μF	
R_{10}=75 K ohms.		
R_{11}=Bleeder		

(B) SUPERHETERODYNE RECEIVER :

Principle : In the process of superheterodyne reception, the signal voltage is combined with a local oscillator voltage and is converted into a new signal with the *same modulation*, but of lower carrier frequency called the intermediate frequency, which is then amplified and detected to produce the original modulation.

Advantages : Advantages of superhetrodyne receiver over tuned radio frequency receiver are :

(*i*) improved selectivity interms of stations on adjacent channels,

(*ii*) quite stable operation,

(*iii*) higher stage gain due to amplification being at lower frequency value of the i.f. stages, and

(*iv*) selectivity remains uniform over the broadcast band.

Since the characteristics of intermediate frequency amplifier, which provides the major portion of the amplification and selectivity of receiver, are independent of the frequency to which the

receiver is tuned, the selectivity of a superhetrodyne is fairly receiver uniform throughout its tuning range.

Block Diagram ; Block diagram of a superheterodyne receiver is shown in fig. 20.

The superheterodyne receiver has the essential components of a t.r.f. receiver, with a frequency converter and intermediate frequency amplifier in addition. The frequency converter consists of an oscillator and a modulator. By means of a gang capacitor, oscillator and t.r.f. circuits are tuned simultaneouly.

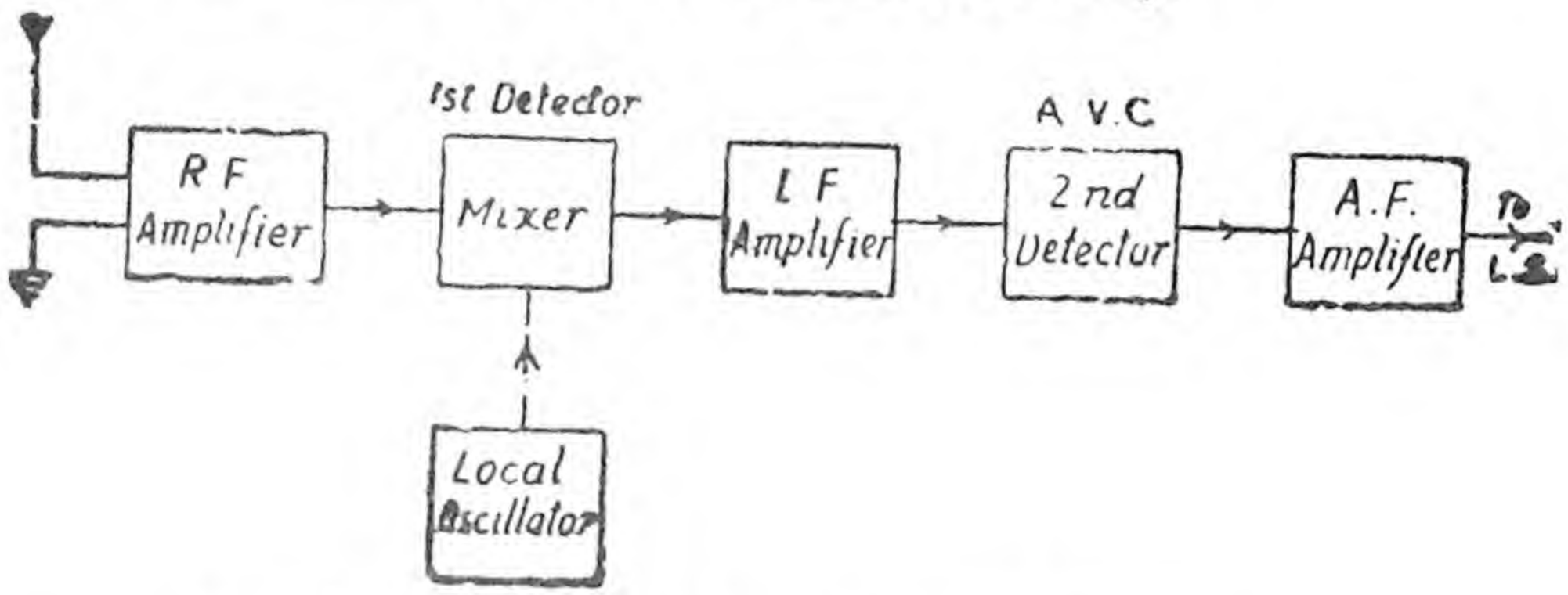


Fig. 20. Block diagram superheterodyne receiver.

Operation in brief : The incoming r. f. signal f_s is mixed with slightly lower or higher frequency from the local oscillator f_0. When mixed, the sum (f_0+f_s) and difference (f_0-f_s) frequencies are formed. By suitable arrangement only difference frequency is extracted. The difference frequency (f_0-f_s), known as intermediate frequency, is then further amplified by the I. F. amplifier. The output of this stage is then applied to second detector where rectification and a.f. amplification takes place simultaneously. Audio frequency output is again amplified by the next stage and then applied to a loudspeaker.

Description of Individual Stages : A complete circuit diagram is shown in figure 2 . In previous chapters of demodulation, amplification etc., we have discussed fully the machanism of these operations. In this article, therefore, we shall discuss converter stage in detail while simply mention other stages.

(*a*) *R.F. Amplifier* : The r. f. circuits are used primarily for eliminating certain forms of interference common to this receiver. This is known as preselection stage. R. F. amplifiers are generally operated with untuned primaries and tuned secondaries and employ pentode tubes. This stage leads to :

(*i*) better selectivity,

(*ii*) improved signal to noise ratio, and

(*iii*) good image frequency suppression.

In circuit diagram, pentode tube 6K7 is used for this stage. Combination of resistor R_1 and capacitor C_1 provides cathode biassing voltage. R_{14} is the resistance to give suitable screen grid

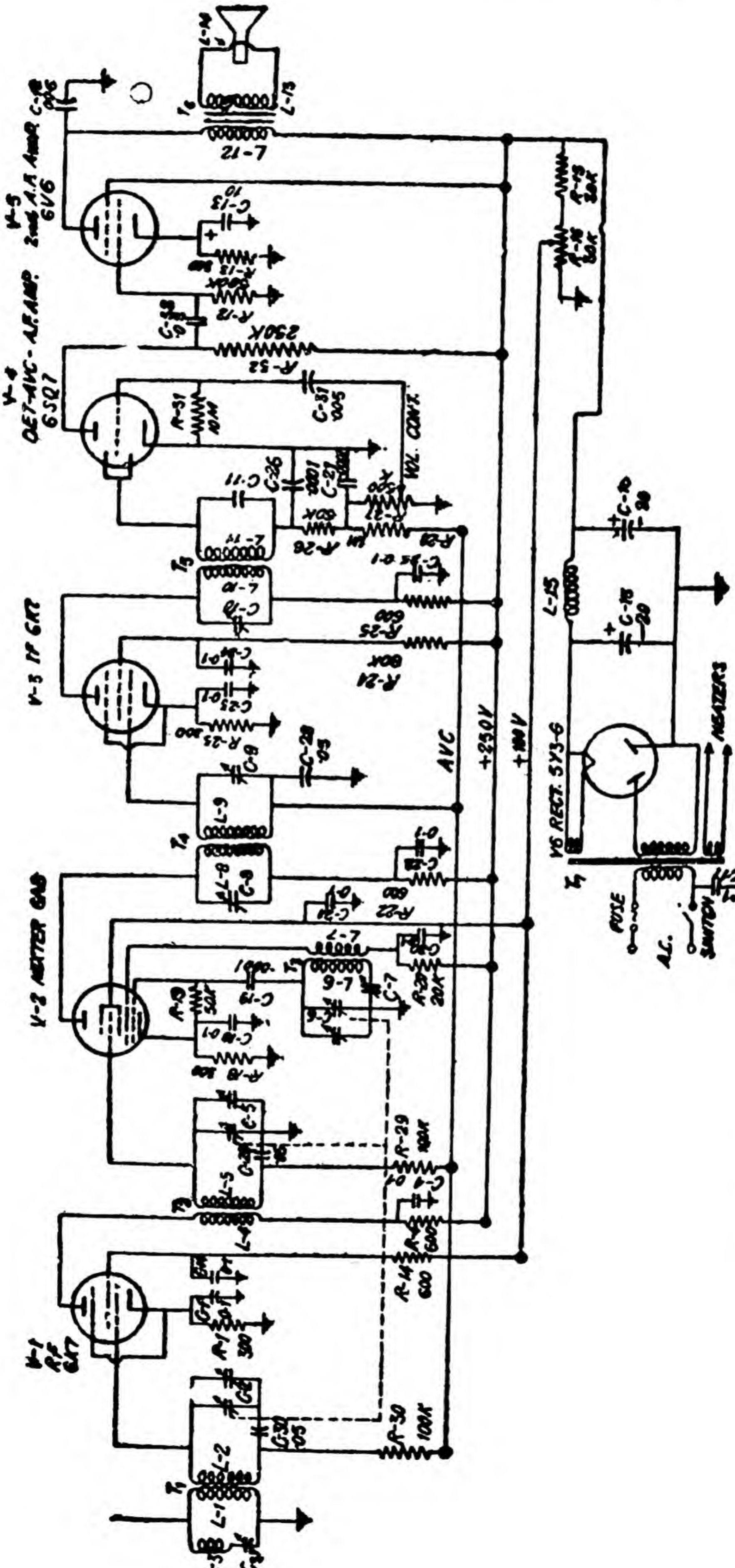


Fig. 21. Circuit diagram of 6 tube a.c. operated superheterodyne receiver.

voltage. The output of this stage is coupled to converter stage by means of a transformer.

(*b*) *Pentagrid Converter* : Mixing operation of this converter has been discussed already in art. 32·7-2. All the advantages of superhet set are based on the fact that frequency of various r.f. input signals is reduced to a fix difference frequency called intermediate frequency. This need for frequency conversion is fulfilled by means of a pentagrid converter. The converter section is shown in fig. 22.

Grids 1, 2 and cathode, work as the components of oscillator, Grid 2 acts as anode for local oscillator containing coil L_7 connected to high tension through a resistor R_{20}-R_{18}-C_{18} combination is

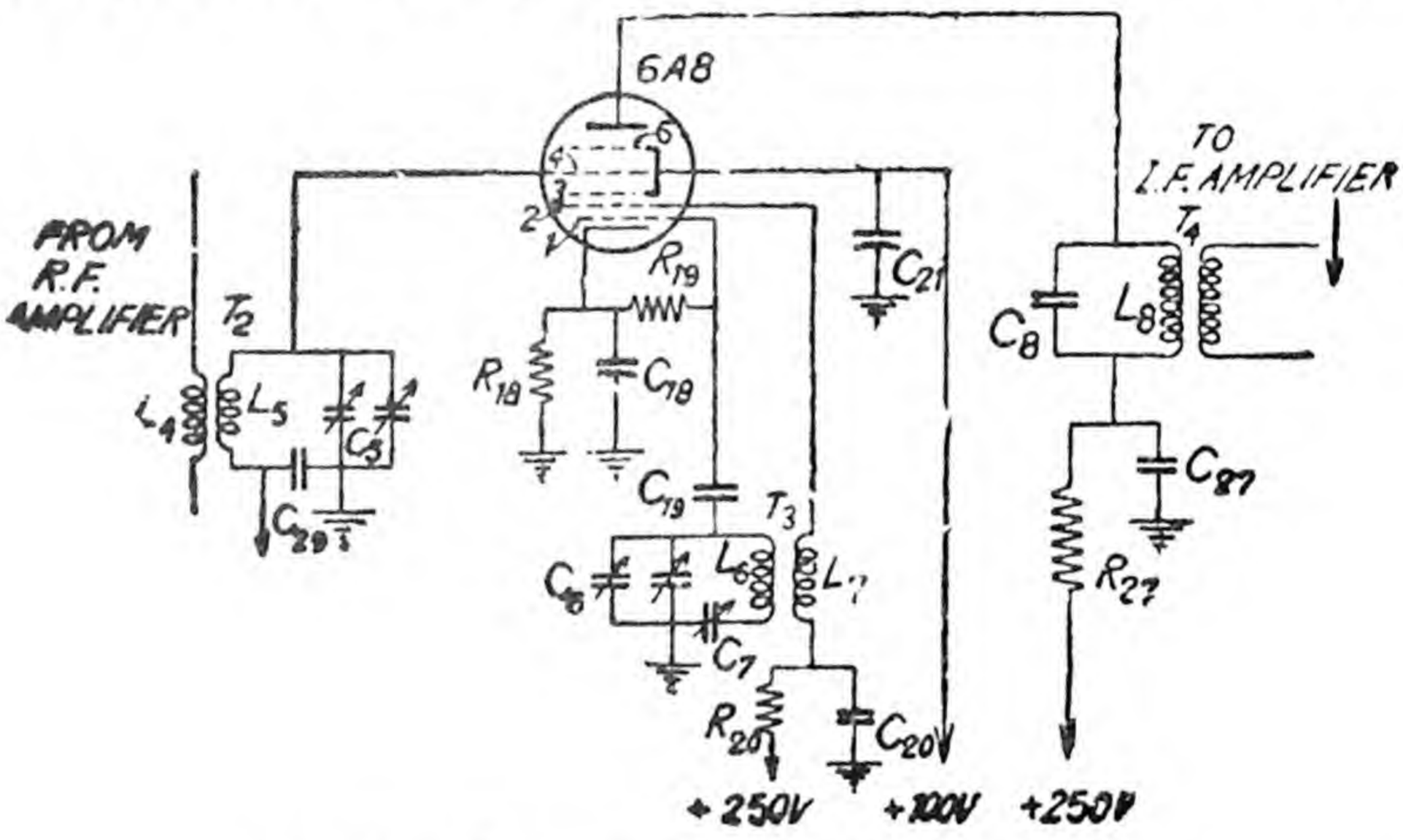


Fig. 22 Converter section of receiver

autobiassing arrangement. Radio frequency signal after preselection by r.f. amplifier stage is applied to the grid 4 by means of r.f. transformer T_2.

When high tension is switched on, the local oscillations are set up in the oscillatory circuit which send oscillatory electron stream towards grid 3. The r.f. signal applied to grid 4 further modulates the electron stream. The plate current of 6*A*8 will then contain :

(*i*) the r.f. signal frequency f_s,
(*ii*) the oscillator frequency f_0,
(*iii*) sum of oscillator and signal frequencies, *i.e.* (f_0+f_s),
(*iv*) the difference of r.f. signal and oscillator frequencies, *i.e.* (f_0-f_s).
(*v*) other frequencies due to combination of harmonics of above frequencies.

For example, r.f. circuit is tuned to 550 kc/s, and the oscillator is tuned to 1,015 kc/s, then plate current will have frequencies 550 kc/s, 1015 kc/s. 1565 kc/s, and 465 kc/s. and harmonics of these frequencies. We require only difference frequency 465 kc/s.

In practice, the output load for the converter is a resonant circuit, tuned to the inermediate frequency (465 kc/s) which will result in acceptance of the 465 kc/s current and rejection of currents of all other frequencies. The local oscillator is arranged such that the component ($f_o - f_s$) is always equal to the intermediate frequency 465 kc/s. or 455 kc/s in home receivers.

(*c*) *I.F. Amplifier Stage* : The anode load of converter stage consists of a tuned transformer in which the primary and secondary are tuned to the intermediate frequency. The transformer offers high impedance to the valve at the desired intermediate frequency and a low impedance at other frequencies. This i.f. output is applied to the control grid of the tube 6K7. The load of this stage is again an intermediate frequency transformer.

These input and ontput intermediate frequency transformers of this stage, each with tuned primary and tuned secondary, have coupling about 1 percent and the Q about 90. The intermediate frequency gain of such a stage is ordinarily around 40 db. and the bandwidth about 1 kc/s, at 6 db. or 44 kc/s at 20 db. The inductance of each winding is generally about 1·0 millihenry, requiring a capacitance of about 120$\mu\mu F$.

The stage provides better gain and high selectivity.

(*d*) *Second Detector Stage* (*Detection, A.V.C. A.F. amplification*) : This stage-employs a duodiode-triode tube *6SQ7*.

The function of various resistors and capacitors can be pointed out easily if we remember detection and amplification processes. It is obvious that combination R_{26}, C_{26} and C_{27} (100 pf) serves as π type resistance capacitance filter in output circuit of diode. R_{27} serves as diode load resistor. C_{31} is for audio frequency coupling and R_{31} serves as grid load resistor.

The operation of the circuit is quite simple. At the positive peak of intermediate frequency signal the diode plates attract electrons from cathode These flow through the secondary windings of the i.f. transformer, through R_{26} filter resistor and through R_{27} resistor for volume control, returning to cathode. For modulated wave, the rectified diode current will have both d.c and a.c. components. The desired amount of a.f. voltage is picked off by the tap on volume control resistor R_{27}

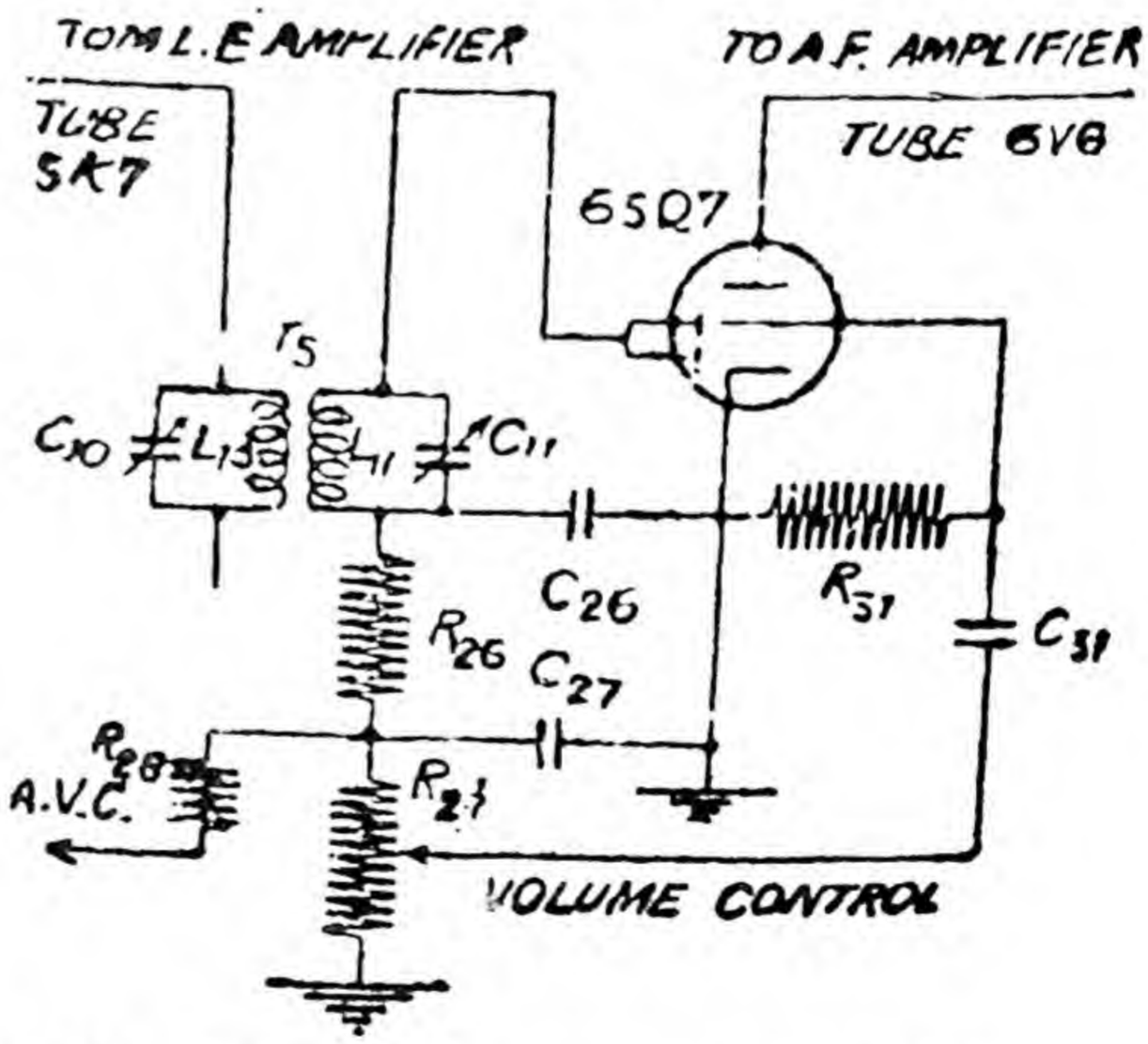


Fig. 23. Second detector section of the receiver.

and applied to the grid of the triode which serves as first audio amplifier. The rectified d.c. voltage across R_{27} acts through R_{28} filter resistor in the a v.c. bus to control the grid potential of tubes of previous stages. As the signal strength at the antenna increases, more signals will be rectified by the diode and a greater negative bias will be applied to the grids of controlled tubes. The change of bias reduces the gain of previous stages with the result that the output of the receiver changes by much smaller ratio than the input. Thus a uniform response is obtained by the receiver. For very weak signal reception delayed a.v.c. is preferred in receiving sets.

The output of this stage is applied to next amplifying stage through *RC* coupling.

(*e*) *Final A.F. Amplifier*: The amplified input from second detector stage is amplified by it and then the output is coupled to a loudspeaker. The tube used is 6*V*6.

(*f*) *Power supply unit*: A full wave rectifier tube 5*Y*3 is used Combination of C_{15} L_{16} serves as filter. The required output is taken from various tappings on output load resistor

(*g*) *Tuning Indicator*: The commonly used tuning indicator in broadcast receivers is electron ray tube also, called a magic eye or a cathode ray indicator tube. Fig 24 shows the circuit connection

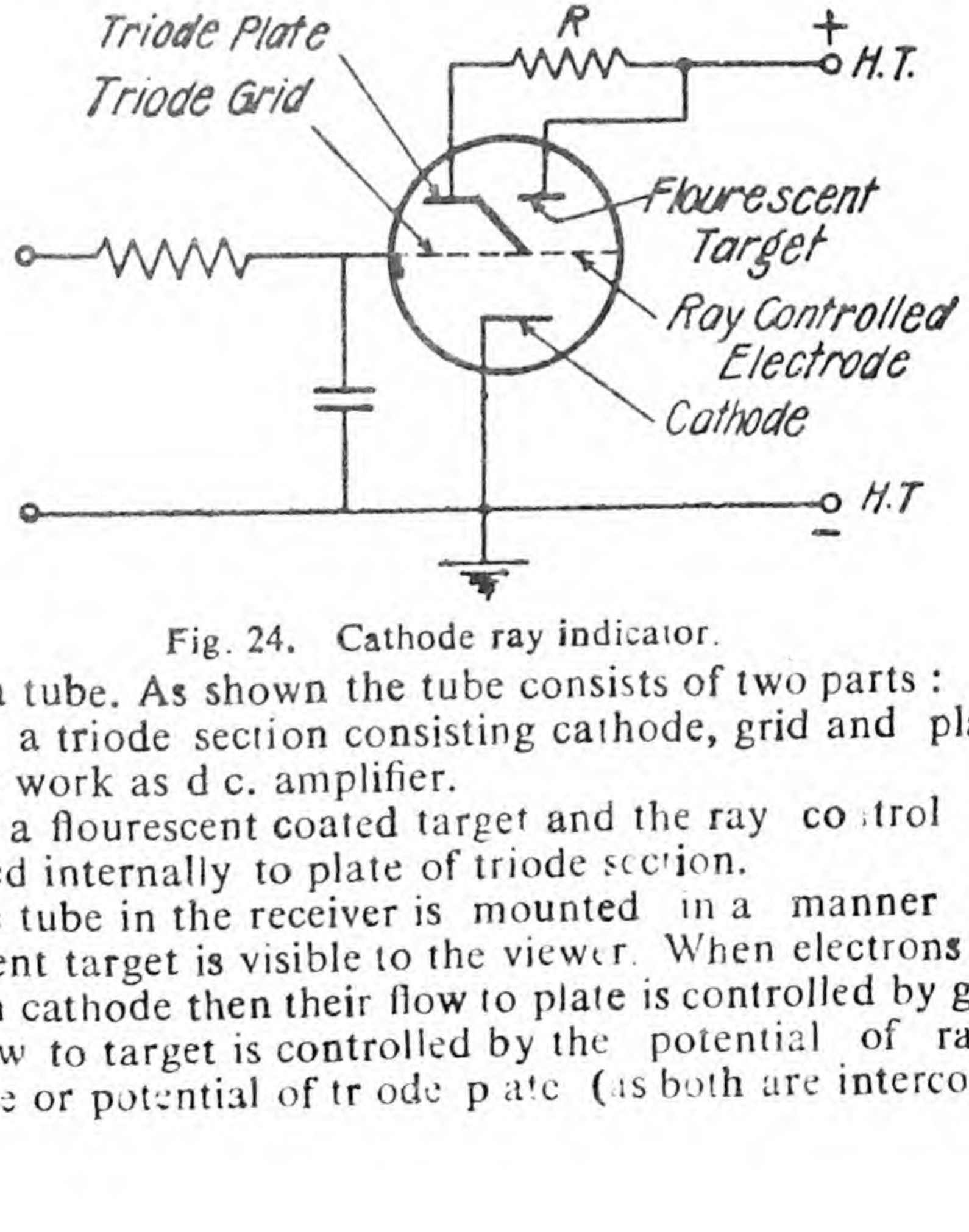


Fig. 24. Cathode ray indicator.

of such a tube. As shown the tube consists of two parts :

(*i*) a triode section consisting cathode, grid and plate which together work as d c. amplifier.

(*ii*) a flourescent coated target and the ray control electrode connected internally to plate of triode section.

The tube in the receiver is mounted in a manner that the fluorescent target is visible to the viewer. When electrons are emitted from cathode then their flow to plate is controlled by grid while their flow to target is controlled by the potential of ray-control electrode or potential of triode plate (as both are interconnected).

When electrons strike the target, it gives faint green fluorescent light. When the ray-control electrode and the target are at the *same* potential, the target will be *evenly illuminated* or may have *small* shadow sector but if ray-control electrode is at a less potential than target (*i.e. negative with respect to target*), most of the electrons are repelled back by the electrostatic field of this electrode and only a few electrons reach the target so that illumined area of the target *decreases* (shadow increases).

Triode control electrode *i.e.* grid is connected to the a v.c. line of the receiver. When receiver is *correctly* tuned to a station, a.v.c. voltage attains its highest negative value so that, grid being most negative, plate current decreases to its lowest value. Therefore d.c. voltage drop across resistor *R* decreases to its minimum value. This in effect increases the potential of ray control electrode to maximum value. Thus most of the electrons will impinge on the target and its maximum area will be illuminated *i.e.* shadow area will be minimum. Thus smallest shadow area is an indication of correct tuning.

(*h*) *Tone Control* : This is used to adjust relative amplitudes of higher frequency tones in the audio output of the receiver according to the taste of the listener. The commonly used arrangement for tone control consists of a series combination of a fixed capacitor,

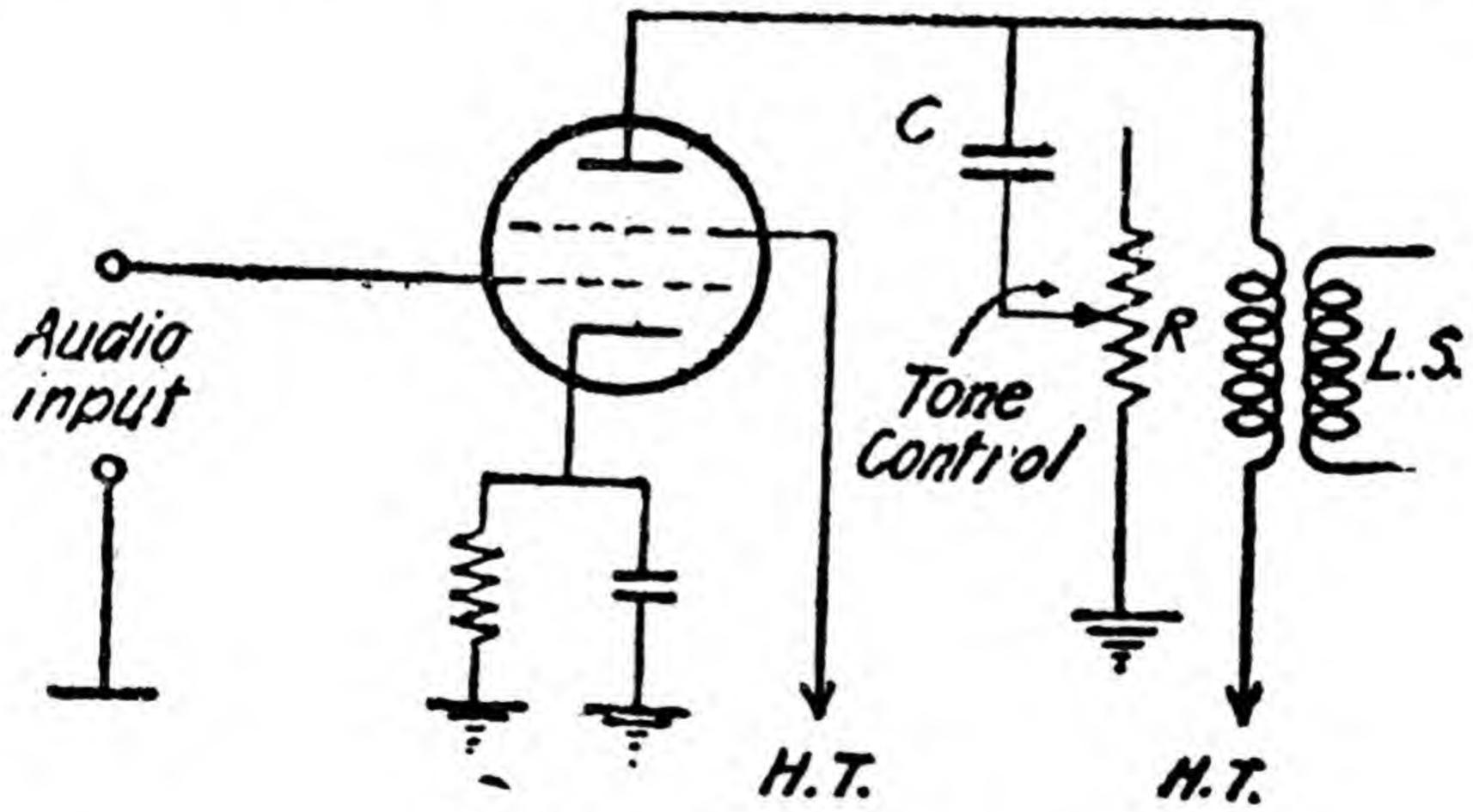


Fig. 25. Tone Control.

C_1 and a variable resistor, *R*, placed in shunt with the load resistance in an audio voltage amplifier (fig. 25). The combination acts a shunting path for a portion of a.f. plate current and thereby tends to lower the amplitude of signal. As the value of resistor *R* is increased relative amplitudes of higher frequency terms are increased.

INTERMEDIATE FREQUENCY CHARACTERISTICS :

While choosing intermediate frequency for the set, we keep two main considerations in view :

(*i*) for the rejection of image frequency interference, higher values of i f. should be preferred,

(*ii*) for better gain and selectivity, low values of i.f. should be chosen.

Therefore, for a compromise between two considerations, we can not choose too high or too low a value of intermediate frequency. In different receiving circuits, different i.f. values are used as tabulated below :

Receiving sets	Most widely used i f.
A.M Broadcast r ceiver	455 kc/s.
F.M. Broadcast receiver	10 7 mc/s.
Television receivers	{Picture carrier —26 mc/s {Sound carrier —21·5 mc/s.
Radir receiver with carrier frequency in u. h f and s h f. range	0 to 60 mc/s.

We consider the choice of intermediate frequency in view of the following main points :

(*a*) *Image Frequency Interference* : With a certain value of oscillator frequency, it is possible to obtain same intern ediate frequency with two different carrier frequencies. Suppose the required station sends carrier frequency of 100 kc/s. and oscillator frequen y is 1030. The i f. produced will be 30 kc/s. For another station sending 1060 kc/s. carrier frequency, oscillator frequency being the same, *i.e.* 1030 kc/s. the i.f. produced will be again 30 kc/s. This undesired i.f. due to second station is called the *image frequency signal*. This probability of image frequency interference will obviously decrease if we choose higher value of intermediate frequency The effect of 'mage frequency signals may also be eliminated by the use of one or more pre-selection stages. The ratio of the output from the desired r.f. signal to that from undesired r f signal is called signal to image ratio The image frequency effect becomes important in short wave receivers.

(*b*) *Selectivity* : Selectivity of stations on adjacent broadcast channels is improved if we use low i.f. values such as 130 kc/s and 262 kc/s This can be shown easily by expressing the difference in frequency between stations on adjacent channels as a percentage of the frequency at which tuning and amplification take place. Suppose 10 kc/s is difference in frequency between stations on adjacent channels ; then percentage of frequency for an i.f of 130 kc/s, will be 7·7% For 262 kc/s, i f it will be 3·8% and for 455 kc/s i.f., it will be 2·2 perecent. Thus it shows that for better selectivity low i.f. values should be adopted.

(*c*) *Tuning ratio* : The intermediate frequency can be obtained by making the oscillator frequency either lower or higher than the

r. f. signal by an amount equal to intermediate frequency chosen. But as regards tuning ratio, it is advantageous to keep the oscillator frequency greater than the r. f. signal. Suppose we require the receiver to operate over a r.f. range of 550 kc/s. to 1600 kc/s. and maintain 465 kc/s. as intermediate frequency. Then the oscillator will have to be tuned from 85 to 1,135 kc/s. if the oscillator is to be kept lower, or from 1,015 to 2,065 kc/s. if the oscillator frequency is to be kept higher We see that in the former case tuning ratio is 13 to 1, while in latter it is 2 to 1. For this reason in most broadcast receivers oscillator frequency is higher than r.f. signal.

(*d*) *Spurious Responses* : Harmonics generated in some sections of receiver produce interferences at its output. These interferences are generally named as spurious responses. The principal sources of these responses are : (1) harmonics of the intermediate frequency generated by the second detector, (2) harmonics of the incoming r.f. signal generated in the converter tube, (3) reception of r.f signal of a transmitter operating at a frequency equal to intermediate frequency. For elimination of these effects receiver is carefully designed.

Alignment of Superheterodyne Receiver : In aligning superheterodyne receiver we proceed from last a. f. output stage towards antenna Therefore the order in which the circuits are aligned will be as follows :

(1) The I.f. amplifier circuit.
(2) The oscillator circuit,
(3) The r.f. or preselector circuit.

Aligning I.F. Amplifier Circuit : An output circuit is connected across the secondary terminals of output transformer T_6 and is kept connected throughout the alignment procedure. While tuning i.f. circuit, the local oscillator and automatic volume control should be made inoperative. A signal generator is connected with its ground terminal to the chassis of the receiver and its high side terminal connected through a 0·01 μF capacitor to the control grid of 6*K*7. The receiver volume control should be tuned on fully, the signal generator frequency is set to correspond with the recommended value of i.f. for the set. Output of signal generator is varied enough to produce an indication on the output meter. The trimmer capacitors C_{10} and C_{11} (3 to 40 p. f.) are then adjusted to obtain maximum indication on the output meter. When the i. f. transformer T_5 is property adjusted, the next step is to adjust i.f. transformer T_4. The procedure is same except that the output of signal generator is now connected to r.f. control grid of 6*A*8. Trimmers C_8 and C_9 are adjusted for maximum indication on the output meter.

Aligning the Oscillator and R.F. Circuit : Now local oscillator is made operative With the signal generator still connected to the r.f. control grid of 6*A*8, set the receiver dial and the signal

generators frequency each to a value near 1400 kc/s, but not to a point at which a station is received and adjust the oscillator trimmer capacitor C_6 to obtain maximum indication on the output meter.

The local oscillator and preselector circuits can also be adjusted by connecting the signal generator output to the antenna terminal through a 200 $\mu\mu F$ capacitor and then setting the receiver and signal generator frequency near 1400 kc/s., we adjust the trimmer capacitors C_2, C_5, C_6 for the maximum indication on the output meter.

The local oscillator must also be adjusted for some value at the low frequency end of the dial. The low frequency alignment may be made by setting the receiver dial and the signal generator frequency at 600 kc/s. The oscillator padding capacitor C_7 is then adjusted to obtain maximum indication on the output meter while gently rocking the ganged capacitors back and forth to keep the signal tuned in The low frequency adjustment should be made only at the padder capacitors and no adjustment of trimmer capacitors

Tracking : For tracking the oscillator circuit of the superheterodyne receiver with the preselector circuits, padder capacitors are employed. All the ganged tuning capacitors are of the same construction. The capacitors used to tune the r.f. circuits must have a greater range than the capacitors used to tune the oscillator, the reason being that as the oscillator is tuned to a higher frequency, the ratio of high to low frequency values is less for the oscillator than for the r.f. circuit In order to use identical capacitors for all the tuning circuits, padder capacitor is connected in series with the original tuning capacitor to reduce the ratio of the maximum or minimum values of the series combination. Padder is made adjustable for tracking procedure. Radio receivers may be designed so that exact tracking will be obtained at these frequencies in the band and satisfactory tracking also provided over remainder of the band.

32·9. TRANSISTOR RECEIVER CIRCUITS :

The basic principles and circuit applications using vacuum tubes, described earlier, for local oscillator, frequency converter, and i f. amplifier can also be applied to similar transistor applications.

32·9-1. CONVERTER CIRCUITS (Mixer circuit) :

A transistor converter circuit consists of an oscillator and a mixer. The mixer section is shown in fig. 27. Through C_1 and R_2, input from i.f. amplifier is fed to the base of CE transistor. The required positive base voltage is obtained from the voltage

divider R_1 R_2. R_3 is to provide stabilisation of the collector current.

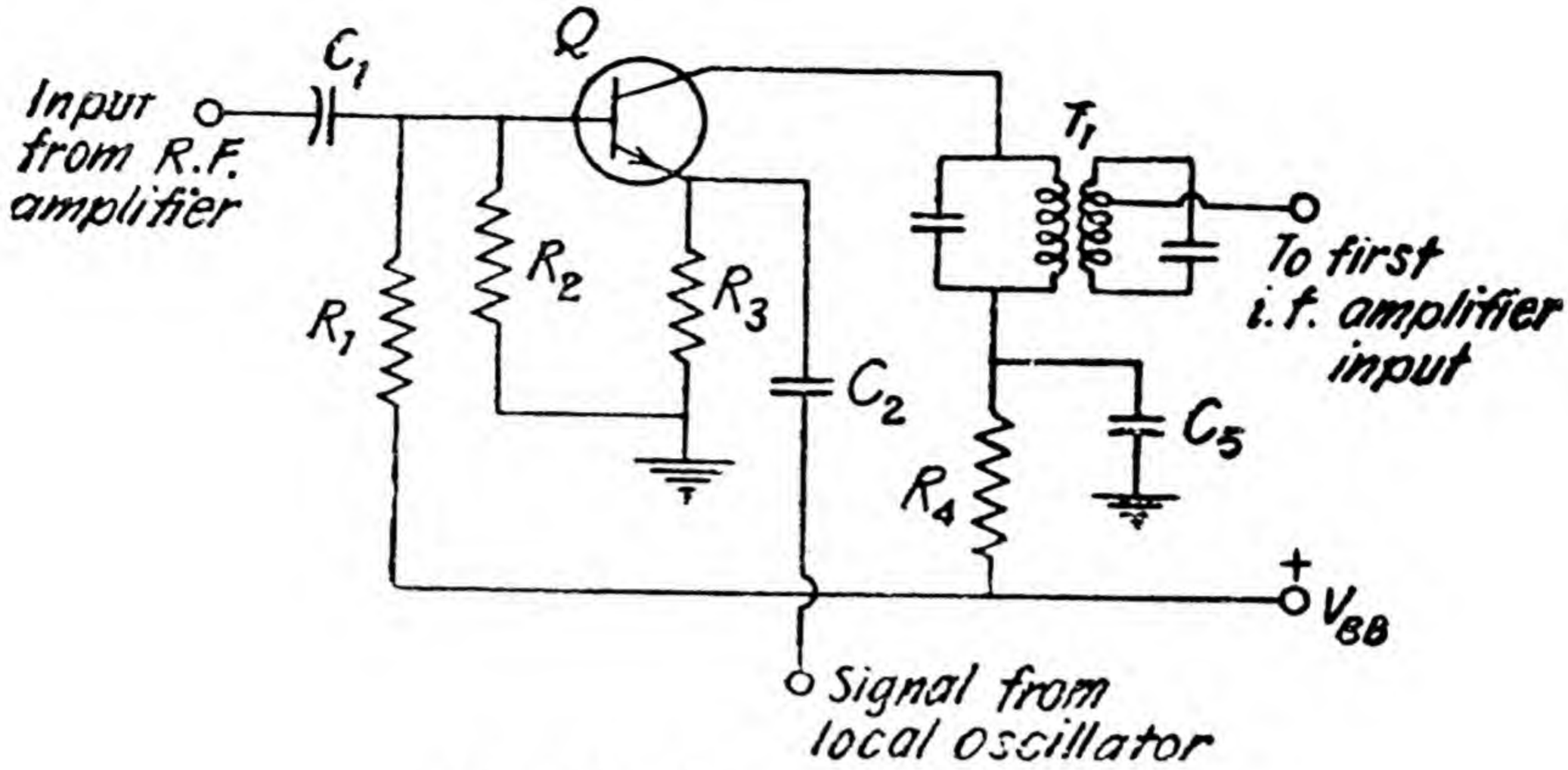


Fig. 27. A mixer circuit.

The incoming r.f. signal of frequency, f_i, is mixed with local oscillator signal of frequency, f_0. as follows : the oscillator signal is coupled to the emitter through C_2. Application of this additional signal voltage to emitter drives the emitter into conduction during negative half cycles Thus transistor operates at class *B* operating point. The base emitter diode current is modulated by both r.f. input and oscillator signal. Consequently, collector current will vary at both r.f. and oscillator signal; The difference frequency ($f_0 \sim f_i$) is selected by i f. transformer T_1.

Autodyne Converter :

It is a self oscillating mixer, *i e.*, in it the same transistor elements are simultaneously used as an oscillator and a mixer. The circuit is shown in fig. 28. R.F. input is inductively coupled to

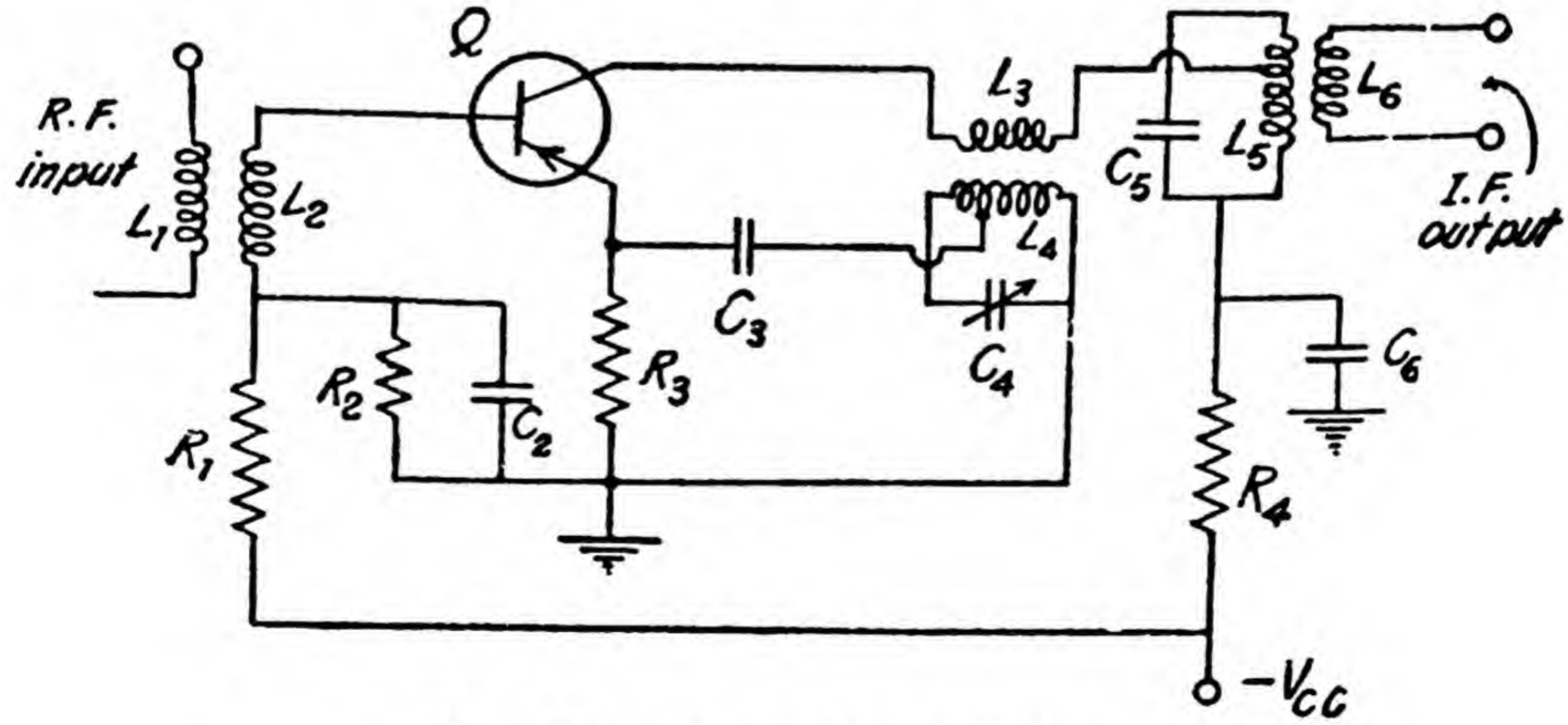


Fig. 28. Autodyne converter.

the base of the transistor. Base bias is obtained *via* R_1 and R_2. L_4C_4 constitute the oscillator tank circuit and is tuned to the desired

oscillator frequency, f_0. Energy from the collector circuit is fed to this circuit and tank circuit develops r.f. voltage at frequency, f_0. This oscillator signal is injected into the emitter circuit by R_3–C_3 coupling. Thus base emitter diode current is modulated by both r.f. input and oscillator signal. The difference frequency ($f_0 \sim f_i$) is selected by output tank circuit L_5C_5. Note that feedback from output *via* L_3 keeps the circuit L_4C_4 oscillating.

A dual gate MOSFET mixer circuit :

The circuit is shown in fig. 29. The input r.f. signal is fed to onegate and the oscillator voltage, from a separate oscillator circuit, is fed to the second gate through C_2R_1 coupling. This minimises the chances of interaction between two input signals. FET is biased into its non-linear region using source bias R_3C_3 combination. As a result, drain current is distorted and will con-

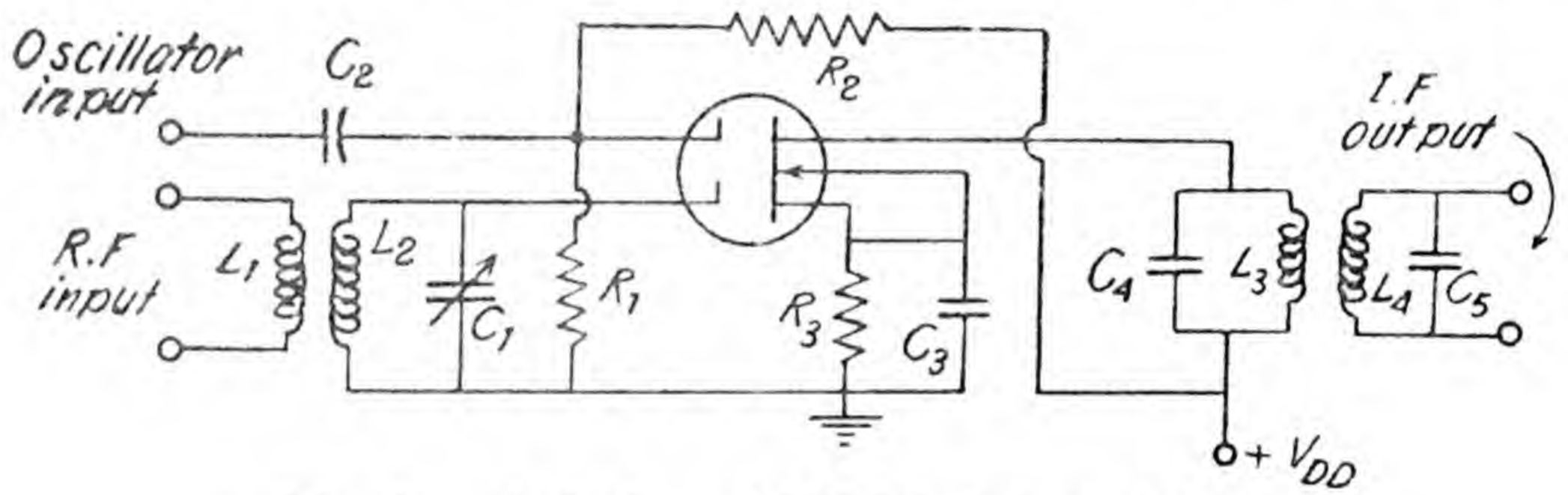


Fig. 29. A dual gate MOSFET mixer circuit.

tain many components—one of which is at the difference frequency ($f_0 \sim f_i$). The drain tank circuit L_3C_4 selects the difference frequency.

I. F. AMPLIFIER CIRCUITS :

A single stage i f. amplifier is shown is fig. 30. I. F. amplifier is a fixed radio frequency amplifier. It rejects adjacent unwanted frequencies. Since i.f. amplifier operates at a fixed frequency, the double tuned inductively coupled circuits can be used for improved band pass action, and for better selectivity it is also necessary that Q of the transformer is maintained. Compared to vacuum tube circuits, in transistor i.f. amplifier stage, there are two main problems :

(*i*) *Loading problem* : The high output impedance of one stage is to be matched with low input impedance of the following stage. It is done by

(a) tuning the primary circuit of each transformer to the desired frequency by variable inductance and fixed capacitors C_1 and C_4.

(b) having the secondary circuits of each transformer untuned, and

(c) tapping the primary winding of each tranformer.

(*ii*) *Feedback problem* : The collector to base capacitance of the transistor may be high enough to cause undesirable feed-

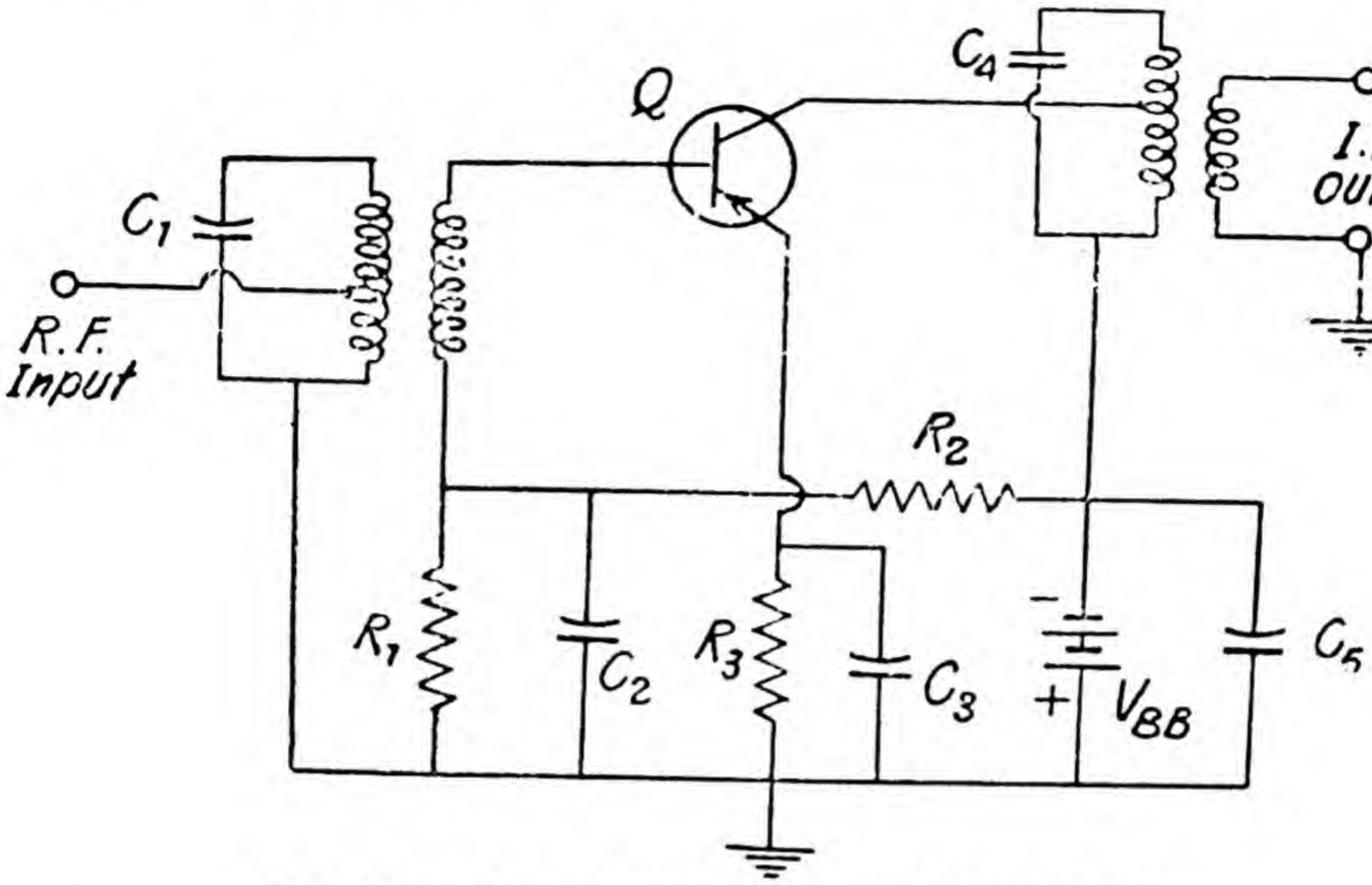


Fig. 30. Single stage i.f. amplifier circuit.

back. Therefore neutralising circuits are to be used to achieve stability.

In circuit of fig. 30 proper forward bias is obtained by the voltage divider network R_1R_2 and degeneration is minimised by C_2 and C_3. For high Q (better selectivity) i.f. transformer should have an output resistance of about 25 kΩ. Further collector is connected to a tap on primary winding of i.f. transformer and not shunted across the entire primary winding so that selectivity is not seriously affected. The desired circuit Q and proper impedance match can be obtained by selecting the correct position of the tap.

Usually it is impractical to use a single stage i.f. amplifier due to low detection efficiency at low signal levels and also due to bandwidth consideration. *Better gain, bandwidth and selectivity is obtained with two stages of i.f. amplification.*

Neutralising Circuits :

When collector to base capacitance of the transistor, being used, is large enough to produce undesirable feedback, neutralising circuits are then used. Neutralisation of i.f. amplifier stages may be accomplished by connecting a feedback capacitor, C_n, from the base of the transistor in one stage to the base of the transistor in the preceding stage. The value of this capacitor is given by

$$C_n = nC_o,$$

where C_c is collector-base capacitance (in pF) and n is the turns ratio (between the tap on the primary and turns on the secondary of the i.f. transformer).

Automatic Gain Control : Refer to art. 31·5.

DETECTOR CIRCUITS :

The crystal diode as a half wave rectifier is frequently used in a.m. detectors. Their operation as an a.m. detector for transistor radio receivers is the same as for vacuum tube radio receivers. They offer the advantages of (i) small size, (ii) low internal capacitance, (iii) low dynamic resistance, (iv) high frequency and (v) requiring no operating power. Refer to chapter on demodulation.

32·9-2. TRANSISTOR RADIO RECEIVER CIRCUIT :

Refer to the circuit of fig. 31. It uses four transistors (*NPN* type) and one diode (second detector). This unit has five stages, described as follows :

(i) **First Stage :** It is an autodyne converter (transistor Q_1). Resistor R_2 ($=10$ kΩ) in emitter circuit is to provide *dc* stabilisation against temperature changes etc. The positive voltage that emitter current develops across R_2 is counterbalanced by a positive voltage fed to the base from the battery. Resistor R_3 ($=2{\cdot}2$ kΩ) provides proper biasing voltage for the collector of Q_1. Bypass capacitor C_7 ($=0{\cdot}001$ μF) keeps the signal currents out of *dc* supply source.

(ii) **Second and third stage** (transistors Q_2 and Q_3) :

There are two i.f. amplifier stages They are necessary to attain better gain, bandwidth and selectivity. To match the high output impedance (collector impedance) of the preceding stage to the low input impedance of the following stage, the secondary of the i f. transformer is kept untuned while primary is tuned with a fixed capacitor. Neutralisation is achieved through C_{10} and C_{11} capacitors, each of value 10 pF.

AVC voltage is applied to the first i.f. stage (Q_2) only. It is the negative voltage obtained from the second detector.

The voltage divider network of R_{17} and R_{18} provide bias for second i.f. stage (Q_3). This bias voltage is bypassed by C_{22} and C_{13}.

Emitter resistor R_5 ($=1{\cdot}2$ kΩ) and R_8 ($=2{\cdot}7$ kΩ) provide d.c. stabilisation.

Each of the collectors of Q_2 and Q_3 receives its operating voltage through a $2{\cdot}2$ kΩ dropping resistor (R_7 and R_{10}). C_{11} and C_{15} are the decoupling and bypass capacitors (each $0{\cdot}001$ μF).

(iii) **Fourth stage :** There is a second detector D_1, beyond second i.f. stage. The load resistor ($R_{12}=2$ kΩ) for the detector is the volume control. The low value of this resistor is needed to match the input impedance of the next audio stage (Q_4).

(iv) **Final stage :** The final amplifier (Q_4) is an audio frequency amplifier. Base bias is provided by voltage divider network R_{13} and R_{14}. The emitter is grounded through R_{15} ($=1$ kΩ) resistor. The output transformer matches 10 kΩ collector impedance of Q_4.

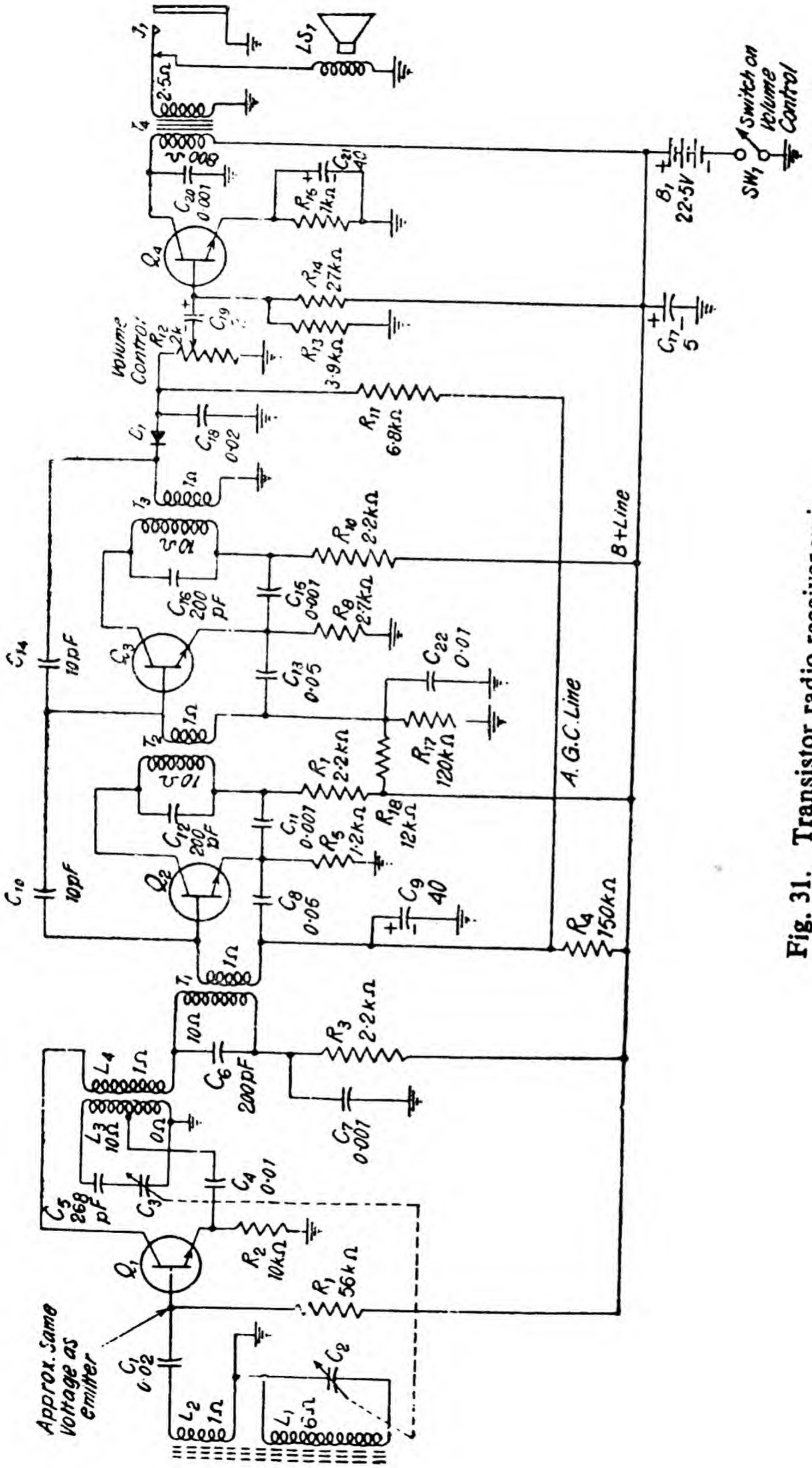
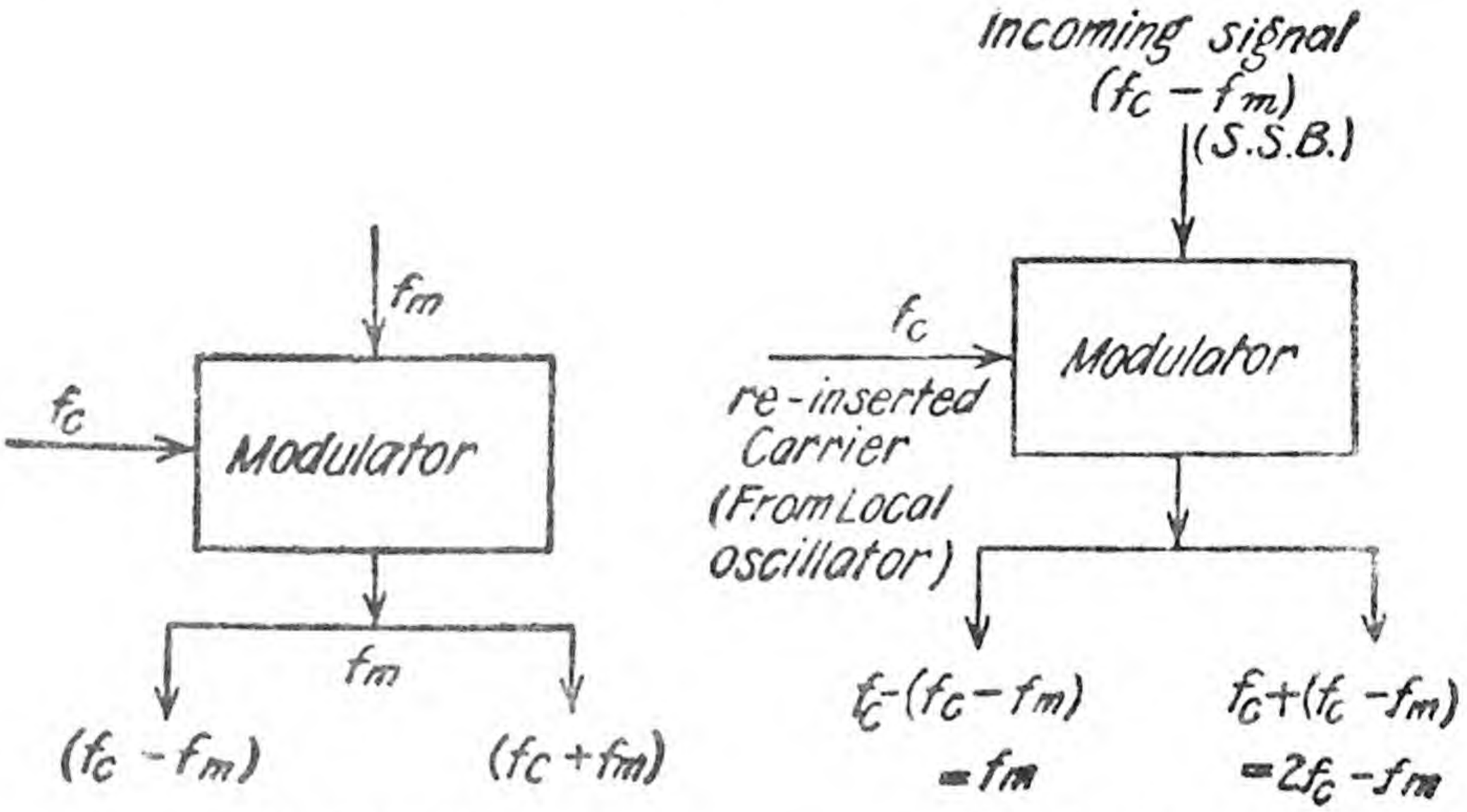


Fig. 31. Transistor radio receiver unit.

32·10. SINGLE SIDE BAND RECEIVER :

For the demodulation of a single side band signal, carrier is to be reinserted at the receiver. Then demodulation of *SSB* signal is carried out by modulating the received waveform using a local oscillator for generating the carrier frequency as illustrated in fig. 32 (*b*). The lower side band of frequency $(f_c - f_m)$, (as *SSB* signal input at the receiver) when modulated with re-inserted carrier, f_c, will produce an upper sideband of frequency $2f_c$ f_m and a lower sideband of frequency f_m. Since only the lower sideband is required, upper sideband is filtered out by some suitable circuits.



(*a*) Modulation (*b*) Demodulation

Fig. 32 *SSB* demodulation.

A balanced modulator, shown in fig. 33, may be used for demodulation of *SSB* signal. The carrier frequency generated by the local oscillator is re-inserted at the centre taps of the transformers and the incoming *SSB* signal is applied to the transformer

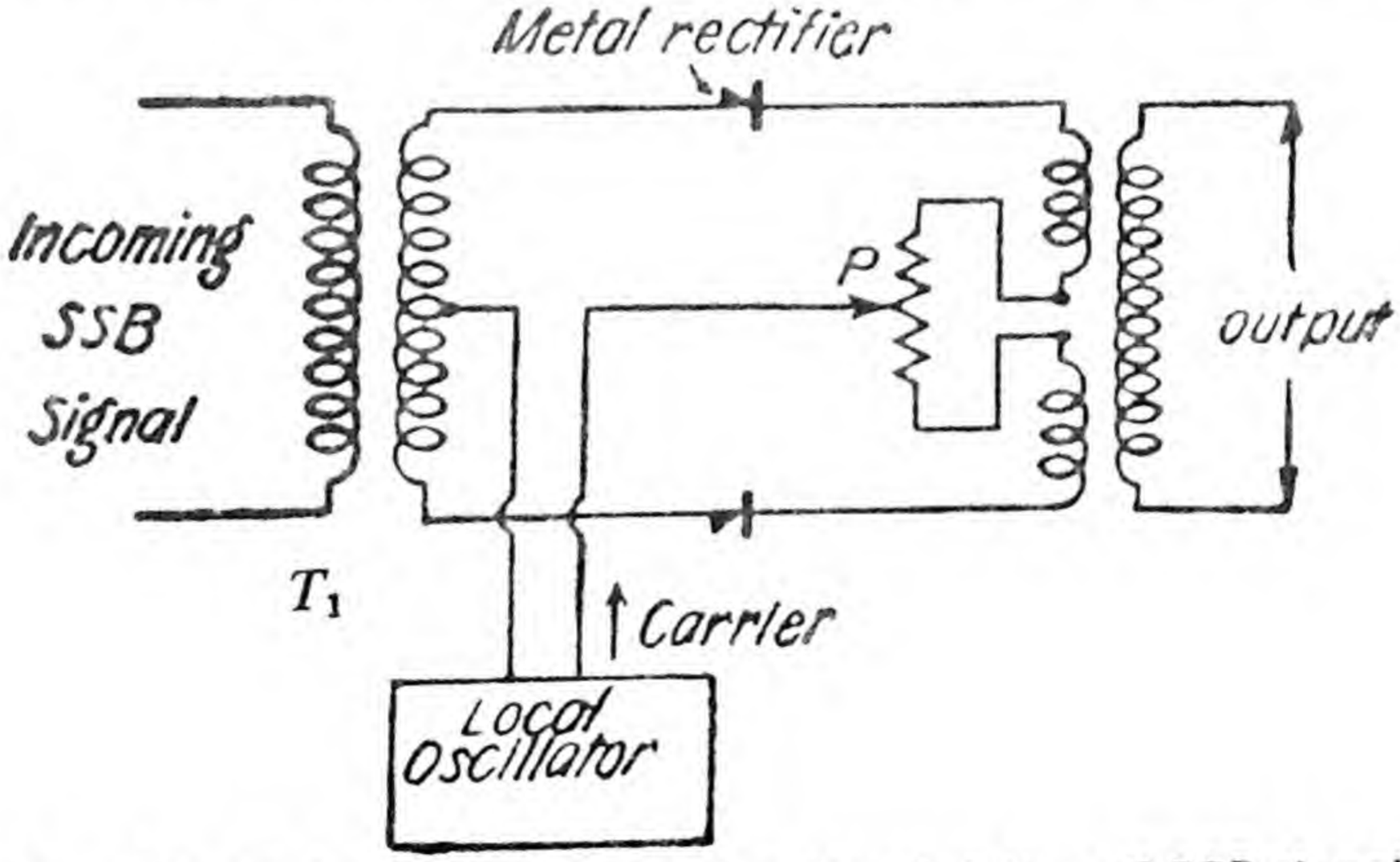


Fig. 33. Balanced modulator for demodulation of *SSB* signal.

T_1. The potentiometer P, known as the carrier leak control, is adjusted until the minimum carrier output is obtained. The metal rectifier is employed because of its small impedance when forward biased and very large impedance when back-biased, so that the output waveform is produced by the sideband being switched on and off at the carrier frequency. This output waveform contains a component of the original modulating signal.

The re-inserted carrier as required for the demodulation of *SSB* signal must have :

(i) same frequency as the original suppressed carrier,

(ii) same phase, and

(iii) requisite strength.

For this requirement the two carrier oscillators-one at transmitter and other of receiver must be in perfect synchronism otherwise the frequency of the demodulated output signal (audio) will be changed as discussed below :

Let the carrier of the sending end be

$$e_c = E_o \cos \omega_c t$$

and the modulating signal be

$$e_m = E_m \cos \omega_m t,$$

then the modulated wave is

$$e = E_o (1 + m_a \cos \omega_m t) \cos \omega_c t$$

$$= E_c \cos \omega_c t + \frac{m_a E_c}{2} \cos (\omega_o - \omega_m) t + \frac{m_a E_c}{2} \cos (\omega_c + \omega_m) t$$

If the carrier and upper sideband are completely suppressed at the transmitting end then transmitted wave is

$$e_t = \frac{m_a E_o}{2} \cos (\omega_o - \omega_m)\, t.$$

At the receiver this sideband is added to a local carrier which has a phase displacement, θ, from the original carrier, the voltage waveform available for demodulation is

$$e_r = E_o \cos (\omega_c t + \theta) + \frac{m_a E_c}{2} \cos (\omega_o - \omega_m) t$$

$$= E_o \cos \omega_c t \cos \theta - E_o \sin \omega_c t \sin \theta + \frac{m_a E_c}{2} [\cos \omega_c t \cos \omega_m t + \sin \omega_c t \sin \omega_m t]$$

$$= \left[E_o \cos \theta + \frac{m_a E_c}{2} \cos \omega_m t \right] \cos \omega_o t - \left[E_c \sin \theta - \frac{m_a E_c}{2} \sin \omega_m t \right] \sin \omega_c t$$

$$= \left[E_o^2 \cos^2 \theta + m_a E_c^2 \cos \theta \cos \omega_m t + \frac{m_a^2 E_c^2}{4} \cos^2 \omega_m t + E_c^2 \sin^2 \theta - m_a E_c^2 \sin \theta \sin \omega_m t + \frac{m_a^2 E_c^2}{4} \sin^2 \omega_m t \right]^{1/2} \cos (\omega_o t + \alpha),$$

where $$\alpha=\tan^{-1}\left[\frac{\sin\ \theta-\frac{m_a}{2}\sin\ \omega_m t}{\cos\ \theta+\frac{m_a}{2}\cos\ \omega_m t}\right]$$

or $$e_r=\left[E_c^2+\frac{m_a^2\ E_c^2}{4}+m_a\ E_c^2\cos\ (\omega_m t+\theta)\right]^{1/2}\cos\ (\omega_c t+\alpha)$$

if depth of modulation is small so that $m_a < 1$ and $m_a^2\ E_c^2$ term is neglected as compared to E_c^2 then

$$e_r=[E_c^2+m_a\ E_c^2\cos\ (\omega_m t+\theta)]^{1/2}\cos\ (\omega_c t+\alpha)$$

Which, on expansion by binomial theorem, gives

$$e_r=\left[\ E_c+\frac{m_a\ E_c}{2}\cos\ (\omega_m t+\theta)\ \right]\cos\ (\omega_o t+\alpha)$$

$$=E_c\left[\ 1+\frac{m_a}{2}\cos\ (\omega_m t+\theta)\ \right]\cos\ (\omega_c t+\alpha).$$

When $\theta=0$, modulation index, m_a, is the same as in modulated wave at the transmitter. Thus modulation of signal at the receiver reproduces the modulation envelope which corresponds to modulating signal at the transmitting end. Any variation of θ from zero simply causes a change in phase shift of the output modulation signal from its correct position. If the phase angle varies slowly with time on account of slight difference in frequency between suppressed carrier and reintroduced carrier, the phase of modulating signal varies accordingly. A slow variation of phase of output modulating signal at the receiving end with respect to that at transmitting end with time is equivalent to a change in modulating frequency. In other words, the two carrier oscillators (transmitting end and receiving end) should be exactly synchronised otherwise frequency of the derived modulating voltage at the receiver output will be changed.

In order to maintain the condition $m_a < 1$, the amplitude of re-inserted carrier must be large enough with respect to sideband. This is important otherwise harmonics of modulating signal will appear at the output of the receiver, causing distortion.

SSB Receiver :

A SSB receiver block diagram is shown in fig. 34.

The input signal (freq. 13·1—13·103 *mc/s*) is amplified by r.f. amplifier and then fed to first mixer where it is mixed with local oscillator frequency (10 *mc/s*) and beat frequency 3·1—3·103 *mc/s*, thus produced, is then fed to intermediate frequency amplifier. This amplified i.f. signal is further heterodyned in second mixer with second local oscillator frequency (3 mc/s) to produce a lower i f. of 100 kc/s. This i f. is then amplified and fed to amplitude demodulator where it is combined with reinserted 100 kc/s carrier (corresponding to pilot carrier). In order to keep $m_a < 1$ so that envelope of the combined wave varies inaccordance with the original modulation, the amplitude of the reinserted carrier is kept large compared with the amplitude of single sideband. The original

modulating signal is received by the conventional a. m. diode detector.

For faithful reproduction of signal it is required that the reinserted carrier be within 10 cycles of the orignal one. For this the output of second mixer which consists lower i.f. (100 kc/s), also called pilot carrier, is fed to a crystal filter which extracts 100 kc/s. carrier voltage. This 100 kc/s pilot carrier is amplified and passed through a limiter circuit to remove fading. This gives a constant

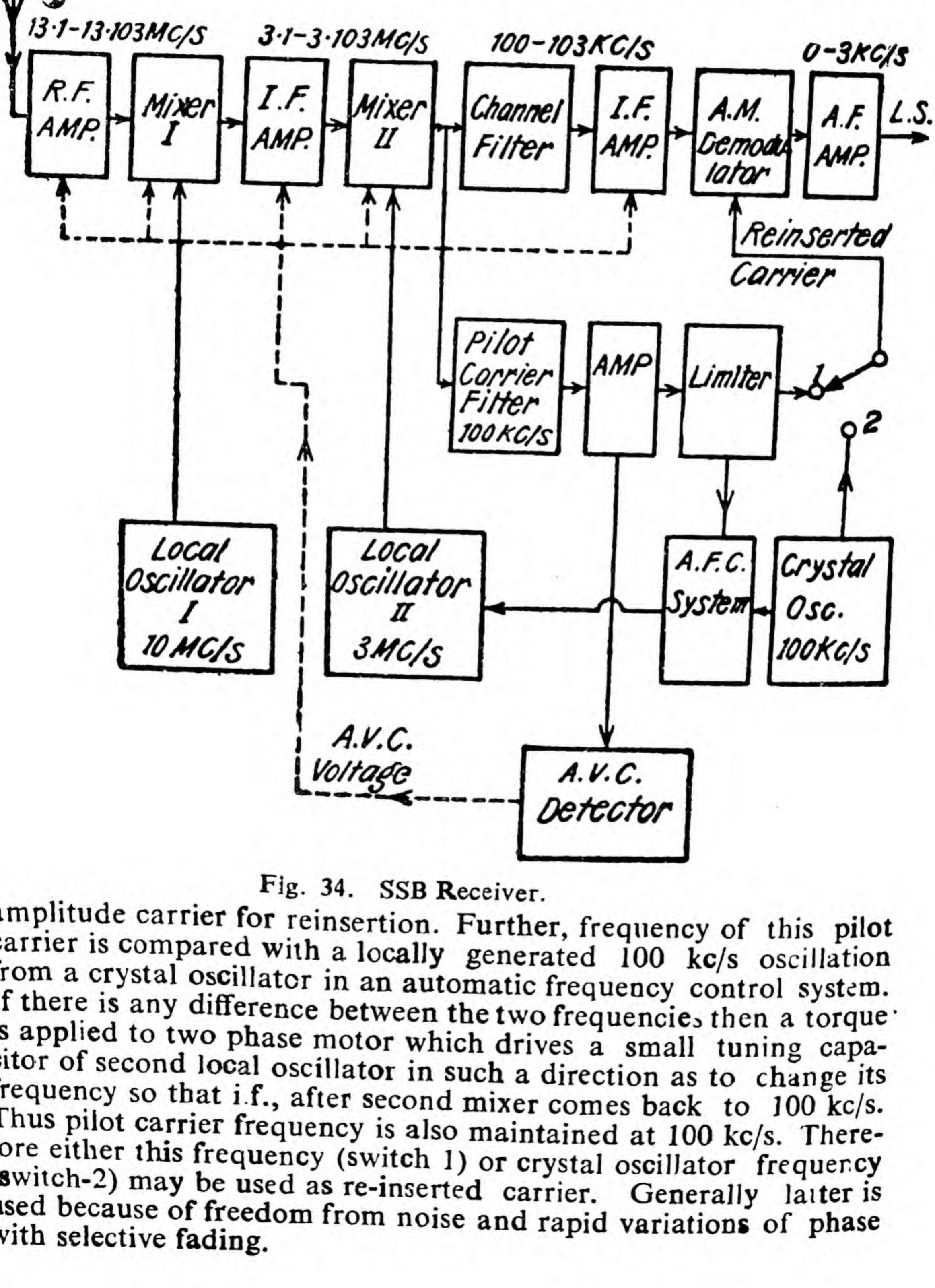


Fig. 34. SSB Receiver.

amplitude carrier for reinsertion. Further, frequency of this pilot carrier is compared with a locally generated 100 kc/s oscillation from a crystal oscillator in an automatic frequency control system. If there is any difference between the two frequencies then a torque is applied to two phase motor which drives a small tuning capacitor of second local oscillator in such a direction as to change its frequency so that i.f., after second mixer comes back to 100 kc/s. Thus pilot carrier frequency is also maintained at 100 kc/s. Therefore either this frequency (switch 1) or crystal oscillator frequency (switch-2) may be used as re-inserted carrier. Generally latter is used because of freedom from noise and rapid variations of phase with selective fading.

32·11. F.M. RECEIVER :

These days f. m. receiver, that overcomes the disadvantages encountered in a. m. broadcasting, finds extensive use in broadcast services. The carrier frequency in f.m. system can be varied over a wide range than the extent of the signal components. In frequency modulation broadcasting, the standard deviation is 75 kc/s. when high quality signals are to be transmitted, while military and police communications employ a deviation of 1 kc/s. F.M. broadcasting stations operate on assigned frequencies ranging from 88 mc/s to 108 mc/s.

SUPERIORITY OF F. M. RECEPTION OVER A.M. RECEPTION :

1. Frequency band : The bandwidth in the case of f.m. receiver are broader as compared to a.m. receiver. Bandwidth varies with the type of services for which the receiver is intended. A.M. broadcasting stations operate on band frequencies ranging from 540 to 1600 kc/s, while as stated above f.m. stations operate on assigned frequencies ranging from 88 to 108 mc/s. These frequency ranges have been assigned by F.C.C.

2. **Selectivity** : Selectivity depends upon the ratio of channel width to carrier fiequency which is greater in f.m. receivers. Therefore f.m. receivers will have a greater adjacent channel selectivity

2. **Noise Reduction** : F.m. receivers have greatest advantage of suppressing external noise in comparison to a.m. receiver. This noise can be :

(*i*) noise picked up by the antenna from passing noise or 'static' radio fields.

(*ii*) thermal noise generated by the resistance that the antenna system presents to the input terminals of the receiver, and

(*iii*) noise generated within the receiver consisting of circuit (thermal) and tube noise.

The superimposition of the noise over f.m. wave will cause fluctuations in amplitude and likewise produce phase variations. The noise contained in the wave can be prevented from producing appreciable noise in the output of receiver by

(*i*) employing a limiter or ratio detector so that the amplitude variation are smoothed out, and

(*ii*) using a frequency deviation at the transmitter so great that, for full modulation, the frequency deviation is large compared with unity, which, inturn, helps to suppress the fluctuation due to phase variations.

Thus the output of the receiver is practically free from noise due to static.

4. **Image Interference** : Image frequency interference is caused

by the signal having the frequency double that of intermediate frequency. In f.m. receiver intermediate frequency value is very high of the order of 10·7 mc/s : so double of this frequency will be 21·4 mc/s. which exceeds the complete span of f.m. band of 88 mc/s. to 108 mc/s. Thus image frequency interference is rare in f.m. receivers.

5. **Reduction of Co-channel Interference :** When two frequency-modulated transmitters located within the range of a receiver are operated on the same frequency, then stronger signal tries to suppress the weaker signal almost completely. As mentioned above in noise reduction, with f.m. receiver it is possible to separate the signals of two stations operating on the same frequency, if the interfering signal is not greater than fifty percent of desired signal.

6. **Fidelity :** The essential requirements for good fidelity set are :

(*i*) the set must be practically free from hum,

(*ii*) the reproduction must be practically free from harmonics and intermodulation frequencies at all volume levels.

(*iii*) the electrical characteristic must be practically uniform over a wide audio range, and

(*iv*) a good loudspeaker must be provided.

F.m. receiver satisfies all the above requirements to a greater extent as compared to a m. receiver. Therefore f.m. receiver has good fidelity, particularly because of suppressing noise to minimum.

BLOCK DIAGRAM :

A block diagram of basic receiver is shown in figure 35. F.M. system usually employs superheterodyne type receiver, which as shown in block diagram, includes r.f. amplifier circuit, converter circuit, r.f. amplifier section, f.m detector and audio amplifier section connected to a loudspeaker.

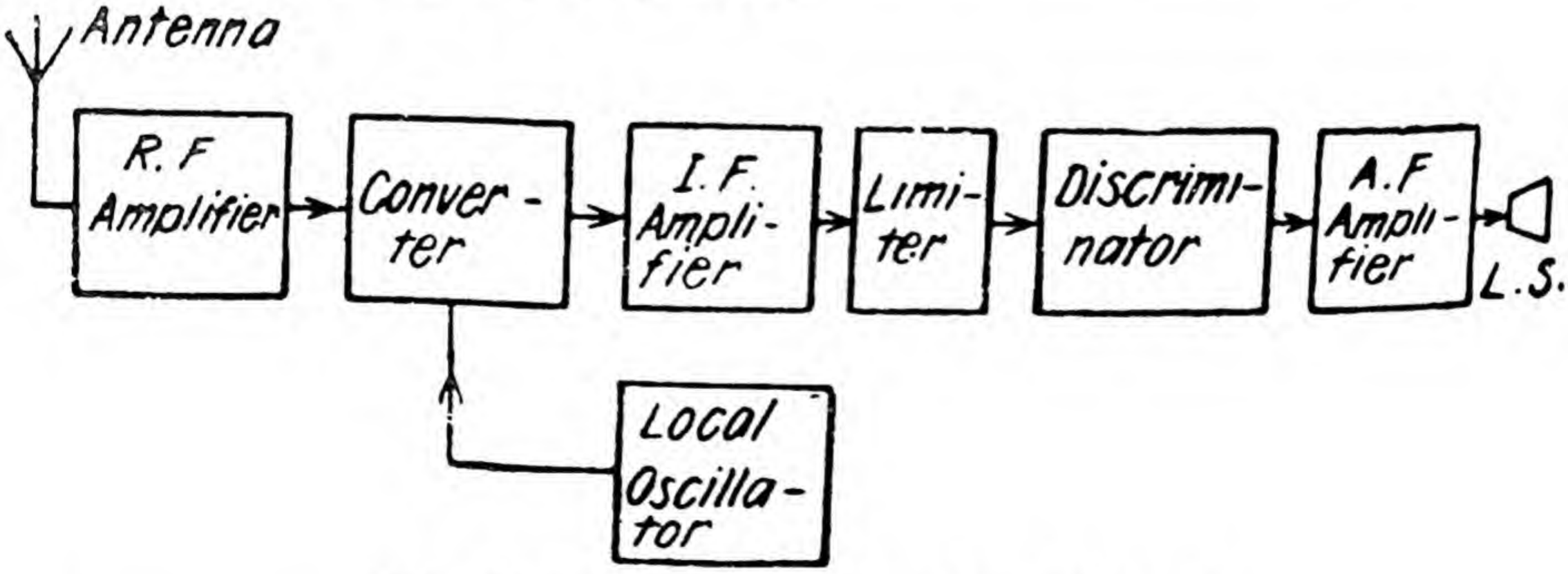


Fig 35 Block diagram of f m. receiver.

DIFFERENCE IN F.M. AND A M. RECEIVERS :

F.M receiver differs in the following respects from medium and high frequency a.m. broadcast receivers :

(*i*) The carrier frequencies are much higher.

(*ii*) The bandwidth is wider to accommodate the wide swing of the signal and the drift of the local oscillator.

(*iii*) The intermediate frequency used is higher being of the order of 10·7 mc/s.

(*iv*) The r.f. amplifier may contain one or more limiter stages.

v) Frequency-sensitive device to reproduce the original modulating signals is used in place of a.m. detector.

(*vi*) De-emphasis circuit is employed along with the audio circuit to remove the pre-emphasis inserted at the transmitter.

(*vii*) For good fidelity, appropriate loudspeaker is provided.

OPERATION OF F.M. RECEIVER :

The limiter and discriminator stages of f.m. receiver have been described in the previous chapter of detection. The stages before limiter are also similar to their counter-stages in a.m. receiver except some differences in the design of the circuit and tubes used. Therefore we shall not discuss much about f m. receiver

Due to some differences in the stages previous to limiter, we shall explain here the front end portion (the section of the receiver which first acts upon the received signal) of f.m. receiver in brief.

Figure 36 shows the front end portion of f.m. receiver.

The front end employs three tubes :

(*i*) a pentode VT_1 used as r.f. broadband ground cathode amplifier.

(*ii*) a triode mixer VT_2, and

(*iii*) a twin triode VT_3 of which one section is used as a tuned grid oscillator and the other section as a variable reactance tube for the a.f c. circuit.

The antenna coupler $C_1C_2L_1$ is of unbalanced type and its r f. signals are coupled to VT_1 by C_3 and R_1. The r.f. amplifier has wideband response and increases the strength of all signals in the f.m. band ; its output is coupled by C_5 to the tuning circuit $L_3C_6C_7$. The r.f. tuning circuit selects the signal of one station and rejects all other ; the signal of the selected station is fed to the grid of the mixer tube VT_2. VT_{3A} operates as a tuned grid triode oscillator that is tuned by $L_5C_{10}C_{11}$. The oscillator signal is coupled to the grid of the mixer tube by C_9. Capacitors C_6 and C_{10} are mechanically controlled by a single tuning dial. VT_{3B} is operated as a variable reactance tube which is coupled to the oscillator tuning circuit by C_{13} and keeps the oscillator frequency constant.

Description of Individual Stages :

(*i*) *Antenna and its coupling units* : The type of the antenna used with f.m. receiver depends upon the distance of broadcasting station from the location of receiver When the receiver is used

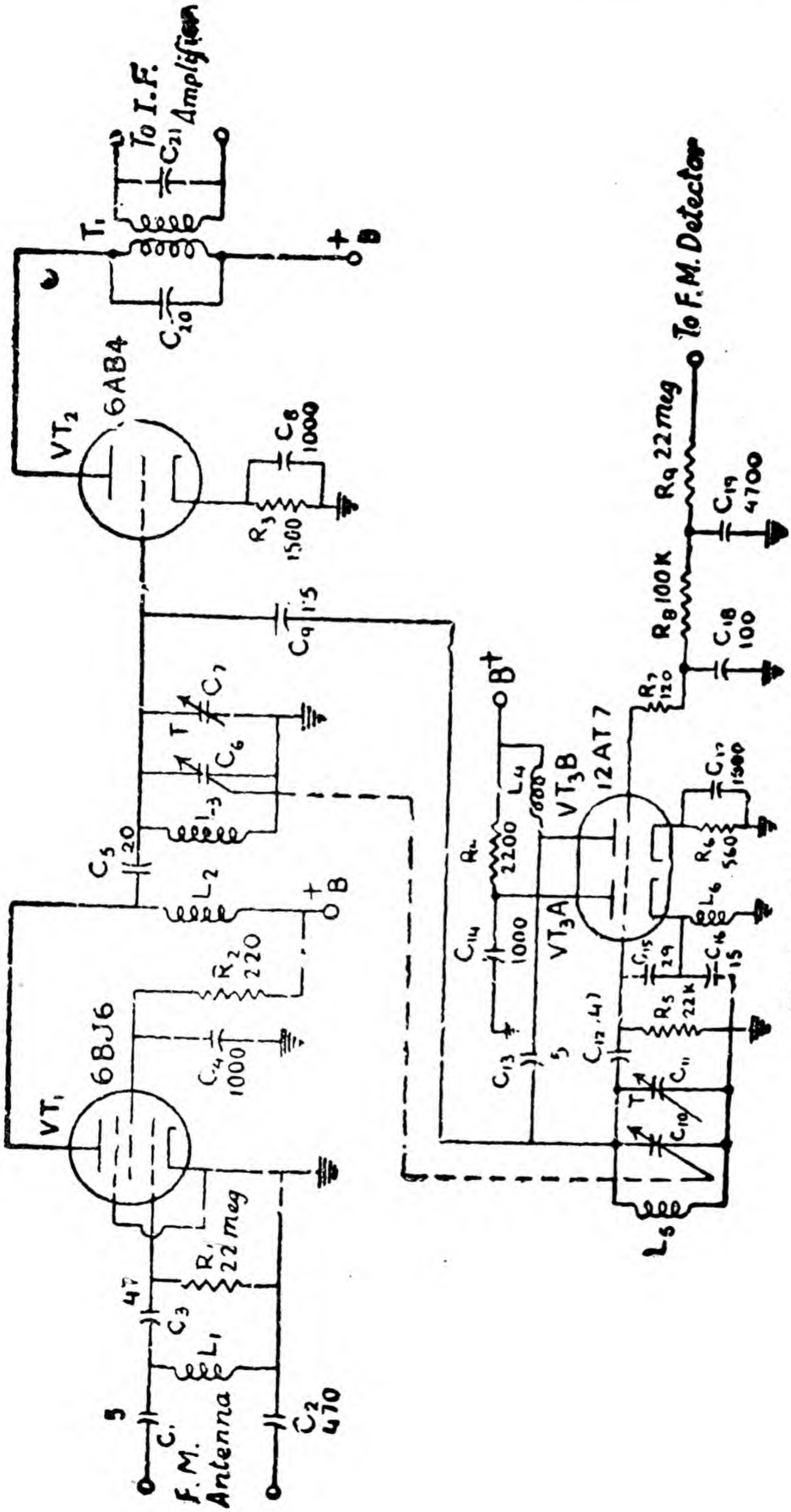


Fig. 36. Front end portion of a f. m. receiver.

in strong signal area, a self contained type of antenna having inherent ability of power lines to pick up radio signals is employed. In weak signal area, a commercial type of outdoor antenna is used.

The antenna coupling unit is a net-work of inductance, capacitance and resistance located between the input of the first tube and the signal input terminals of the receiver coupling the transmission line to the receiver properly. For improved selectivity and to develop maximum signal voltage at the input of the first tube, this line should be terminated by a resistive load of magnitude equal to the characteristic impedance of the line. When each side of the transmission line is at the same impedance above the ground then it is said to be balance type antenna coupler.

(*ii*) *R.F. Amplifier* : The r.f. amplifier stage increases the signal strength at the same frequency at which it is received. The functions of this stage are (*i*) to improve selectivity, (*ii*) to reduce image frequency effect, (*iii*) to maintain appropriate signal to noise ratio. Since high carrier frequencies are employed in f.m. system, the tubes used in r.f. sections are specially designed to minimise the difficulties that arise in high frequency amplifier circuits. Generally miniature tubes are used because of (*i*) close inter-electrode spacing so that electron transit time is reduced, (*ii*) small electrode areas to avoid excessive inter electrode capatance (*iii*) short leads to minimise lead inductance and power loss in leads.

The r.f amplifier used in f.m. receivers are usually broad-band amplifiers designed to have a wide band frequency response at the antenna coupler and relatively narrow band response at the inter-stage coupler. In order to get the desired bandwidth, these resonant circuits (antenna coupler are designed to have a sufficiently low value of Q. In fig. 36 a grounded cathode r.f. amplifier circuit is used due to its characteristic of high gain and high input impedance.

(*iii*) *Local Oscillator and Automatic Frequency Control* (*A.F.C.*): The local oscillator must supply r.f. voltage of correct frequency to the mixer so that on combining with signal from the desired station it produces the required intermediate frequency of 10·7 mc s. The prime importance in the design of this circuit is the maintenance of the stable oscillations Any frequency drift in the oscillator frequency will affect that i.f. value appreciably.

A.F.C. circuits are used to keep the oscillator frequency constant. In figure 36, VT_{3B} acts as a variable resistance in an f.c. circuit. Any frequency drift at the oscillator tube VT_{3A} will cause a change in the detector output voltage which in turn, as being coupled to the grid of VT_{3B}. cause a change in the reactance of this tube. This reactance by means of C_{13} is connected to the tank circuit of oscillator and hence any change in reactance will affect the

oscillator frequency in the right direction as to have maximum output at the f.m. detector.

(*iv*) *Mixer stage* : The voltages of both r.f. and oscillator signals are applied to the grid of tube VT_2. The oscillator voltage is made relatively larger than r.f. signal voltage which will cause the plate current to vary between zero and maximum of its value. This, in turn, will vary the transconductance of the tube inaccordance with its own variations. These variations of plate current and transconductance due to large oscillator voltage will form a series of pulses whose peaks will be modulated by relatively small r.f. signal. Therefore, the plate current will consist of pulses occurring at the oscillator frequency rate and also the side-bands above and below the oscillator signal frequency ; the side bands still contain r. f. signal modulations.

The transformer T_1 is tuned to select the lower side band frequency named as intermediate frequency which is 10·7 mc/s. in f.m. receiver.

(*v*) *I. F. Amplifier* : The stage is quite similar to its counterstage used in a.m receiver except that here it operates at higher frequencies, 10·7 mc/s and requires a wide bandpass.

The rest stages of f.m. receiver have been discussed in previous chapter. For de-emphasis circuit connected at the a.f. output of the f.m. detector, a simple arrangement is shown in fig. 37. This

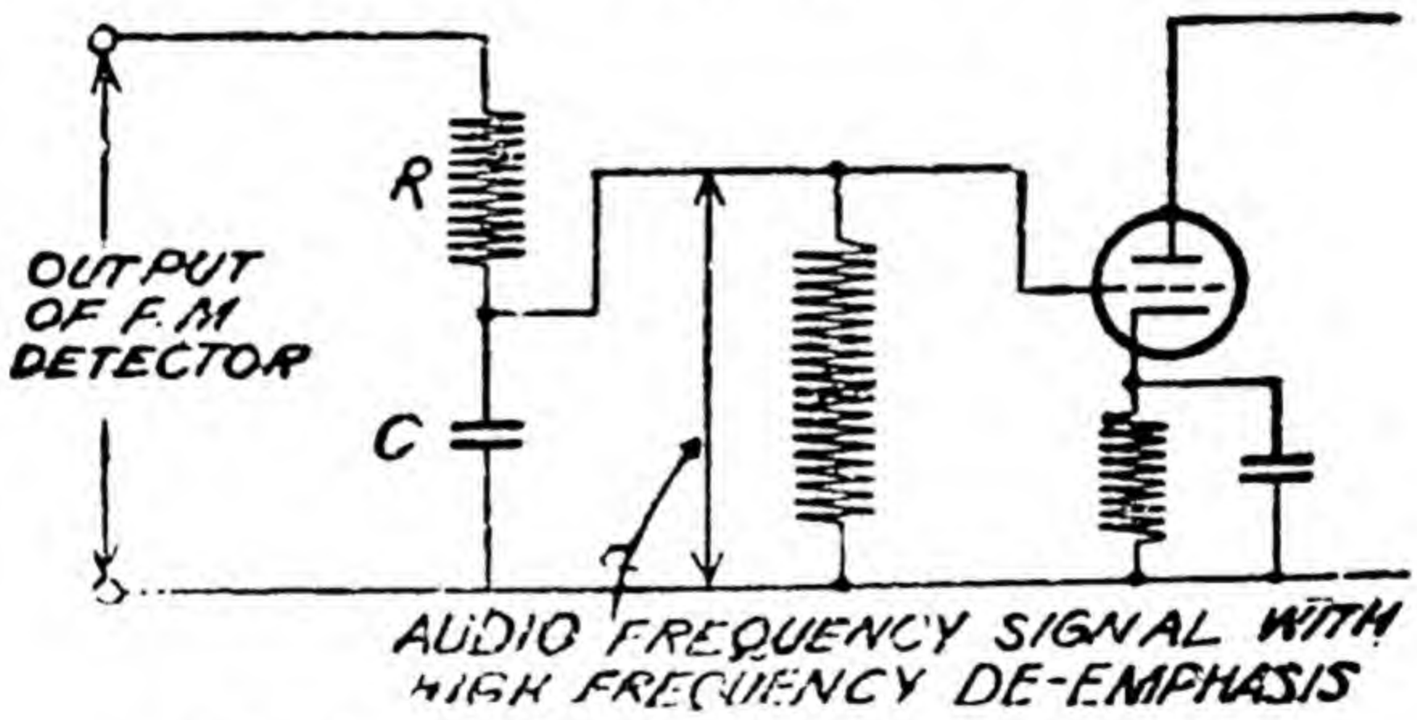


Fig. 37. De-emphasis Circuit.

consists of a simple RC network connected across the output of the f. m. detector. The higher the frequency, lower is the reactance of capacitance C and hence signal amplitude applied to the audio amplifier will be relatively lower. The circuit, thus, removes the pre-empasis inserted at the transmitter.

The de-emphasis network can be separately sketched as shown in fig. 38.

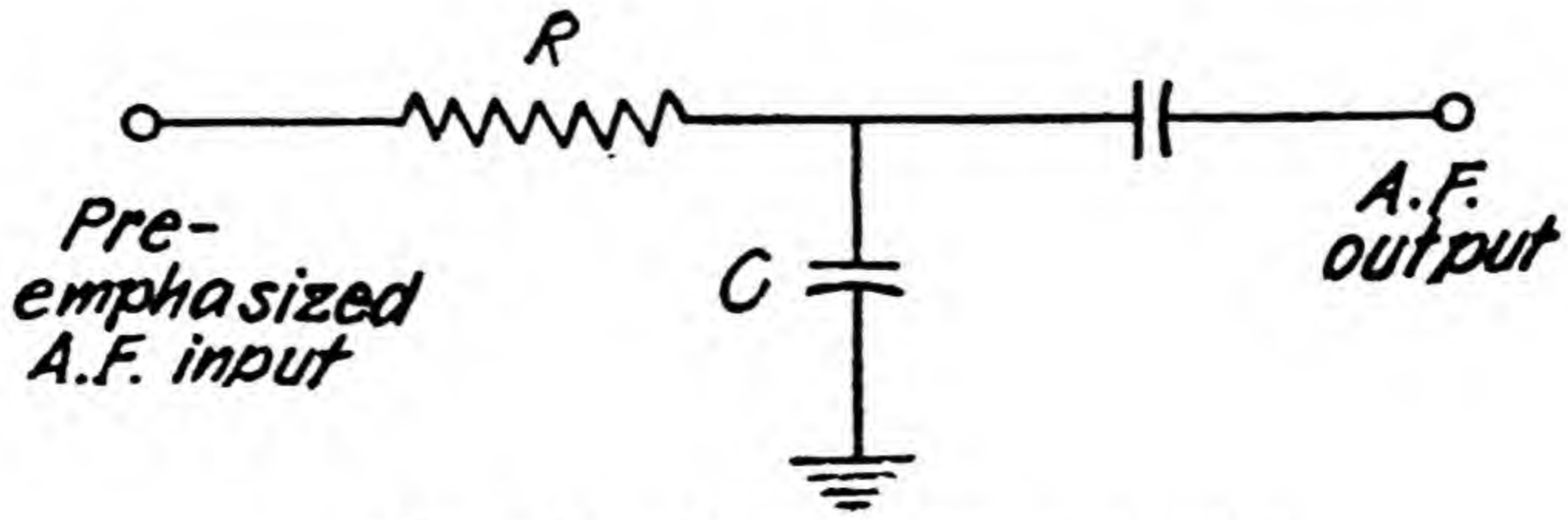


Fig. 38. De-emphasis network.

EXERCISES AND PROBLEMS

1. Describe briefly a modulated amplitude type of a broadcast transmitter
2. Draw the block diagram of a typical f. m. transmitter and explain the parts.
3. Describe phase modulated type f. m, transmitter. Discuss the advantages.
4. Give a circuit diagram with a brief description of the output stage of a radio telephone transmitter designed to work on shortwaves. The output of the transmitter is 0 KW, the carrier power, which is fed to a balanced transmission line of 500 ohms impedance. If the modulation is 100 percent, what is the maximum peak voltage across the output ?
5. What are the principal advantages and disadvantages of high level modu ation as compared with low-level modulation in a radio transmitter ?
6 Sketch pre-emphasis network and discuss why it is necessary to increase the amplitude of the higher modulation frequencies before modulation at the transmit er.
7. Give the block diagram of an amplitude modulated radio telephone transmitter.

   Describe how the modulation is obtained with the modulated plate class *C* amplifier.

   Describe the regative feed back amplitude modulated transmitter.
8. Draw block diagram of a typical F.M. transmitter using reactance tube and explain he parts
9 Select the correct answer :

(a) It is not desirable to modulate an oscillator directly because of the
   (*i*) large power required by the modulator ;
   (*ii*) non-linear output ;
   (*iii*) generation of sub-harmonics ;
   (*iv*) incidental frequency modulation. **Ans. (*iv*)**

(b) The highest frequency stability is achieved by using an oscillator of type
   (*i*) Hartley (*ii*) Colpitts
   (*iii*) Crystal controlled (*iv*) Electron coupled. **Ans. (*iii*)**

(c) Efficient r.f. amplification is achieved by using amplifiers of type
   (*i*) Class A (*ii*) Class B
   (*iii*) Class C (*iv*) Push pull. **Ans. (*iii*)**

(d) Chief advantages of pre-emphasis circuit in F.M. tran mitter is to
   (*i*) increase the carrier power ;
   (*ii*) improve the signal to noise ratio at low audio frequencies ;
   (*iii*) improve the signal to noise ratio at high audio frequencies ;
   (*iv*) increase the bandwidth of sidebands **Ans. (*iii*)**

(e) A frequency multiplier usually operates in
   (*i*) Class A (*ii*) Class B
   (*iii*) Class C (*iv*) Pushpull. **Ans. (*iii*)**

(f) Overmodulation of a transmitter signal is corrected by the adjustment of the
   (*i*) R. F. amplifier, (*ii*) power amplifier,
   (*iii*) speech amplifier. (*iv*) frequency doubler. **Ans. (*iii*)**

10. Explain the principle of operation of a superhetrodyne receiver for the reception of amplitude modulated signals.
Describe how the sensitivity and fidelity of a radio receiver are determined.
11. Discuss the factors affecting the choice of a proper intermediate frequency in a superheterodyne receiver.
12. In what respects does the r.f. receiver differ from a. m. receiver Describe front end portion of f.m. receiver.
13. Explain the principles and indicate the advantages of superheterodyne radio receiver compared with the straight (t.r.f.) receiver.
Explain the following terms :
(a) Delayed automatic gain control, (b) Signal frequency tracking,
(c) Cross modulation.
14. Give the circuit diagram of a superheterodyne radio receiver. Explain the process of frequency conversion by pentagrid converter. Discuss the choice of intermediate frequency and the function of a v.c. line.
15. A superheterodyne receiver is required to cover the waveband 200—500 metres and the intermediate frequency is to be 4·5 kc/s. calculate
(a) the maximum and minimum frequencies of the beating oscillator
(b) the maximum capacitance of the oscillator tuning capacitance assuming that the minimum value of 100 pf.

[**Hint.** (*a*) $f_{min}=600+450=1050$ Kc/s.

$f_{max}=1500+450=1950$ Kc/s.

(*b*) $f=\frac{1}{2\pi\sqrt{(LC)}}$, $f_{max}=\frac{1}{2\pi\sqrt{(LC_{min})}}$

and $f_{min}=\frac{1}{2\pi\sqrt{(LC_{max})}}$

so that $\frac{f^2_{min}}{f^2_{max}}=\frac{C_{min}}{C_{max}}$

or $\left(\frac{1050}{1950}\right)^2=\frac{10\times100^{-12}}{C_{max}}$]

16. Indicate the false statement :
(*a*) Straight receiver had the following drawbacks :
(*i*) gain variation over the frequency coverage range,
(*ii*) insufficient gain and sensitivity,
iii) inadequate selectivity at high frequencies.
(*iv*) instability ; **Ans.** (*ii*)
(*b*) If IF is very high :
(*i*) image frequency rejection is very good,
(*ii*) the local oscillator need not be extremely stable,
iii) the selectivity will be improved,
(*iv*) tracking will be improved. **Ans.** (*iv*)
(*c*) Main functions of R F. amplifier in a superhet set is to
(*i*) provide improved tracking,
(*ii*) permit better adjacent-channel rejection,

(*iii*) increase the tuning range of the receiver,
(*iv*) improve the rejection of the image frequency.

Ans. (*i*, *ii* and *iii*.

17. Discuss the demodulation of single sideband system. What will happen if the carrier at the transmitting end and that re-inserted at receiving end are not in perfect synchronism. Give the block diagram of *SSB* receiver.
18. Give and explain, sketching the ideal and typical curves, the physical content of, the definition of sensitivity, selectivity, and fidelity of a radio receiver. How is the sensitivity affected if a receiver is provided with an A.G.C. ? Describe a method to measure any one of the above mentioned characteristics.
19. Explain the purpose of pre-emphasis, de-emphasis and the limiter and describe schematically how one of these is incorporated in FM transmission and reception.

33 INTRODUCTION TO RADAR AND TELEVISION

RADAR

The term '*Radar*' is the abbreviation for '*radio detection and ranging*'. Radar is not a single instrument but an electronic device which includes various techniques employed for the purpose of detection and location of objects at distances that far exceed the range of human vision. The function of radar system is very much precise, rapid and is unaffected by darkness, clouds or fog which are impediments to vision.

33·0. PRINCIPLE OF RADAR :

Radar is a self contained unit including a powerful radio transmitter and sensitive receiver. Powerful pulses of very short duration are set out in space by the radar transmitter. This radio energy is bounced back to the sender system by a distant target which should be capable of scattering electro-magnetic wave in space. The time thus taken in two way journey is a measure of the range of the target.

These pulses are sent at regular intervals. For the measurement of long distance, a time interval of $\frac{1}{250}$ sec, and for short distances that of $\frac{1}{50000}$ sec is usually adopted. To send such pulses, the radar transmitter comes in operation for a duration of micro-second and then remains inactive for the relatively long time of $\frac{1}{250}$ seconds. Thus the pulses are transmitted with a definite frequency called the pulse repetition frequency.

33·1. BASIC ARRANGEMENT OF RADAR SYSTEM :

A block diagram of a basic pulse type radar system is shown in fig. 1. For a simple description, its function can be stated as follows :

(*i*) Large video pulses with a repetition frequency of several undred per second are supplied by the modulator micro-wave hcillator.

(*ii*) The resulting r.f. pulses are radiated by a highly directional antenna to confine the radiated r.f. power in a definite range and direction.

(*iii*) Since the same antenna is used for transmission and reception purposes, a duplexer is used to *isolate the receiving unit from strong transmitting pulses at the time of transmission* of radio

energy to avoid its damage, and then to switch the antenna to the receiver in the intervals between the pulses during which the reflected energy is being received.

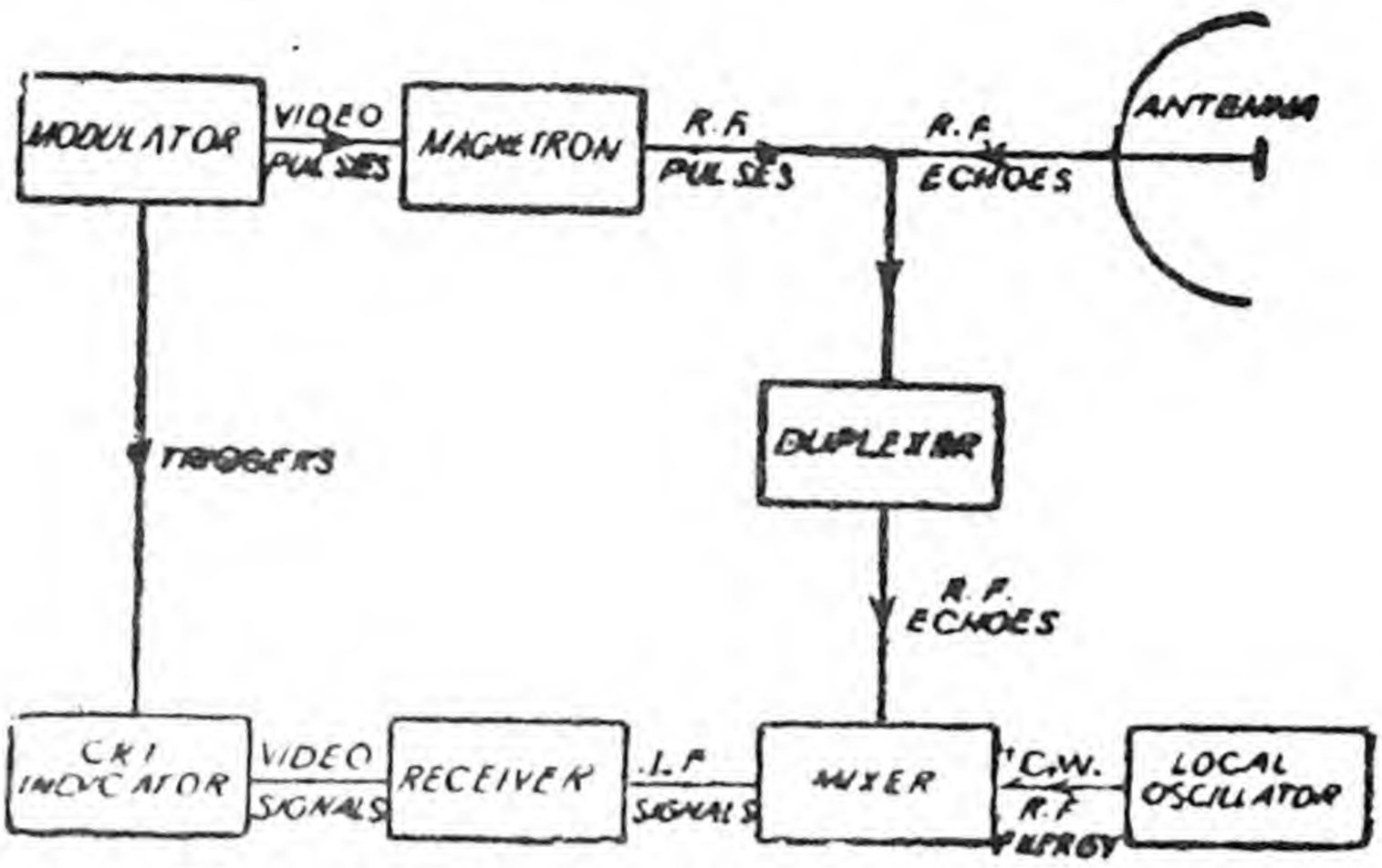


Fig. 1. Block diagram of a simpie pulse transmitter.

(*iv*) The received r.f. pulses are fed to mixer which converts them to I.F. pulses which are then amplified and detected to give video signals.

(*v*) The signals are then fed to the indicator unit. It is essential to synchronise exactly the indicator-sweeps with the transmitted pulses by means of triggers obtained from modulator.

(*vi*) In searching the surrounding space for targets, the directional antenna is given a rotatory motion so as to point the energy beam in different directions. The range and direction are then noted by the indicator unit such as P.P.I. which is properly synchronised with antenna position.

As shown in fig. 2, the direction of target is specified by two angles. One *azimuth* ϕ which is the direction of the projection of the radius vector on the horizontal from a fixed reference direction

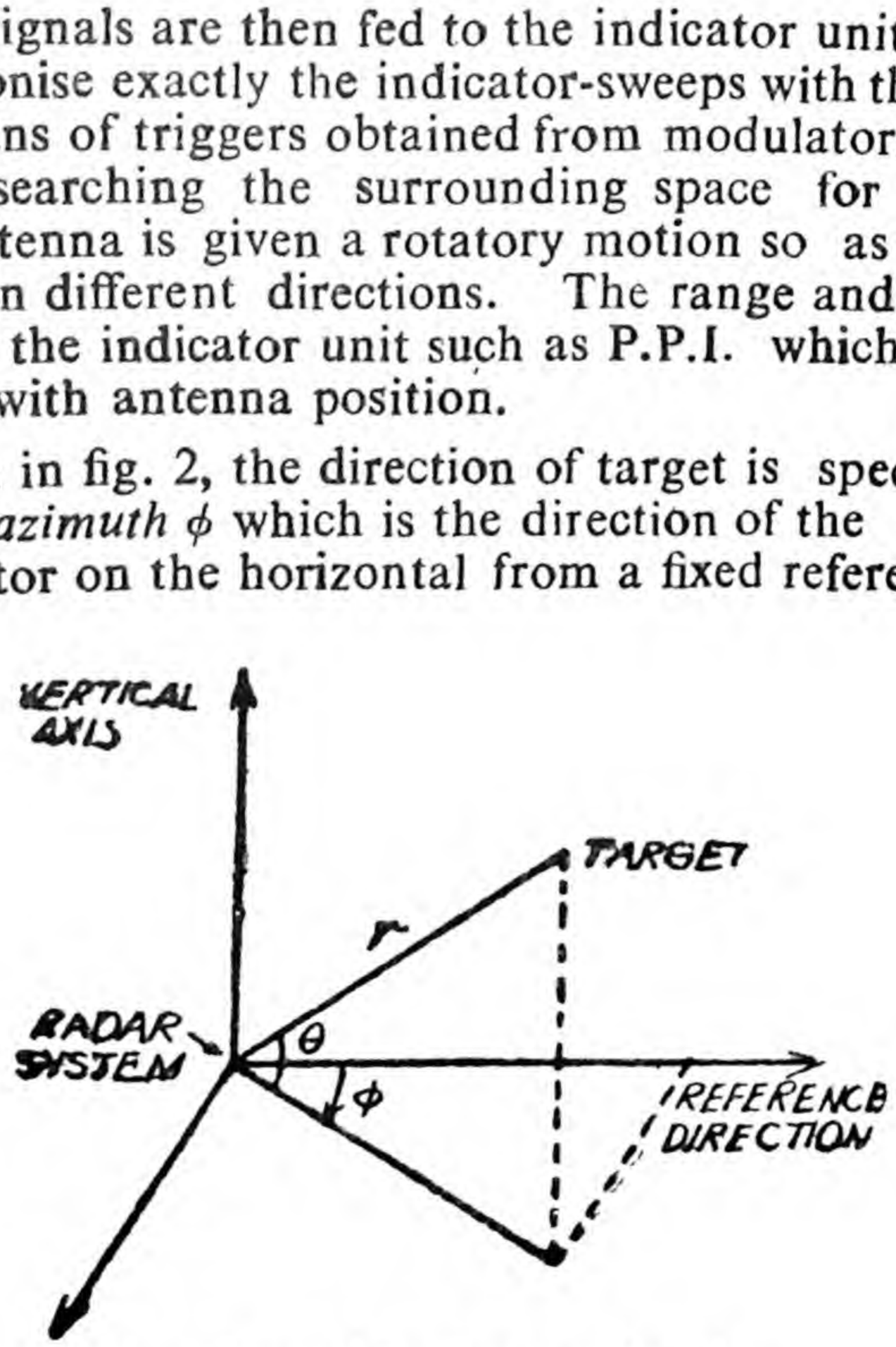


Fig. 2. Target direction specification.

(usually taken as north) and the second *elevation* θ which is the angle between this projection and radius vector *r* from radar to the target.

33·2. AZIMUTH AND RANGE MEASUREMENT :

The function of indicator : Indicators furnish the information about the presence and position of the target. Cathode ray tube is used as an indicator in pulse radars whose beam is modulated in intensity to show the presence of target and distinguish it from other targets and deflected in position to indicate the angular position (azimuth and elevation) of the target. All these arrangements lead to class *C* indicator. We are mentioning here, in brief, the types of indicator which are commonly employed.

Type A indicator : In this type of indicator, this beam of the cathode ray tube is deflected across the screen by the application of linear saw tooth voltage, synchronised with transmitted pulses, to the horizontal plates. The output of the receiver is applied to the vertical deflection plates. Thus the target signal may deflect the beam upward or downward. The upward deflection is more commonly employed. The distance of the each pulse from the transmitter pulse on the left hand side, which sets *zero of the linear scale*, gives the range of the target as shown in figure 3. This type of indicator only reveals the presence and gives the range of the target. It does not give any information about the angular position of the target. The greater advantage of such an indicator lies in revealing the nature of the target by the strength of the received signal.

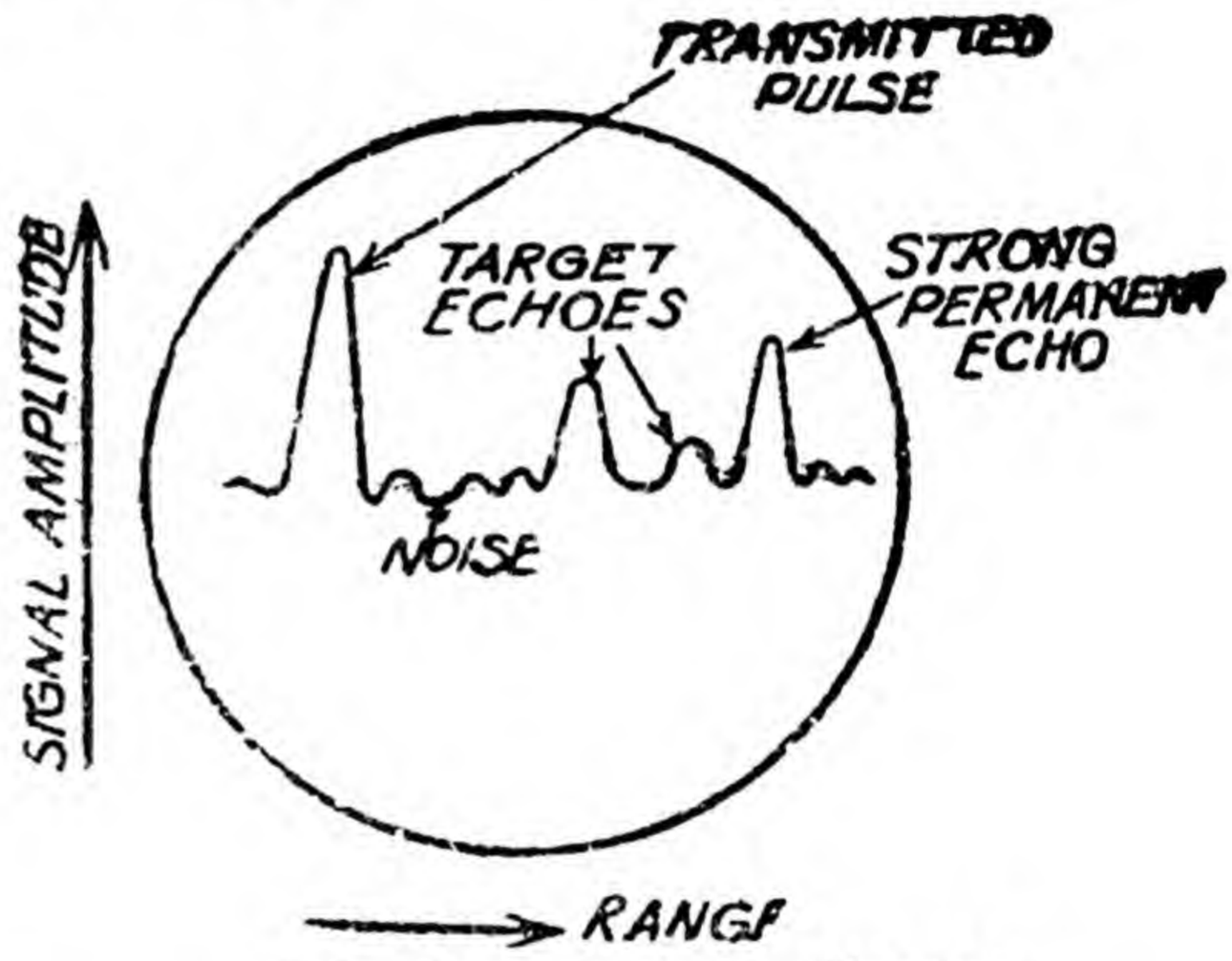


Fig. 3. Type *A* indicator.

Type B Indicator. In this type of indicator, the cathode ray beam is modulated in intensity to indicate the target. This position of the resulting spot of light shows the range and azimuth of the target in rectanguar co-ordinates as shown in fig. 4. The range is displayed vertically and azimuth horizontally. Since the scanner searches for target in range and azimuth only this type of indicator is usually suited to radars employed for circular scanning over a limited sector.

For such type of display, the beam is deflected vertically with linear sweep, starting at the bottom at the instant each pulse is transmitted. The reflected signal intensifies the spot if it is much above the noise. The horizontal deflection is obtained by feeding

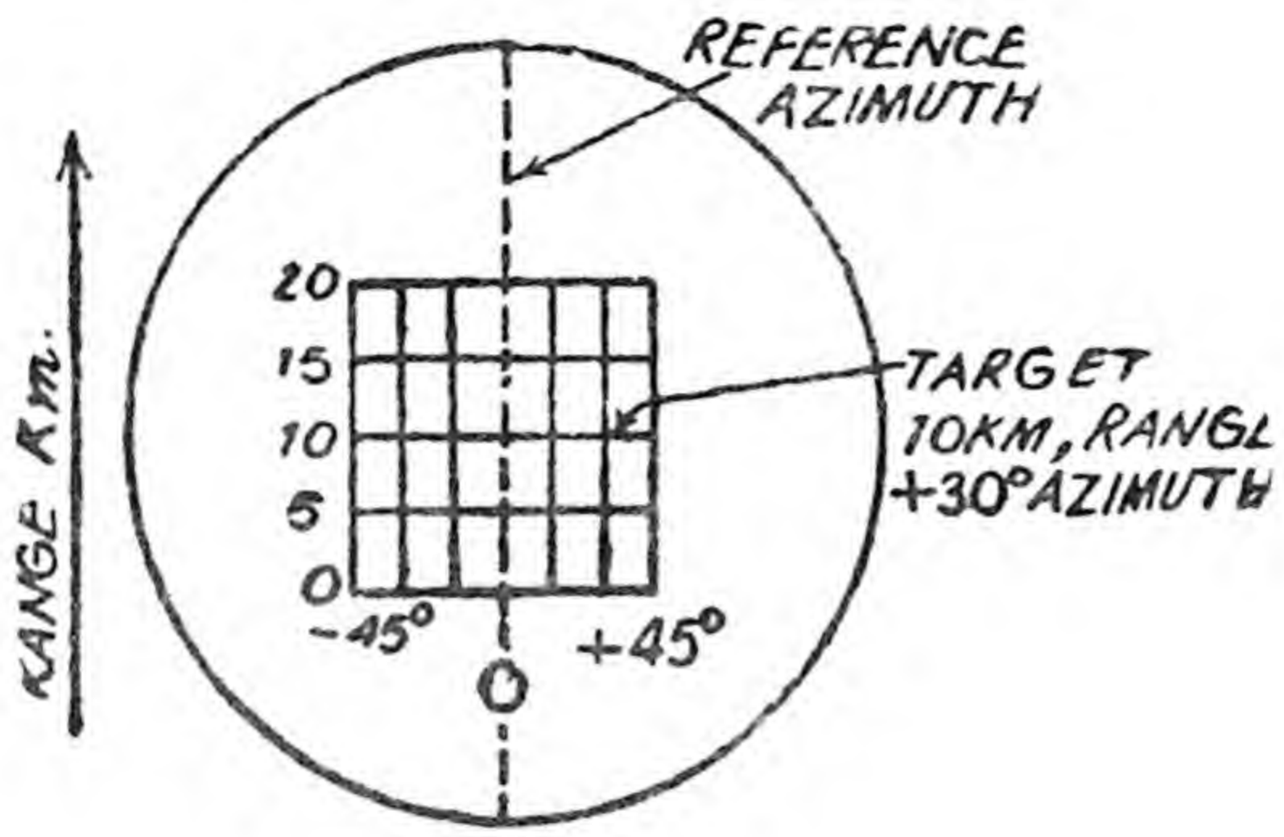


Fig. 4. Type *B* indicator.

the current proportional to the angular rotation in azimuth to the horizontal deflection coil from a potentiometer mounted on the scanner. Thus this indicator gives a distorted picture of space currents—projected on a horizontal plane before the radar.

Type C Indicator : This indicator presents the target by intensity modulation and indicates the angular position (azimuth and elevation) of the target in rectangular co-ordinates. It does not provide range measurements and should, therefore, be used with type *A* or type *B* indicator to furnish target information in three co-ordinate systems.

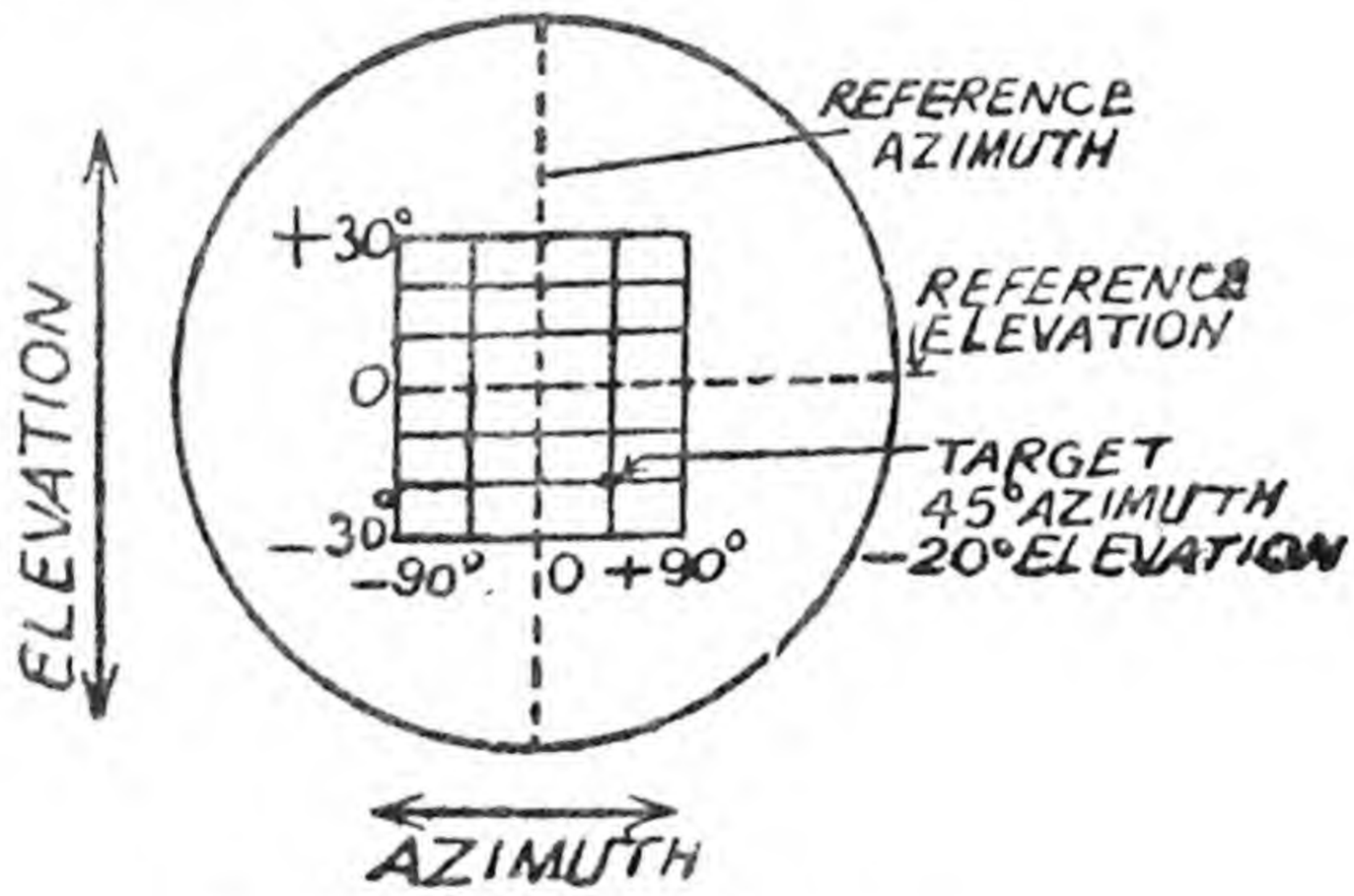


Fig. 5. Type *C* Indicator.

Plan Position Indicator (P. P. I.) : This type of indicator closely resembles with type *B* indicator which presents the target by intensity modulation of cathode ray beam. P. P. I. gives a map-like representation throughout 360° or any particular sector desired, without distortion. In this indicator, a sawtooth timing wave defiects the spot on the screen of cathode ray tube radially outwards from the centre. This saw tooth wave is so synchronised with transmitted pulse that the distance of the spot from the centre gives the target distance. Th eangular deflection of the spot-

caused by the saw tooth timing wave corresponds to the direction in which the antenna beam is directed at that moment. The control

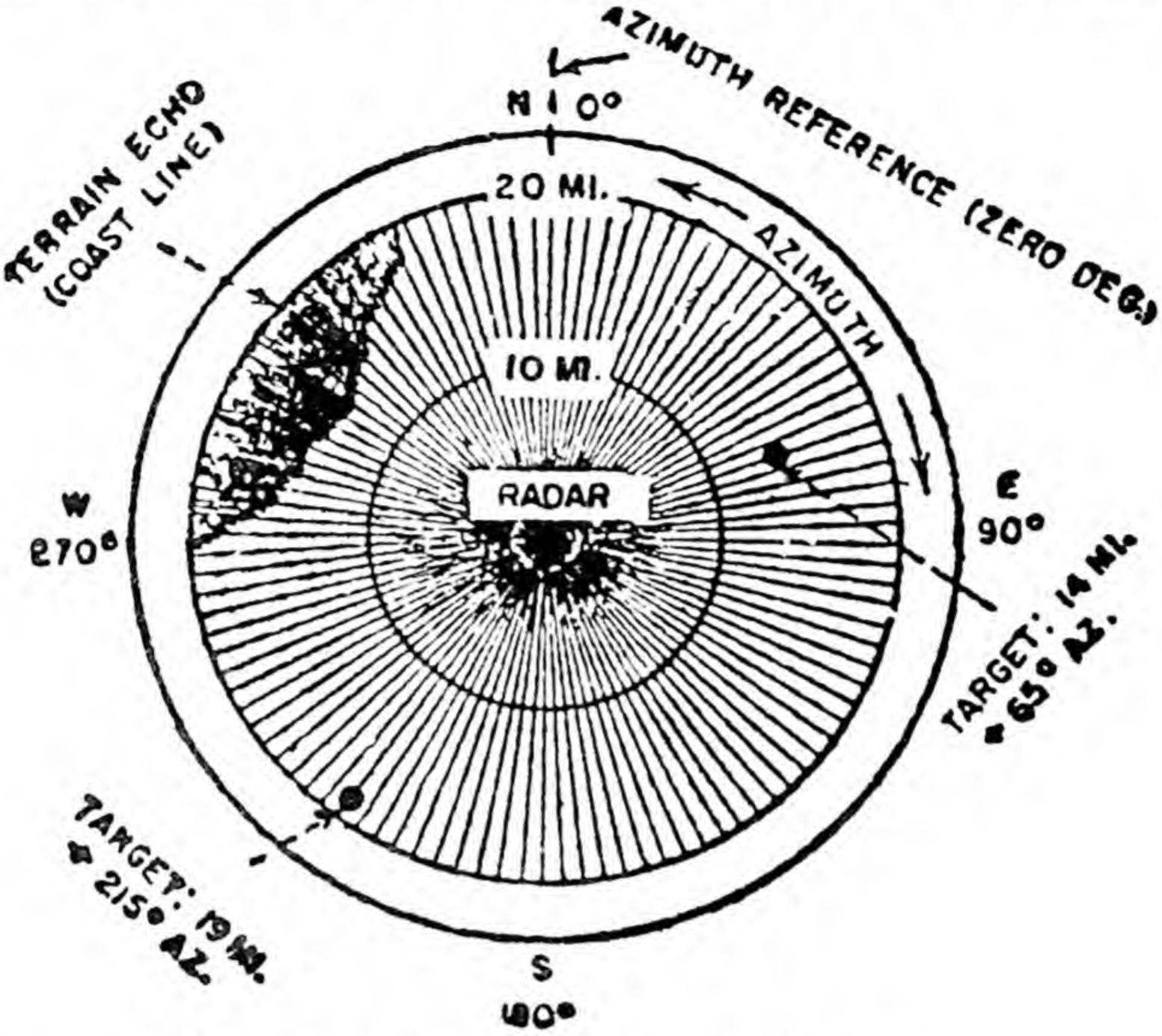


Fig. 6. P.P.I. Representation.

electrode of cathode ray tube is biased just greater than cut-off, so that when echoes are applied to this electrode, the spot is intensified giving range and azimuth on a map. Magnetic deflecting coils that are rotated synchronously with antenna are employed to achieve the radial deflection. A P.P.I. presentation is shown in fig. 6.

33·3. OPERATING CHARACTERISTICS OF RADAR SYSTEMS :

The performance of a radar system depends upon the factors such as the desired operating range, the accuracy of the acquired data and nature of the targets; and these factors, on the other hand, depend upon the *choice of operating frequency, pulse duration and repetition rate, the power output of the transmitter, choice of indicator type, sensitivity and band-width of the receiver, and radiation pattern of antenna.*

Choice of Operating Frequency :

Most pulse radar systems operate at frequencies from above 1000 mc/s. to about 70,000 mc/s The use of such high frequencies has the advantages that ;

(*i*) at such high frequencies a sharp and well confined radiation beam can be achieved with an antenna structure of relatively small physical size,

(*ii*) higher the frequency, shorter may be the transmitted pulses which provide good range resolution.

The use of high frequencies also suffers with the following main disadvantages ;

(*i*) noise figure increases more with frequency.

(*ii*) power generated in the transmitter tends to be less as the frequency is increased :

Pulse Duration :

For good resolving capabilities, short duration pulses in the form a narrow radiating beam should be employed. For a good range resolution, pulse durations of 1 μ-sec or less are employed. Pulse width used, in radar, range from about 0·2 μ-sec. long to values of the order of 30 μ-sec.

Pulse Repetition Frequency :

The pulse repetition frequency is made sufficiently small, so that the time spent in the path between radar and distant target may not exceed the interval between transmitted pulses. Pulse repetition frequencies used in radar work usually vary from 350 to about 10,000 cycles. Short range radars use high pulse repetition rates.

Transmitted power output :

To detect targets at range upto several hundred miles transmitter peak power output of more than one megawatt may be required. For a particular radar transmitted power is determined by the maximum distance over which it is desired to receive target information.

A typical radar used for the detection of conventional aircraft at range of 100 or 200 miles might employ a peak power of the order of $1MW$, a pulsewidth of several micro-seconds, and a pulse repetition frequency of several hundred pulses per second.

Maximum range :

It depends upon the energy content of the transmitted pulses and the sensitivity of receiver. For better sensitivity random noise generated in the input section of the receiver should be minimised.

33·4. MAXIMUM RANGE OF A RADAR SET :

We shall obtain an expression for the range of a radar set in terms of the factors on which it is expected to depend. Let us denote :

R = distance to a target,

S = equivalent area of target,

P_t = power radiated from transmitter,

P_r = echo power absorbed by receiving antenna,

G_t = power gain of transmitting antenna relative to isotropic radiator,

G_r = power gain of receiving antenna relative to isotropic radiator,

λ=wavelength,

A_0=capture area of receiving antenna,

A=aperture area of receiving antenna,

$\frac{A_0}{A}=k$ (a constant having value 0·5 per parabolic antennas and 0·9 for mattress antennas).

The surface area of sphere of radius $R=4\pi R^2$.

Transmitter power in beam direction $=P_tG_t$.

Therefore, the power per unit area of the target

$$=\frac{P_tG_t}{4\pi R^2} \qquad ...(1)$$

This energy strikes the target and a part of it is scattered in the direction of the radar antenna. We assume that the strength of echo received is equivalent to the strength of wave received by this antenna if it were radiated by a section of wavefront of area S from an isotropic radiator located on the target *i.e.* the amount of echo power received by the antenna is equivalent to the power scattered by an effective area S of the target.

$\therefore$ Echo power per unit area at the receiver

$$\frac{P_tG_t}{4\pi R^2}\times S\times\frac{1}{4\pi R^2}. \qquad ...(2)$$

Therefore, the total power delivered to the receiving system or receiving power,

P_r=echo power per unit area $\times$ capture area

$$=\frac{P_tG_t}{(4\pi R^2)^2}.S.A_0. \qquad ...(3)$$

Now gain of receiving antenna

$$*G_r=\frac{4\pi A_0}{\lambda^2}$$

$$=\frac{4\pi kA}{\lambda^2} \qquad \left(\text{since } k=\frac{A_0}{A}\right). \qquad ...(4)$$

since the same antenna is used for reception and transmission,

$$G_r=G_t.$$

Therefore, the equation (3) becomes

$$P_r=\frac{P_tG_r}{(4\pi R^2)^2}.SA_0.$$

Substituting the values of G_r and A_0, we get

$$P_r=\frac{P_t.\ \frac{4\pi kA}{\lambda^2}\ .SkA}{(4\pi R^2)^2}$$

$$=\frac{P_t.k^2.A^2.S}{4\pi R^4.\lambda^2} \qquad ...(5)$$

**Electromagnetic Waves and Radiating Systems* by E.C. Jordon.

If $(P_r)_{min.}$ denotes the minimum energy, that can be detected by the receiver, then the maximum value of R—the so called range the radar—will be

$$R_{max.} = \left[\frac{P_t\, k^2\, A^2.S}{4\pi\lambda^2\,(P_r)_{min.}}\right]^{1/4} \qquad \ldots(6)$$

The expression, therefore, shows that the maximum range for the minimum detectable energy is proportional to the fourth root of transmitter peak power, to the square root of the antenna area, and inversely to the square root of wavelength. The effective area, S, on which range also depends, is different for different targets, *e.g.* for very large bomber it is 150 sq. metres, for small fighter plane it is 10 sq. metres, $R_{max.}$ is of the order of several kilometres even for small airplanes.

33·5. RADAR TRANSMITTING SYSTEMS :

In radars operating upto 600 mc/s. power is usually drawn from triode tubes in a resonant line oscillator circuit. For higher operating frequency, cavity magnetron is suitable for generating high peak powers. High power klystrons are also being used as an alternative to magnetron. Modulators which are employed in radar transmitting circuits to generate pulse of appropriate length, amplitude and repetition rate for the application between the anode and cathode of the power tube of the radar, transmitter, are commonly of two types :

(*i*) Line pulser,
(*ii*) Hard tube pulser.

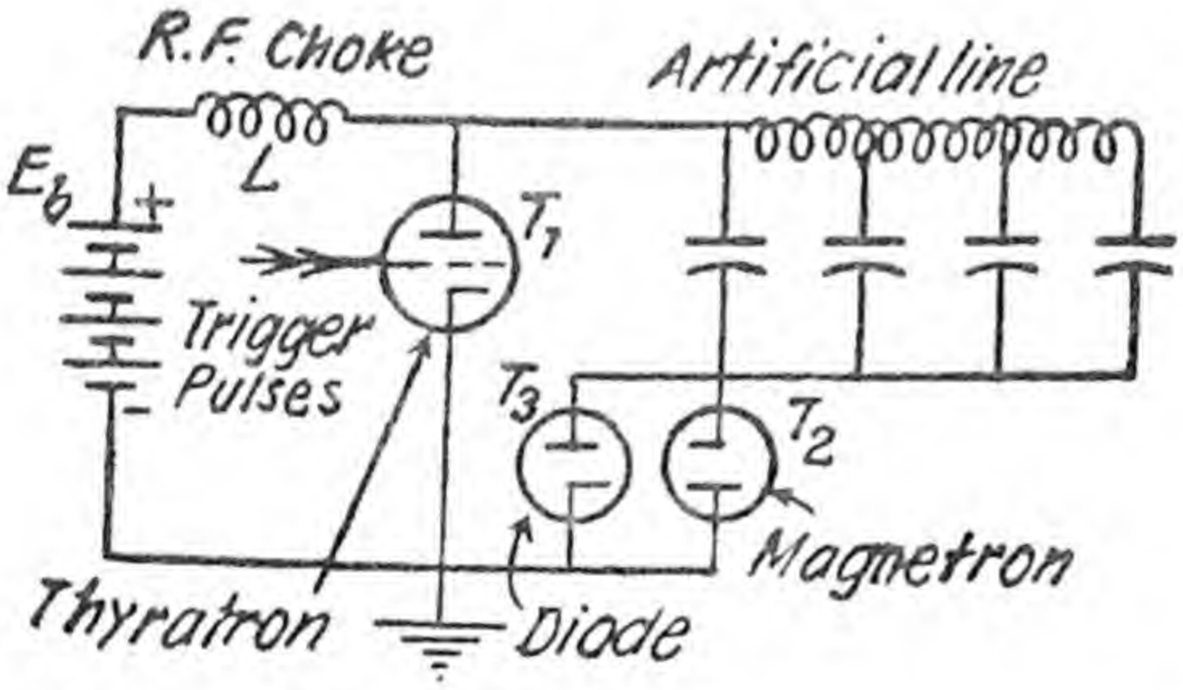


Fig. 7. (a) A line pulser.

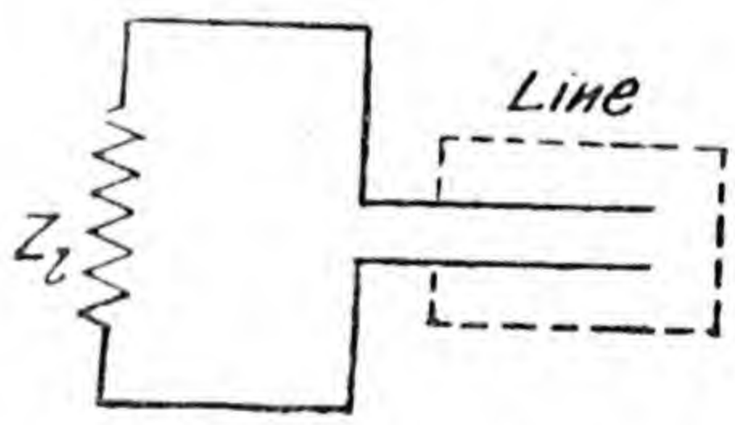


Fig. 7.(b) Equivalent circuit with switch closed.

Line pulser : A typical line pulser is shown in figure 7 (*a*). The thyraton functions as a trigger switch and the load impedance represented by magnetron equals the characteristic impedance of transmission line. Equivalent circuit is shown in figure 7 (*b*). During the interval between the pulses, transmission line is charged to a voltage E_b through inductance L and diode T_3 for the latter is necessary to permit the passage of charging current as magnetron does not conduct when its anode is negative.

By closing trigger switch, pulses are formed and thus the line is terminated with a load Z_1 equal to its characteristic impedance.

The energy stored in the line then discharges into the load impedance and during discharge period, voltage developed across impedance load is $E_b/2$. The duration of discharge is equal to the time taken under such conditions *i.e.* the trigger switch being ideal one—acting simultaneously and load resistance being equal to the characteristic impedance at all frequencies, the line pulser produces an ideal square pulse. Line pulses can be operated at a sufficiently high power level.

(ii) Hard Tube Pulser : Circuit is shown in the fig. 8 (*a*). Triode tube T_1 is biased beyond cut-off, so that during the interval between the pulses, this tube carries no current and capacitor C is

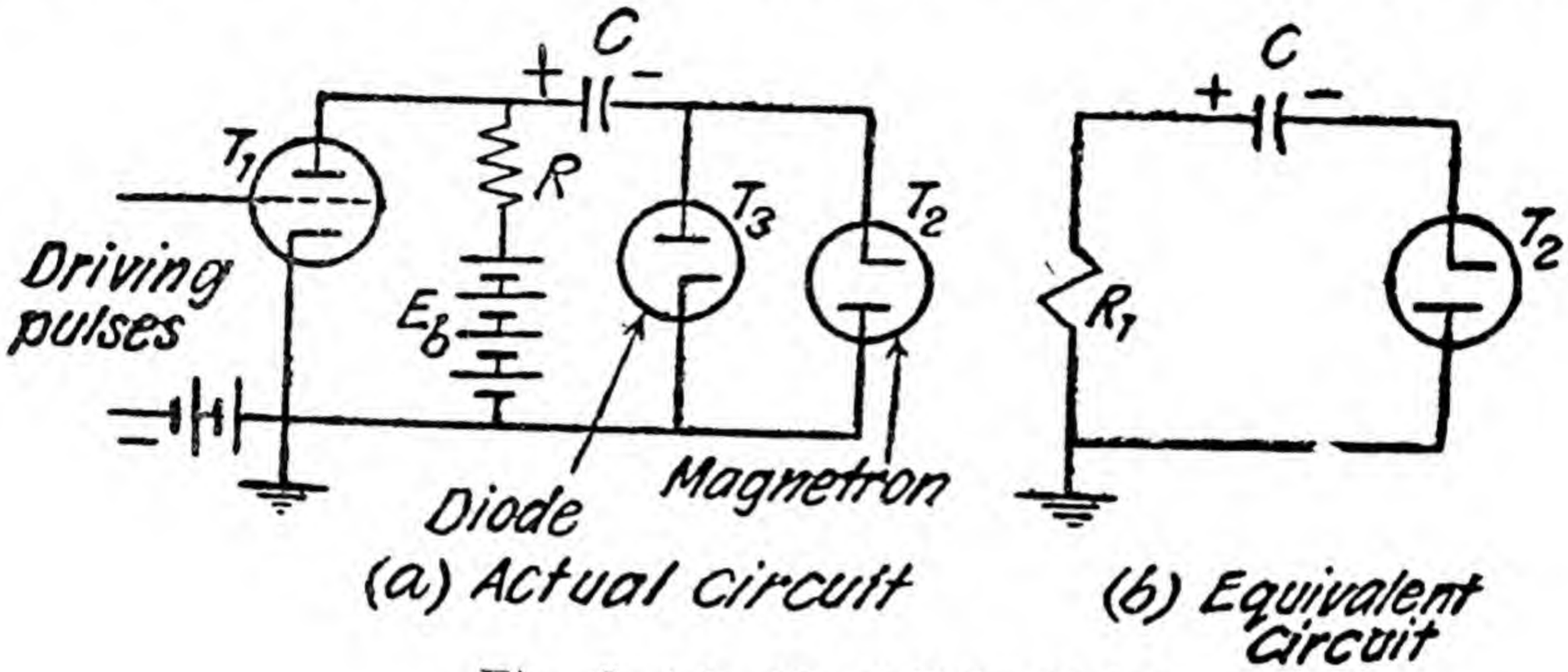


Fig. 8. (*a*) Hard Tube Pulser.
(*b*) Equivalent Circuit.

charged upto plate supply voltage E_b through very high resistance R and series diode T_3. Now some positive pulses are applied to the control grid of T_1 to drive the grid slightly positive due to which it offers very low impedance. The equivalent circuit of the system takes the form as shown in fig. 8 (*b*) during the period of pulse. In this case the condenser C, initially charged to plate supply voltage E_b is connected directly across the load impedance, *i.e.* across the oscillator, and thus applies full voltage E_b to the oscillator. The capacitor C should be sufficiently large so that during its discharge through impedance, the frequency and power output of the oscillator are not markedly changed by the potential across this capacitor. This type of pulser is more bulky and more expensive.

The pulse repetition frequency of the modulator is determined by a blocking oscillator or multivibrator. The pulses of required character are produced by exciting the multivibrator by a sine wave which has been differentiated and clipped properly.

33·6. RADAR ANTENNAS :

Directional antennas are always used in radar system to increase the range and better angular resolution.

The measure of sharpness and directivity of the radiated beam from the antenna is, however, dependent on its requirements. If the radar is being used simply to locate the position of the target

and not the height from horizontal, then a sharp beam is radiated in the horizontal direction to give high azimuth resolutions while a less sharp beam in the vertical direction. If, on the other hand, height of the target is an important factor, then for an accurate determination of the angle of elevation, an intense narrow beam is sent in the vertical plane. The previous is named as 'ground based' air-search radars' type while the latter as 'height finder radars'. Air borne radar (used on aeroplane) is used for navigational aid, the direction pattern of which is shown in figure 9. The field strength of such an antenna pattern is proportional to the cosec of angle θ except when θ approaches 0 and 90° and power radiated is proportional to $\text{cosec}^2\,\theta$. Thus the strength of echoes received from a particular target at a given altitude becomes independent of distance over the range of vertical angle θ.

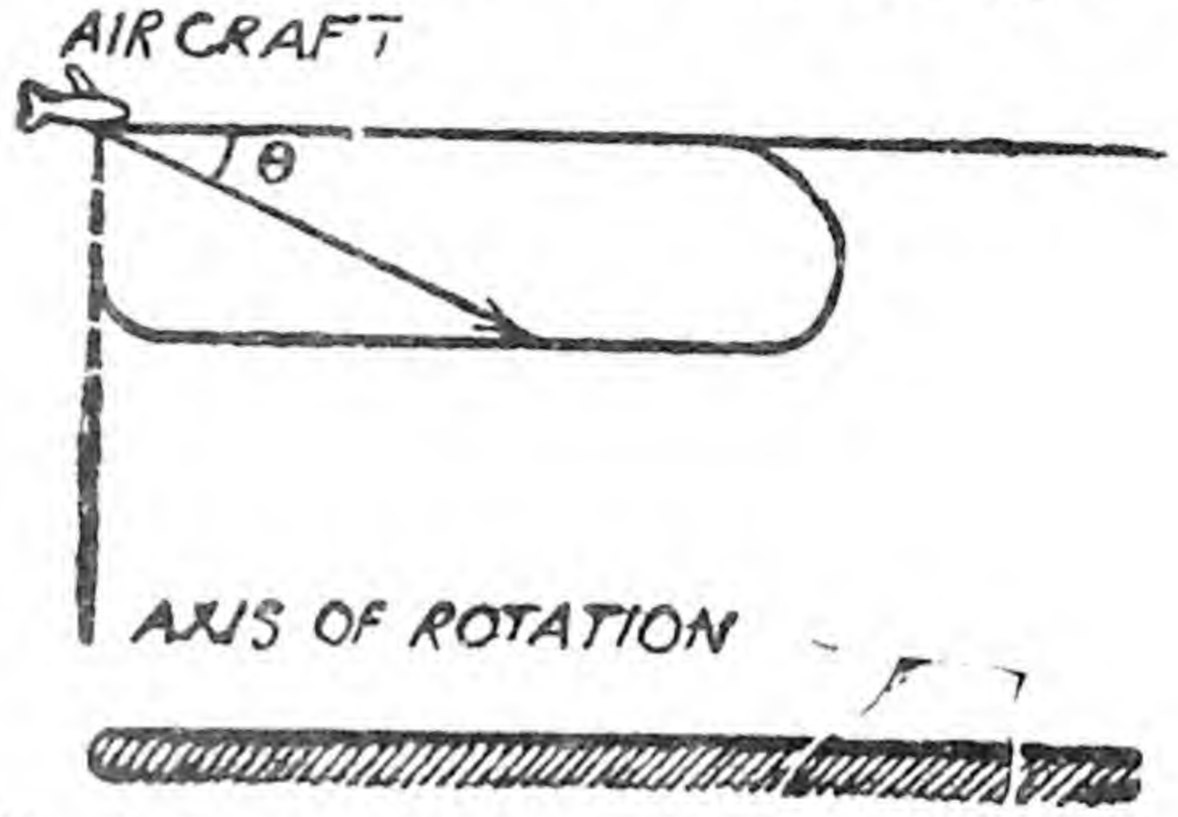


Fig. 9. Constant directional pattern.

The directivity is also a function of frequency used in radar because at low frequencies of the radar of 100 to 200 mc/s. a sharp narrow beam is difficult to be produced even with antenna of large physical size, while at frequencies 3000 mc/s. an antenna of smaller size can serve the better directivity purposes. Ground radars, operating in frequency range 100 to 200 mc/s., normally employ *mattress antennas* with reflectors. In air borne radars, *Yagi antennas* are usually used in the range 200 to 500 mc/s. while at frequencies of 3000 mc/s. and above *parabolic reflectors* are standard. To avoid confusion, it is necessary that the direction pattern of the radar antenna should be free from secondary lobes. Techniques are, therefore, employed for minimizing minor lobes.

In all types of radar, a mechanical motion of the antenna is accomplished to vary the direction of beam radiated from it. The process is called *antenna scanning*. In radars used for navigational

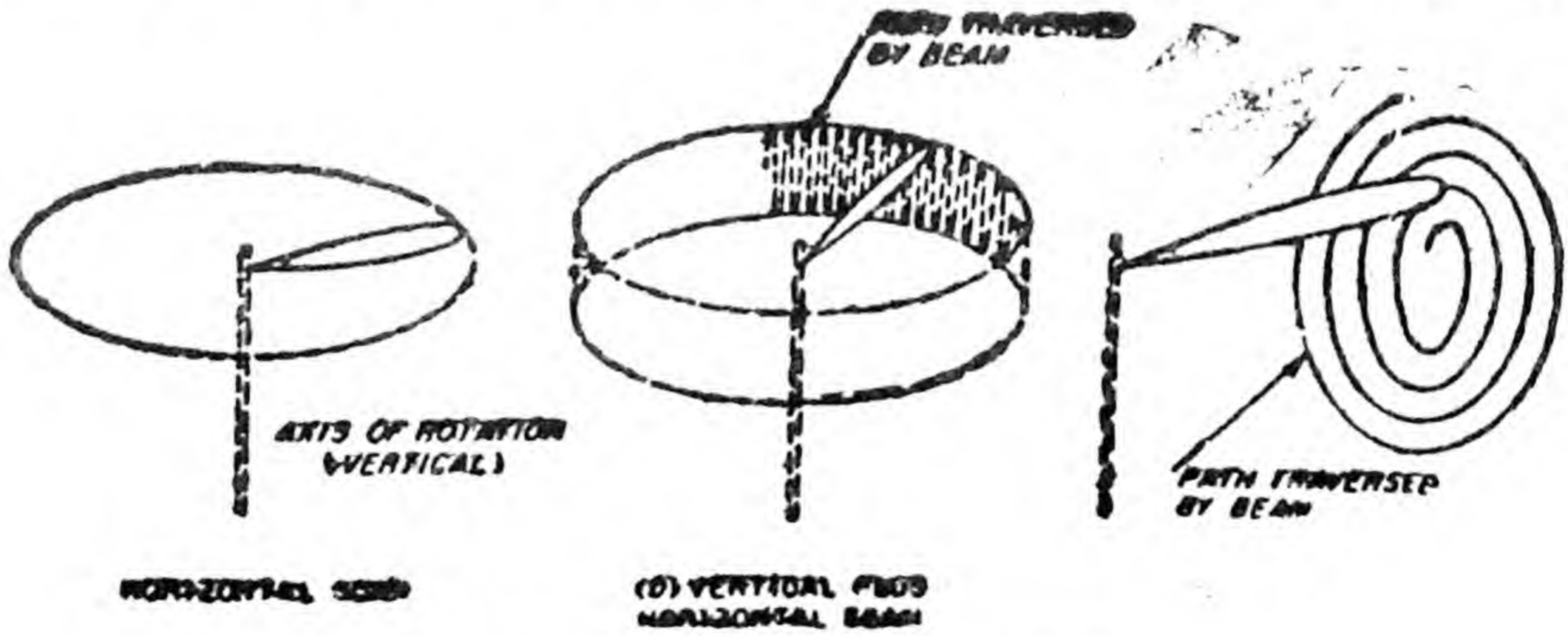


Fig. 10. Scanning method's representation.

purposes, scanning is performed by rotating the antenna about a vertical axis. In height-finder radars, scanning consists of a rapid rocker motion in the vertical plane with much lower azimuth scan about a vertical axis.

Lobe switching : The angular direction of a radar target can be determined accurately by a process called lobe switching. In lobe switching, direction pattern of the antenna is switched rapidly in the direction *OA* and *OB* (fig. 11). A radar target in the direction of α_0 will then give equal echo signal for these two lobe positions, while any target in the direction of α_1 or α_2 will give echos of unequal strength for the two lobe positions. Thus the direction α_0, which gives echo of equal strength for the two lobe positions, is sharply defined and hence the angular position of the target is precisely determined.

In mattress antennas, the lobes are switched in two directions by connecting the transmission line alternately to one end and to other end of the antenna system by a mechanical switch as shown in fig. (11 (*a*).

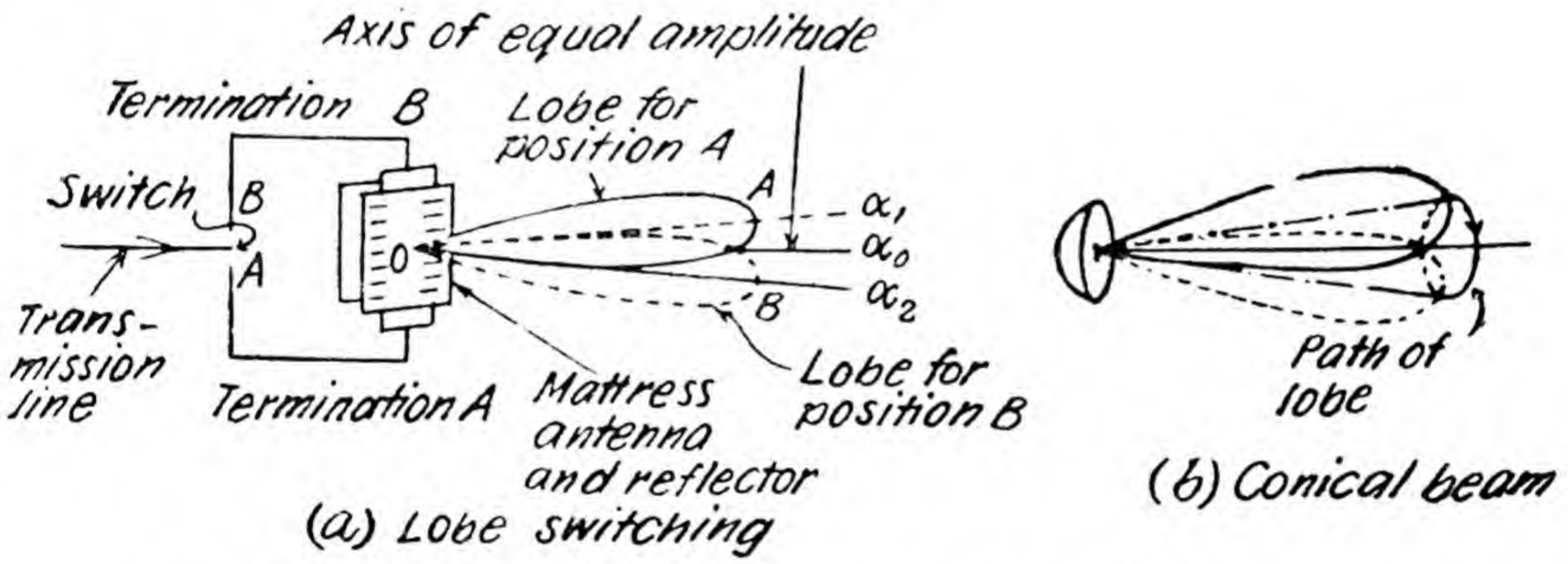


Fig. 11. (a) Lobe switching in mattress antenna,
(b) Lobe switching in parabolic antennas.

In parabolic antennas, the process of lobe switching is accomplished by placing the exciting antenna slightly off centre with respect to the focal point and then rotating it about the axis of parabola at a rate which is less than pulse repetition frequency. This arrangement is termed as conical scanning.

Effect of Ground on Radar Antenna Patterns : When the antenna is located near the ground, the directivity of its vertical patterns is largely affected. Fig. 12 (*a*) represents the directional pattern of an antenna remote from ground. For a grounded antenna of 7·5 metres height above the ground and operating at 10 mc/s., directional pattern assumes the shape as shown in figure 12 (*b*). If the operating frequency is increased, *i.e.* wavelength is decreased keeping antenna height constant, then the angle at which the first lobe appears from the horizon is lowered as shown in fig. 12 (*c*) and thus improves low angle coverage. If the antenna is tilted upwards, a little energy would come in contact with ground and thus eliminating the ground effect, but the low vertical angle coverage

becomes poor. Pattern is shown in figure 12 (*d*).

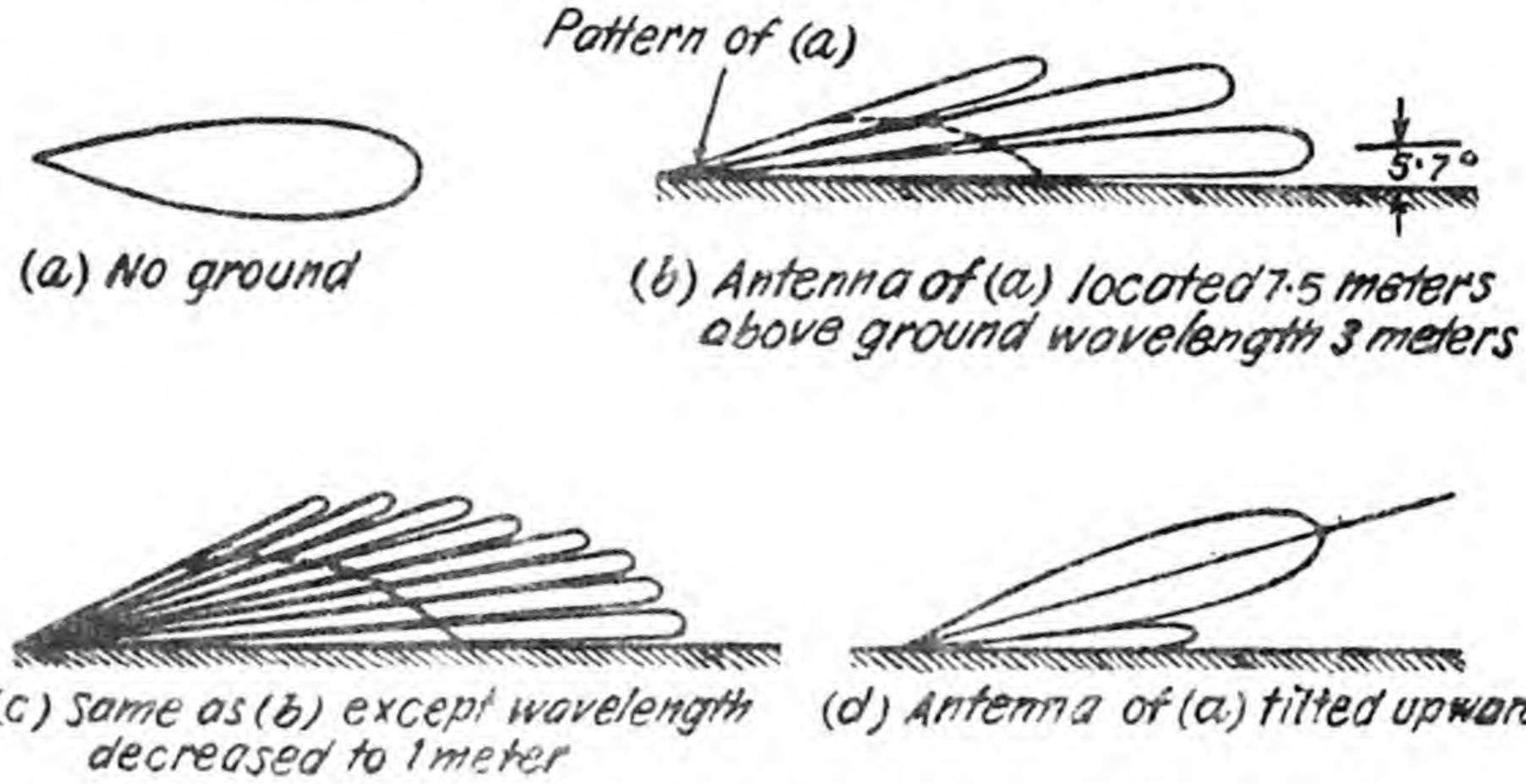


Fig. 12. Effect of ground on directional characteristics of antenna.

33·7. DUPLEXER :

Generally for transmission and reception, the same antenna is used in all radars. This is accomplished with the help of a Duplexer. Duplexer is the device that (1) switches the radar antenna to either the transmitter or the receiver, and (2) serves to protect the receiver from burnout or damage during transmission. It makes use of fast acting r.f. switches called as "T.R. (transmist receive) Boxes" and 'A.T.R. (antitransmit receive) Boxes'. 'A.T.R. Box' is a 'T.R. Box' without output lead.

'T.R. Box' consists of resonant cavities associated with special cold cathode tube containing gas at low pressure as shown in fig. 13. Transmitter pulse causes ionisation across the gap due to which resonance in 'T.R. Box' is disturbed and thus a little transmission of energy from input to the output line occurs. Under such a condition, the box acts as short-circuited and thereby isolating receiver when transmitter is operating. Layout of a Duplexer is shown in fig. 14 (*a*). Line sections *AB* and *CD* are made odd *multiples of* $\lambda/4$ *long*.

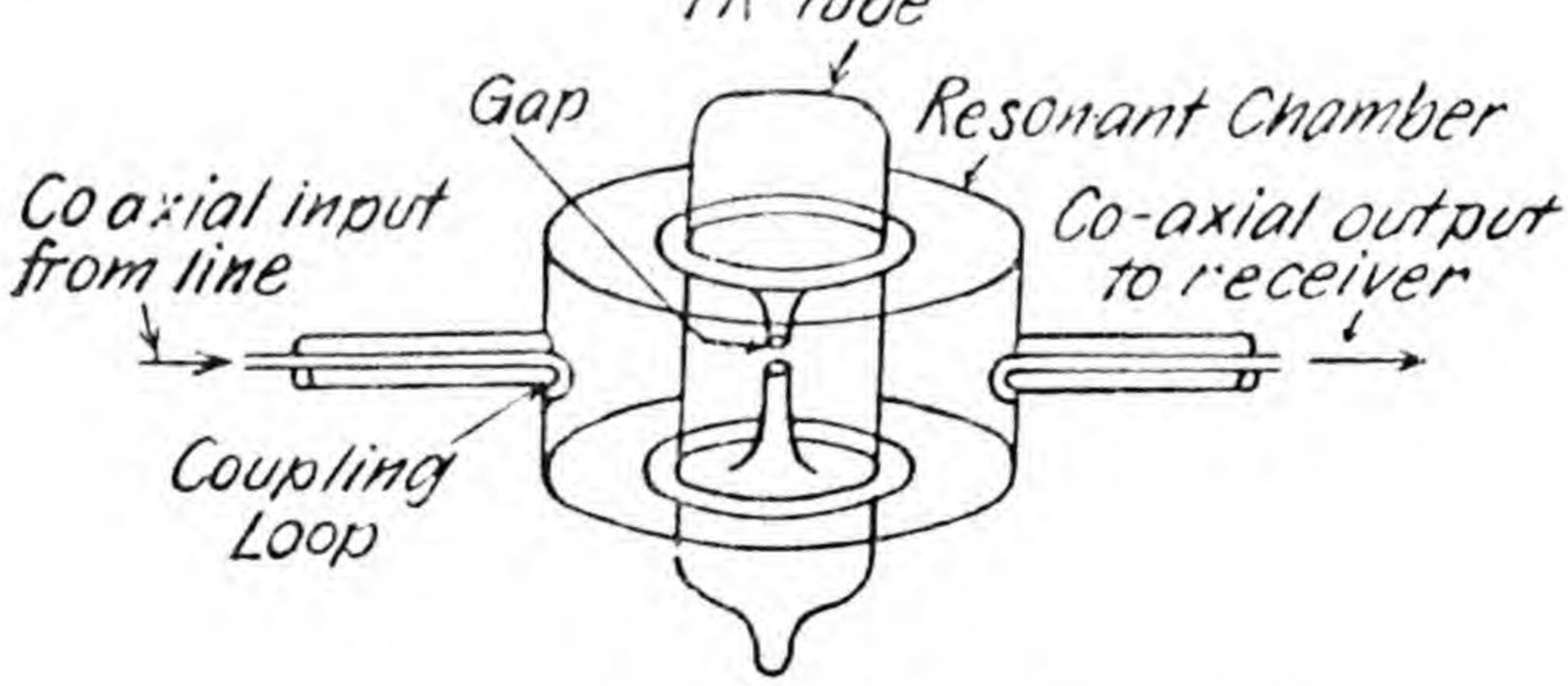
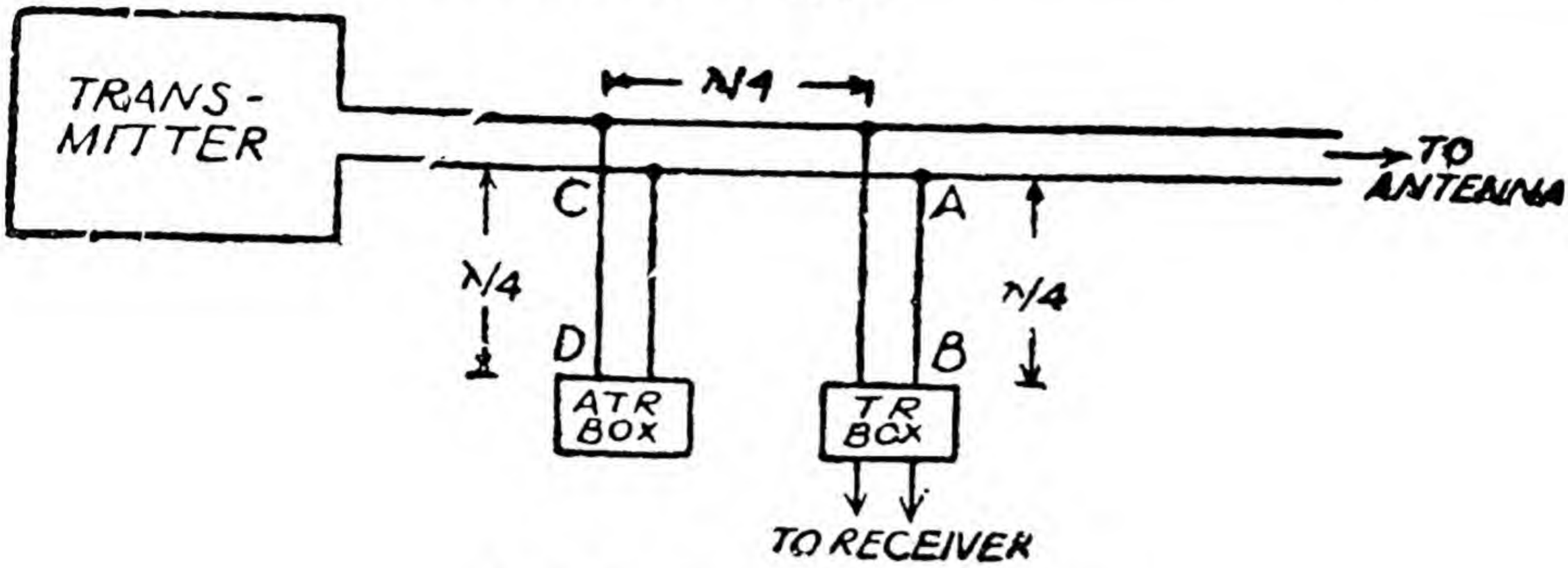


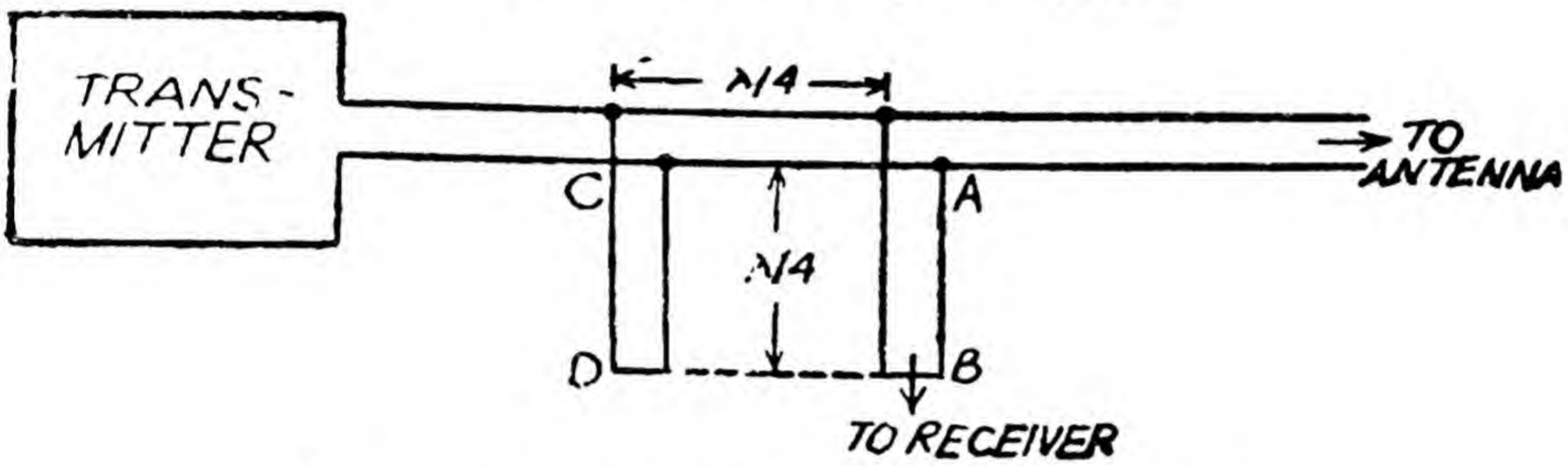
Fig. 13. Duplexer.

Consider when transmitter is operating, A.T.R and T.R. boxes both are *short-circuited* and thus offer high shunt impedance to the line carrying energy from the transmitter to the antenna.

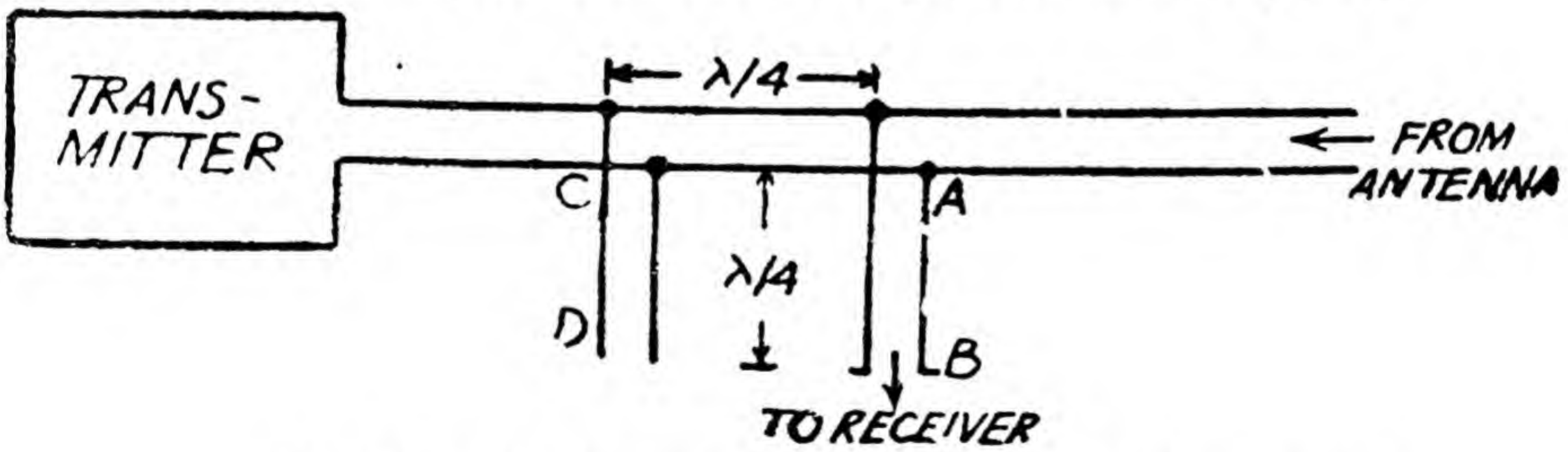
Consequently, the pulses directly reach the antenna and the receiver is isolated as shown in fig. 14 (*b*).



(*a*) Layout of Duplexer system.



(*b*) Equivalent circuit during transmission of pulses.



(*c*) Equivalent circuit when transmitter is inoperative.

Fig. 14. Duplexer action.

Now consider when the transmitter is inoperative, *i.e.* at the end of a pulse. At the instant, the gas deionises atonce and the resonance will take place in the cavity and thus both T.R. and A.T.R. boxes act as *open circuited* as shown in figure 14 (*c*). The open circuit at *D* would place a very low impedance to the line at *C* and high impedance looking towards the transmitter from *A*, with the result that the whole incoming energy goes to receiver and not to the transmitter.

33·8. RADAR RECEIVERS :

Microwave superheterodyne receivers are used in radar systems. The important characteristic is that they must have low noise figure and an appropriate intermediate frequency and video

bandwidths required to produce radar pulse. For the production of intermediate frequency, a local oscillator is chosen according to radar operating frequency. For low frequency radars, high frequency triode oscillator serves the purpose, while for radars operating at 2000 mc/s a reflex klystron is often employed. The popular intermediate frequencies chosen for radar systems are 10 mc/s, or 60 mc/s.

To detect the weakest echo signals, it is necessary that noise figure of the receiver should be as low as possible. About 1000 mc/s., r.f. amplification is employed between antenna and mixer as a means of improving the noise figure. The bandwidth of the receiver is said to be appropriate if at one side it maintains the bandwidth as narrow as possible in order to reduce the minimum detectible power and on the other side it preserves the shape of radar pulse with maximum utilisation of the latter's power. Highest signal to noise ratio is obtained with a band $1{\cdot}5/T$ app. of an I.F. system, T being the pulse width.

Since in radar operating at high frequency *viz.* 3000 mc/s., intermediate frequency is only a small percentage, it is necessary to use some automatic means of checking the intermediate frequency, *i.e.* to correct for any drift in local oscillator frequency. The discriminator circuit for it is shown in figure 15. The frequency of local oscillator and that of transmitter pulse are mixed in a

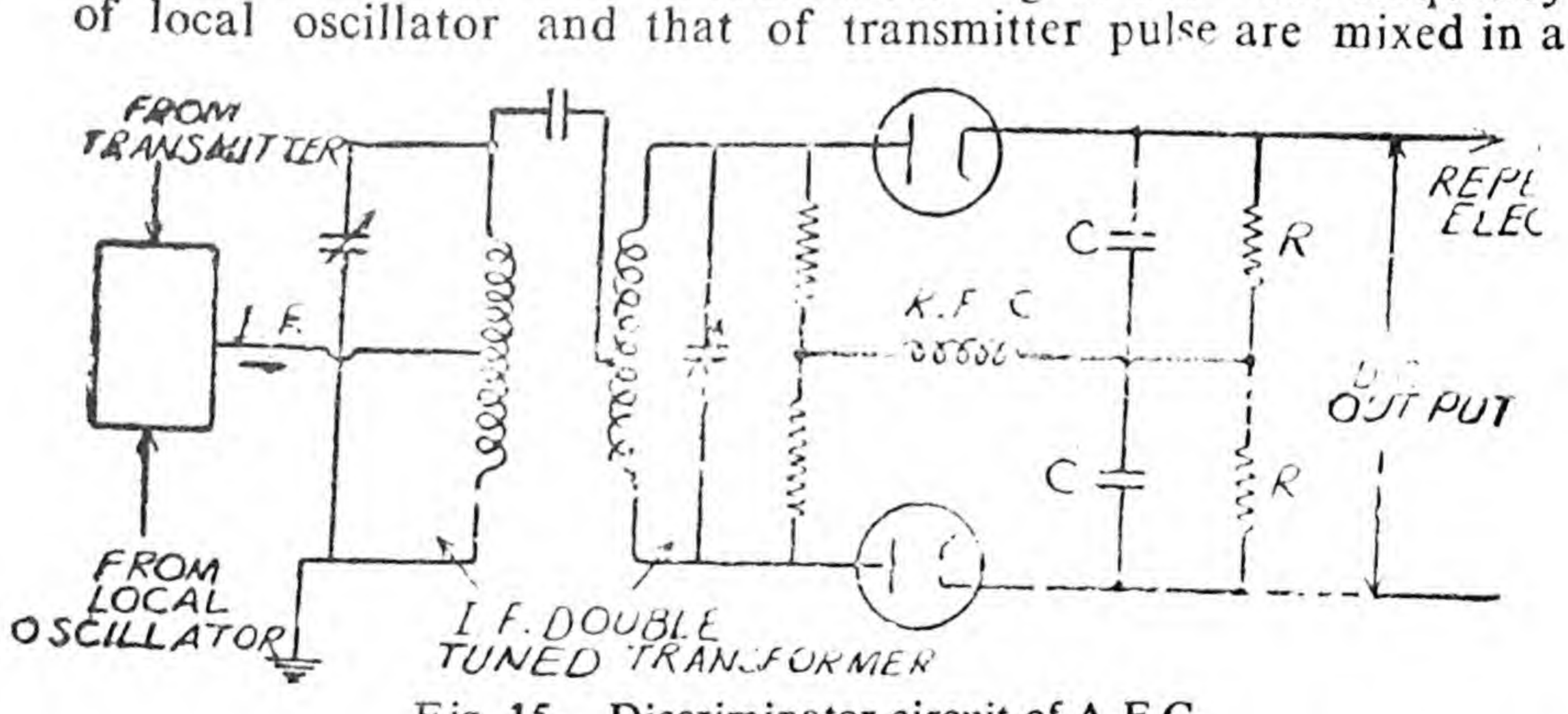


Fig. 15. Discriminator circuit of A.F.C.

separate mixer to reproduce intermediate frequency and is applied to discriminator circuit. If standard intermediate frequency is the centre frequency of the discriminator circuit, then for the correct local oscillator frequency d.c. output developed by this circuit will be zero. If there is any change in local oscillator frequency, then difference of local oscillator frequency and pulse frequency will produce a frequency other than intermediate frequency and hence a d.c. output voltage, polarity of which depends upon oscillator frequency, will be developed across the discriminator circuit which, when applied to repeller electrode of klystron, will modify its frequency in right course.

In interval between the transmitter pulses, frequency of local oscillator, *i.e.* reflex klystron changes. The time constant of discriminator output circuit should be large as to keep the output voltage sufficiently constant during this interval, so that frequency changes in reflex klystron will be approximately less than the bandwidth of intermediate frequency amplifier.

33·9. INDICATOR UNIT :

This unit presents the output of radar in visual indications. Some presentations are due to the deflection of trace on the screen called *deflection modulated presentation* while some other are due to changing brightness of trace called *intensity modulated presentation.* We have previously described both these types of indicators. *A* type indicator gives deflection modulated presentation, while *B* and *C* types utilise intensity modulation of spot. Since the principal aim of the radar is to mark the range of target, all these types of indicators bear such provisions. For better pictorial representation, angular resolution is another characteristic of these indicators.

33·10. DESCRIPTION OF RADAR SYSTEMS

For various purposes, *e.g.* homeland safety and victory over the enemy, different radar systems are designed. A brief discussion is given below :

(1) Chain Home Systems : It was first designed in 1935 for the location and detection of approaching aircraft bombers. The system for transmission consists of an array having fixed horizontal aerials and reflectors suspended from 110 metre masts. Due to this arrangement, pulse is spread out in a wider sector before the station, but a very little behind to avoid confusion with inland aircrafts. Two pairs of the aerials fitted with radiogoniometers constitute the receiving aerials system-one pair of direction finding and other for elevation measurements. The echoes and transmitted pulse are displayed on the screen of a C.R.O. to determine the range of target by the echo delay time.

(2) Chain Home Low System : The system was first introduced in 1959 in which transmitter operating wavelength was 1·5 meters (ultra short wave). The target ranges were more than 100 km. The aerial system, mounted on a turn table, was given a rotatory motion about a vertical axis at a low rate of six revolutions per minute. Beam width of 12° to 15° was employed. Same aerial was used for transmission and reception with the help of duplexer in order to secure maximum sensitivity. It is effective within the range of 5 km. from the target.

(3) Ground Control Interpretation : The major difference with chain low system is that this system involves an indicator called plan position indicator (P.P.I.). Chain home low system is not efficient one for night fighters. P.P.I. was used n this system to bring the night fighters right direction before the 5 km. range is reached. Echo from the target intensifies the spot on the cathode ray tube screen at a distance and in a direction from the centre of screen which corresponds to the target range and angular resolution. When the radar equipment is rotated. echoes from the reflection source in the neighbourhood region of the target produce bright spots on the screen at different angles and thus predicting the whole plan picture.

(4) Centimetre Radar : The exact location of isolated targets, and adequate separation of closely adjacent targets, are difficult, in above types of radar systems because the echo in such systems produces an extended patch and not a sharp well-confined spot on the screen. To produce intense narrow beam, radio pulses of wavelength less than ten centimetres were used This, on the other hand, reduces the physical size of the antenna which in the case, for better directivity, can be a parabolic one. The use of such short pulses demand more power and high sensitivity of the radar receiver.

33·11. USES OF RADAR :

The performance of radar is unaffected by darkness, fog and rain. It, therefore, can be used under all weather conditions to find the positions of mountains, icebergs in sea, shore lines, lakes etc when installed in ships or high flying aeroplanes. The radar pulses can be used for discovering the positions of buried metal, oil and ores. Radar beckons help in safe landing of aircrafts even in poor visibility due to heavy fog.

Civilian Applications : Chief civilian application of radar has been for navigation, both marine and air. Air-traffic-control radar monitors air traffic in the vicinity of airports and enroute between air terminals. In foul weather, radar is used with GCA (ground control of approach) system to guide aircraft to a safe landing. Commercial aircrafts are equipped with altimeters to determine their height above the ground and weather avoidance radar to negative around dangerous storms.

On the sea, radar is used by ships, large and small, for navigation, especially in bad weather or with poor visibility. Their use in detection and tracking of weather disturbances is significant.

Military Applications : Radar is used by military for surveillance and for the control of weapons. Surveillance radars detect and locate hostile targets for the purpose of taking proper military action. Examples of such radars are those in DEW (Distant Early Warning) line for the detection of aircraft ; the BMEWS (Ballistic Missile Early Warning System) radar for detecting and tracking of intercontinental ballistic missiles ; shapboard surveillance radars ; the AEW (Airborne Early Warning) radars ; and AI (Airborne Interception) radars used to guide a fighter aircrafts to its target and bombing radars.

Scientific Applications : It can be used as measuring tool. radar has vastly increased our knowledge of meteorology, aurora, meteors, and other objects of the solar system. Radar can be used to guide space vechicles, satellites and for the exploration of interplanetary space. In addition, the techniques developed for radar have paved the way of more advanced researches in microwave spectroscopy, radio-astronomy.

TELEVISION

Television is the electrical transmission of transient visual images. This system can transmit sound synchronously with visual images.

33·12. GENERAL PRINCIPLE OF IMAGE TRANSMISSION :

The instrument, which responds to the light intensities reaching it from the illuminated object and converts the fluctuations in brightness of light (dark portions of the object surface reflect small light while brighter portions reflect a greater amount of light) it receives into corresponding electrical variations, can serve the purpose of image transmission. Such a device is called '*pick up instrument*'. These electrical variations called *video signal* after amplitude modulation are transmitted by the antenna in conventional fashion. At the receiver, video signal after demodulation is applied to a suitable device for reproducing the corresponding light intensity variations so as to depict the basic structure of the transmitted picture. A *photo-sensitive* device can fulfil such a function in image transmission system as microphone serves in sound-transmission system.

This conversion of an optical image into a set of corresponding voltage impulses is performed in steps. The whole scene, at the transmitter, is broken into many small pictures of varying brightness and then these pictures are viewed instantaneously in a *definite sequence* by the pick up system—the process being called as *scaning*, to produce corresponding electrical signal for transmission after modulation. *In television 25 complete images are scanned and transmitted per second. Each image undergoes 625 line scanning sequence* to provide a desired degree of image detail. The whole process of image transmission can be stated in brief as follows :

(*i*) Optical image of the object of scene is focussed, by means of an optical lens system, on a photo sensitive plate contained in a *camera tube*.

(*ii*) The image is scanned by a cathode ray beam which is deflected either magnetically or electrostatically so that it traverses the photo sensitive plate in an ordered sequence.

(*iii*) Due to scanning, electrical variations corresponding to the brightness of the scanned element are produced. This video signal after modulation is transmitted.

Television transmitter sends out two separate r.f. carriers over a single antenna. One carrier is frequency-modulated by sound (audio signal), while the other is amplitude-modulated by the picture information or the video signal.

A set of pulses—called *synchronising pulses*, is also transmitted with the video signal to keep the scanning sequence at the receiver in step with that at the transmitter.

33·13. RECEPTION OF SIGNALS :

At the television receiver both audio and video signals are picked up by the receiver antenna and then subjected to heterodyne action by the conventional superheterodyne receiving circuits. After suitable amplification by the r.f. amplifier, both signals are applied to video detector. This detector separates audio I.F. signal from video I.F. signal, and then demodulates video signal. Audio I.F. signal after separation is subjected to frequency demodulation circuit.

Video signal is applied to a cathode ray picture tube, for the image reproduction. The synchronising pulses are applied to beam deflecting circuits of the tube to keep picture scanning at the screen in step with that at the transmitter camera tube.

Now we shall discuss television system in the following steps :

(1) Pick up instrument,
(2) Image scanning sequence,
(3) Scanning synchronisation,
(4) Television transmitter,
(5) Television receiver,
(6) Vestigial side band transmission.

33·14. PICK UP INSTRUMENTS :

Iconoscope, image-orthicon and *vidicon* are important electronic scan camera-tubes which find wide applications these days.

(1) Iconoscope : Schematic diagram is shown in figure 16. The

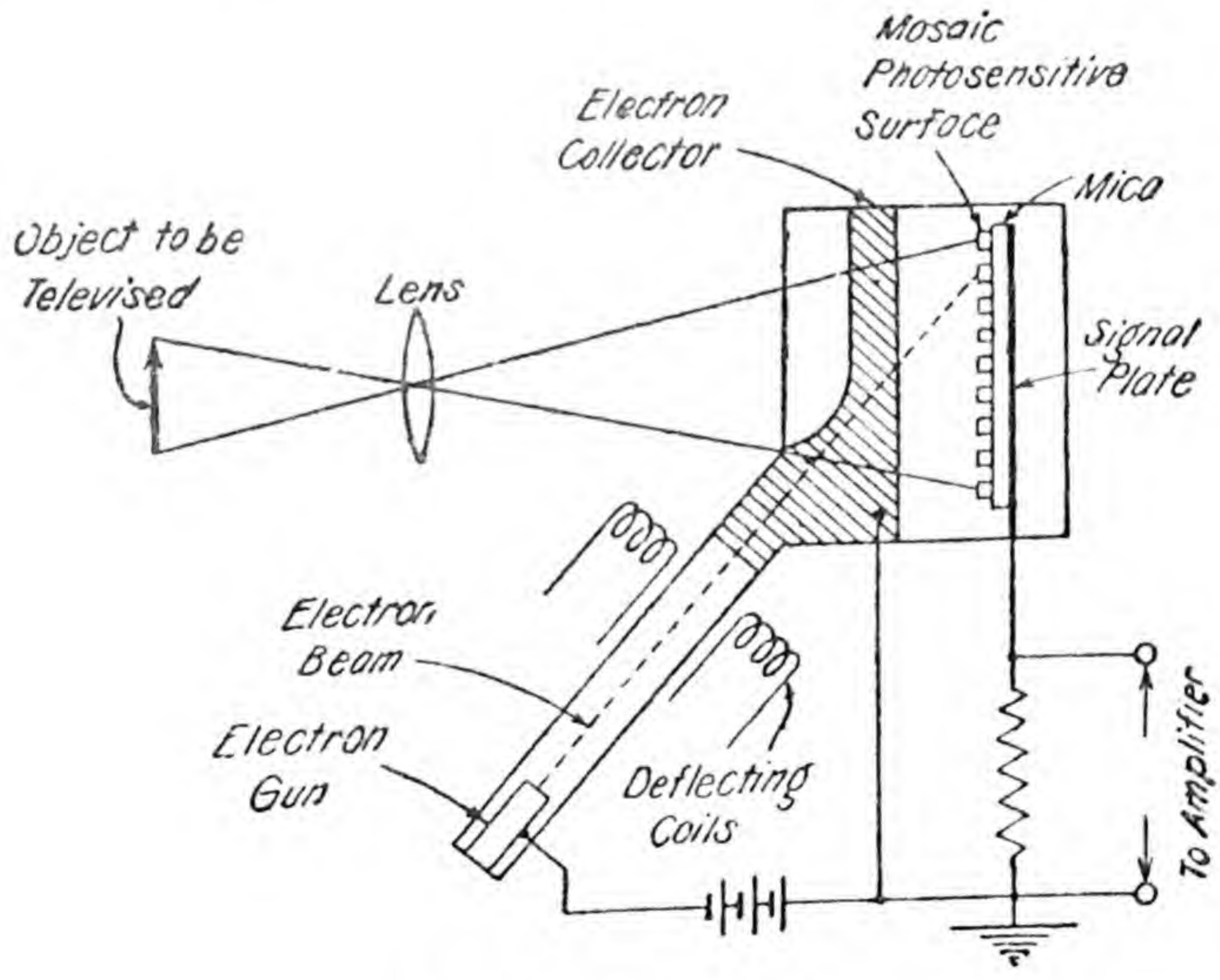


Fig. 16. Iconoscope.

optical image is focussed on a photosensitive surface called *mosaic*, which consists of a large number of minute globules of caesium silver compound discretely deposited and insulated from each other by mica on one side of mica sheet. The other side of mica sheet is coated with conducting material such as a graphite or a metal film. A load resistor, across which signal voltage is developed, is connected with this back electrode. The optical lens system focusses the image on the mosaic. Each globule emits electrons in proportion to the brightness of the particular image portion superimposed on it. Due to this loss of negative charge, the potential of globules is raised, since these tiny globules with conducting globules on the other side of mica sheet form several tiny condensers with mica as dielectric. Thus due to loss of negative charge these condensers are charged inaccordance with brightness of the portion of picture. This electrical image stored on the mosaic cannot be transmitted as a whole but the individual picture elements are scanned one at a time by discharging the globule capacitors in an orderly sequence. This is done by impinging an electron beam on the picture elements in a sequence governed by the deflecting coils for this beam. The scanning is from left to right and top to bottom of the mosaic, one line at a time. When the beam impinges on a particular globule, it restores the charge lost by the photo-electric emission, *i.e.*, tiny condenser is discharged through load resistor ; the discharge current develops the corresponding video signal across the resistor.

The main drawback of this pickup device arises due to the secondary electron emission which occurs when high velocity scanning electrons strike the mosaic. Some of these emitted electrons are drawn off to a collector electrode while others return to mosaic and thus *produce improper charge distribution for a given image* This drawback is removed in another pickup device called image-orthicon.

(2) **Image Orthicon** : It is a sensitive tube and is capable of handling a wide range of light values and contrast. In a single envelope, it includes three sections :

(*a*) *Image section* : This section includes :

(i) a photo sensitive surface, called photo cathode, operated at a very large negative potential.

(ii) a target plate which is a thin plate of glass of low resistivity. Thickness is less than 0·0002 in.

(iii) a screen located very close to target plate and has about 500,000 openings per square inch.

When the optical image is focussed on the photo cathode, photoelectrons, in proportion to the amount of light impinging, are emitted. Most of the photoelectrons pass through the screen and hit the target plate. As the photoelectrons are accelerated to several hundred electron volts, they liberate several secondary electrons

from the target plate surface, and are then collected by the nearby-screen which is at a small positive potential. The emission of

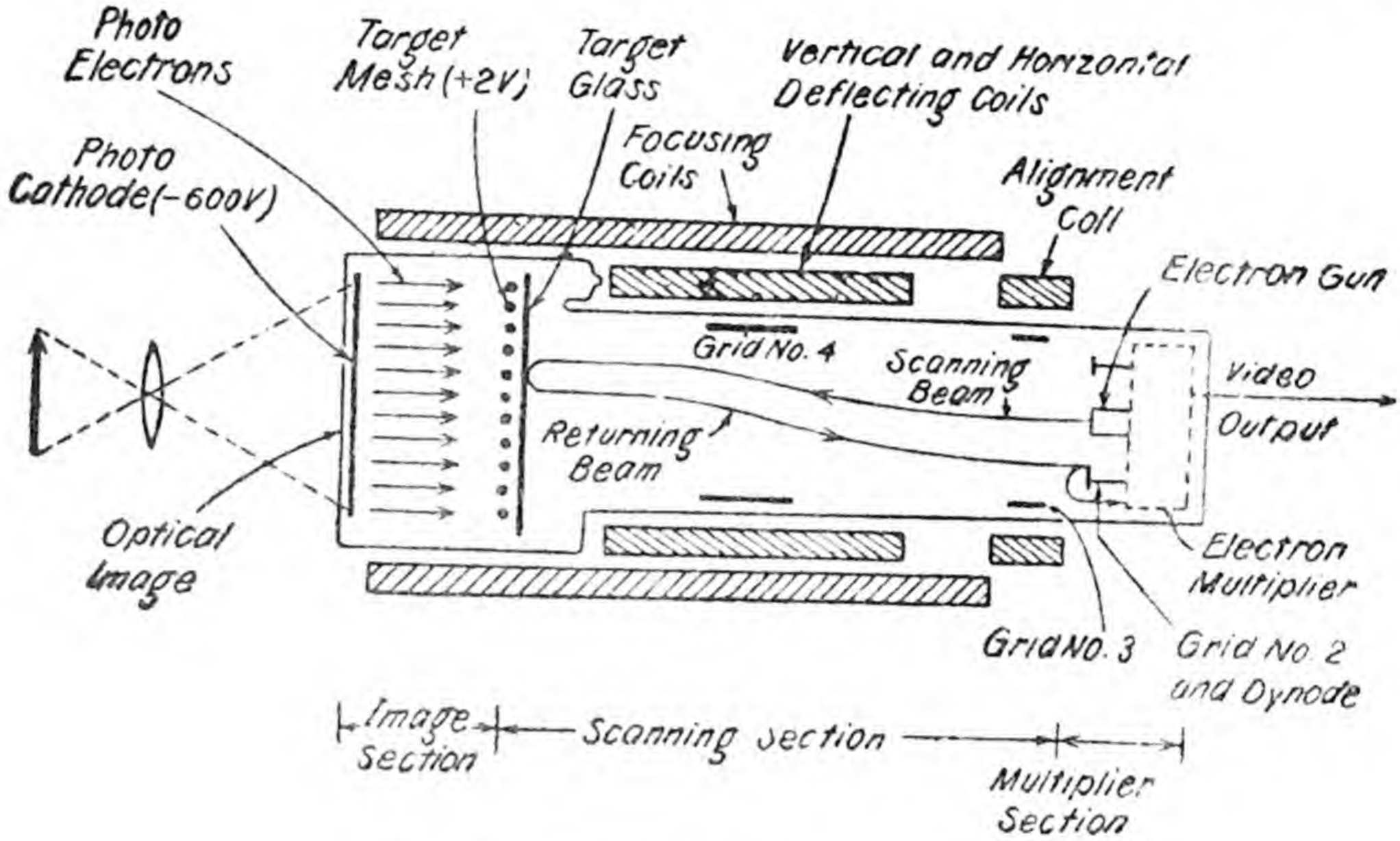


Fig. 17. Image Orthicon.

secondary electrons from target plate leaves a distribution of positive charge on its surface. The low resistivity of target plate resists the lateral charge flow on its surface and thus the image charge pattern, formed on the plate, is truly restored as such. Since the plate is thin, this charge pattern also appears on the other side (away from screen) of the plate.

(*b*) *Scanning section* : The otherside of the plate is now scanned by a beam of low velocity electrons generated by an electron gun. The beam is deflected on the plate by vertical and horizontal deflection coils (shown in fig. 17). Deflection in vertical and horizontal direction as well enables the electron beam to scan the whole plate. This beam gives up the number of electrons required to neutralise the positive charge at that point and thus the returning electron beam varies in magnitude inaccordance with the brightness variation of the image. It should be noted here that since the target portion affected by the white portion of the image will be much positively charged and hence the electron beam has to give up large number of electrons to neutralise the positive charge at that point, *i.e.* the intensity of returning electron beam is much reduced and the video signal developed across the output resistor for this part will be small It, therefore, concludes that the brightest part of the image are transmitted as the signals of low amplitude which is very advantageous in avoiding the effect of strong noise at the receiver.

(*c*) *Electron multiplier section* :

An electron multiplier is located within the pick up tube to amplifying the electron density variation in the returning beam.

(3) Vidicon : This camera tube is based on the photo-conductive properties of semiconductors *i.e.* decrease in resistance with the amount of incident light. The tube is shown in fig. 18. It consists of :

(*i*) *Signal plate* : Which is a conducting metallic film very thin so as to be transparent. The side of this film facing cathode is

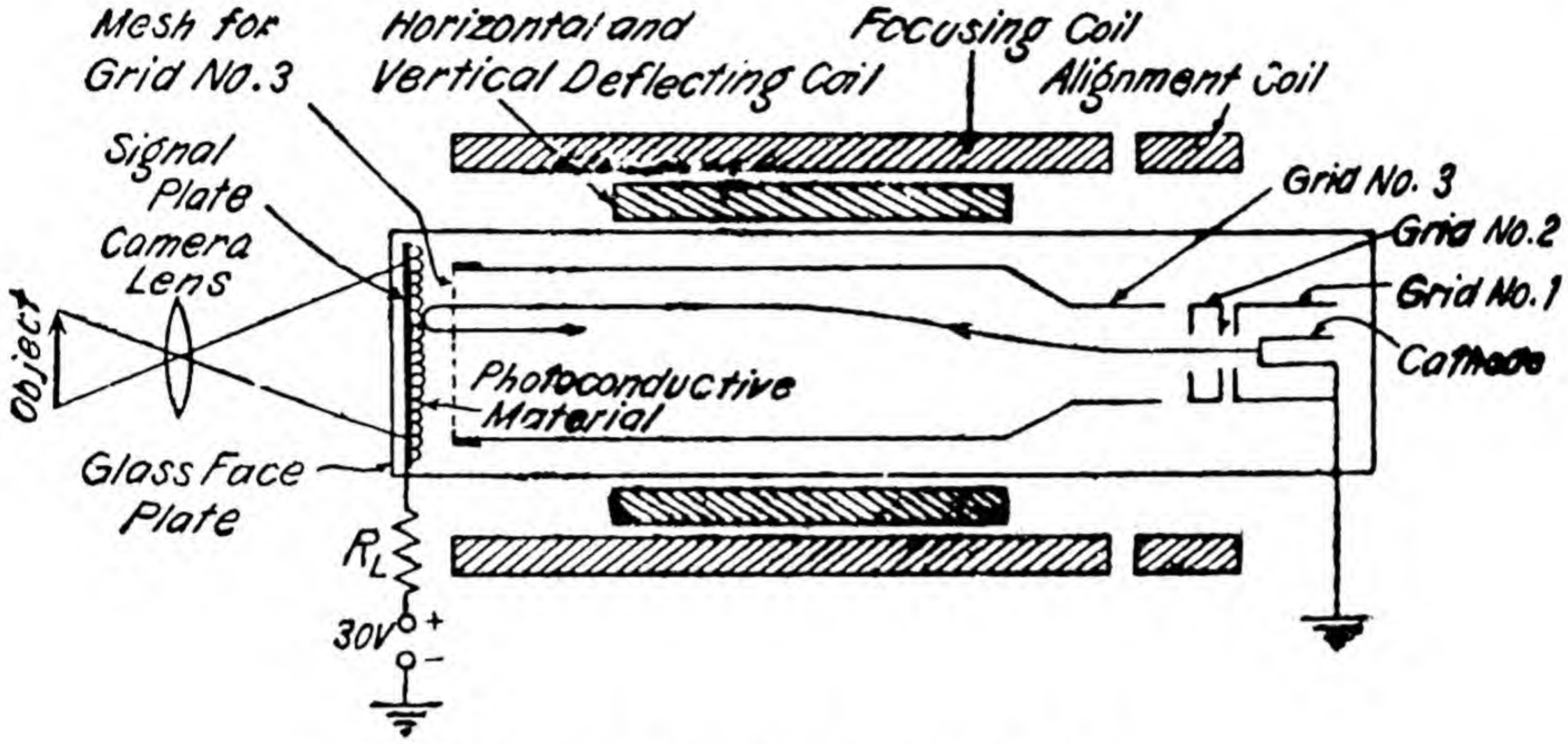


Fig. 18. Vidicon camera tube.

coated with a very thin layer of photoconductive material (amorphous selenium). This side is scanned by electron beam. The optical image is focussed on the otherside of this film.

(*ii*) *Scanning System* : The electron beam for scanning is formed by the combination of cathode, control grid 1, accelerating grid 2 and anode grid 3. The focussing coil produces an axial field which focusses the beam on the film. Vertical and horizontal deflection of the beam, so as to scan the whole film, is accomplished by passing saw-tooth current waves through deflecting coils which thus produce transverse horizontal and vertical magnetic fields respectively. The alignment coils are for initial adjustment of the direction of electron beam.

Operation : When the scanning beam passes over the photo-conductive material of the signal plate, it deposits electrons so that potential of this side of plate is reduced to that of the cathode. But the otherside of the film (plate) is still at its original potential. Consequently a potential difference *across* a given point on the photoconductive material is created. It is approximately 30V. Before the next scanning (which may be done after an interval of 1/50 or 1/25 sec.), the charge leaks through photoconductive material *at a rate determined by the conductivity of the material which, in turn, depends upon the amount of incident light.* White portions of the object will project more light on the film and make it more conductive. Thus charge leaked to photoconductive side of the film will vary according to illumination of the object. As a

result potential at every point on the photoconductive side will vary. Now the electron beam again starts scanning the photo-conductive side of the film but this time the charge deposited by the beam inorder to reduce its potential towards zero (cathode potential) will vary with time. Therefoee current through load R_L (and hence the output voltage) will follow the changes in potential difference between two surfaces of the film and hence follows the variations of light intensity of successive points in the optical image.

Advantages : 1. Low Cost
2 Simple adjustment
3. Sensitivity is large
4. Resolutions of the order of 350 lines can be achieved under practical conditions

Disadvatages :

1 Owing to the fact that the resistance of the photoconductive film does not change instantaneously with change of light intensity, different levels of light intensity are adjusted with slight time lag.

2. The response characteristic is non-linear.

33·15. IMAGE SCANNING SEQUENCE :

There are two types of scanning—one, the progressive or *Simple scanning* [fig. 19 (*a*)] and the other, *interlaced scanning* [fig. 19 (*b*)]. The latter is very much adopted these days.

(*a*) *Progressive Scanning* :

During the process of scanning the horizontal and vertical beam deflection circuits operate simultaneously, so as to cross the image in downwards slant line as shown in fig. 19 (*c*). On reaching at the end of each line, scanning spot retraces horizontally to the starting point of succeeding line. To make this retrace invisible a blanking pulse is applied to the control grid of the electron gun.

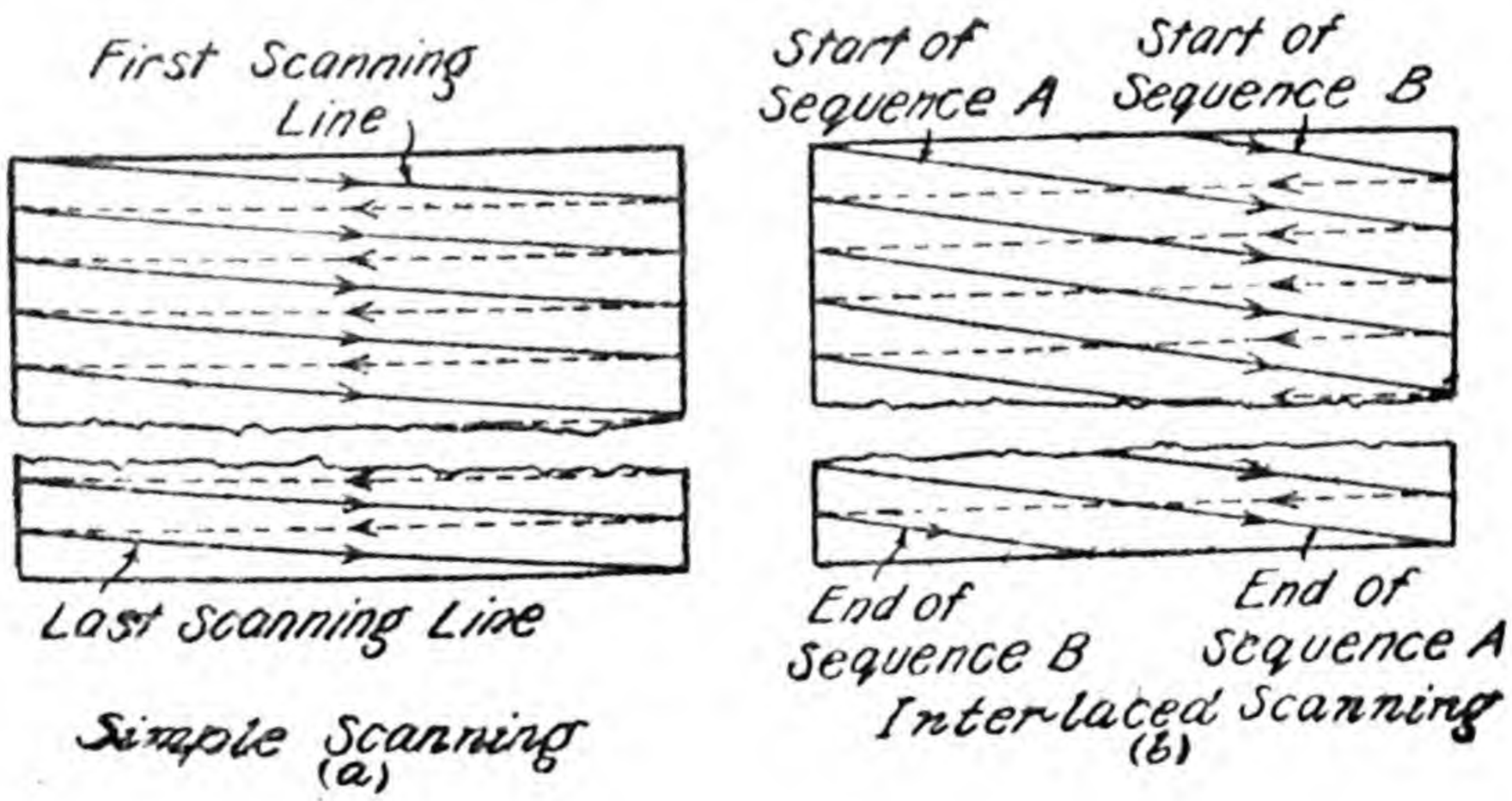


Fig. 19. Scanning sequence.

of picture and camera tube. For picture tube, these blanking pulses are transmitted along with video signal. To keep scanning at picture tube in step with that at camera tube, horizontal and vertical synchronising pulses are also sent along with blanking pulse and are superimposed on the latter. The number of horizontal lines scanned has been standardised at a total of 625 per image or frame and 25 such frames are scanned per second.

(*b*) *Interlaced Scanning* : In India, 25 frames are scanned per second, *i.e.* the frame repetition rate has been standardised at 25 per second. Despite high repetition rate, a certain amount of flicker is still present. To eliminate this residual flicker each frame is shown twice on the face of the picture tube. It is called interlaced scanning. As shown in fig 19 (*b*), first sequence *A* starts and scanning of every line except next occurs, *i e*, 625/2=312·5 lines are scanned. then the spot after reaching the bottom of the image area in 1/50 seconds, returns to the top of the image area and then scanning of sequence *B* starts. Rest 312·5 lines are now scanned in 1/50 seconds. Thus two fields are shown in each frame with repetition rate of 50 per second for each field. The picture element always bears a ratio of width to the height, called aspect ratio, of 4 : 3. That is width is 1·33 times the height of the picture frame.

With 625 horizontal lines in each frame and 25 frames being scanned per second the total number of horizontal lines scanned per second is $625 \times 25 = 15{,}625$ lines. The horizonal repetition rate or line scanning frequency is thus 15,625 c/s. As the frame scanning frequency is 25 c/s and each frame is scanned in two fields, the field scanning frequency is $25 \times 2 = 50$ c/s.

Provision for Scanning in Transmitter :

The movement of the electron beam during the scanning proceses is accomplished by the application of saw tooth waves to the horizontal and vertical deflection elements of the camera tube. Such voltage waves, for picture tube, are generated by saw tooth generator when triggered by the synchronising pulse sent with blanking pulse.

To control the duration of each scanning line and the scanning sequence, the saw tooth generators are triggered by a series of carefully timed pulses For instance, the time '*H*' allowable for sweeping a horizontal line is 1/15,625 second or 64 micro seconds. This includes a retrace time of about 0·15 *H* of the total time *i.e.*, 9·6 microseconds. Therefore duration of each scanning line is normally set at 64–9·6 = 54·4 micro seconds. Thus a train of horizontal synchronising pulses is generated at a rate of 15,625 pulses per second in order to trigger the horizontal-deflection saw tooth generator at intervals of 64 μs.

It is to be noted that time for sweeping a vertical field is considerably longer, being equal to 1/50 seconds. The retrace is only 7% of the trace time. Thus a train of vertical synchronising pulses is generated at the rate of 50 pulses per second to trigger the vertical deflection saw tooth generator to maintain the repetition rate of 50 c/s per second.

33 16. SCANNING SYNCHRONISATION : COMPOSITE VIDEO SIGNAL :

To produce a true and undistorted picture, it is necessary that the scanning process at the transmitter camera tubefshould be quite in step with that at the receiver picture-tube. Thus the timing pulses generated by the synchronising generator to trigger the saw tooth generator for vertical and horizontal plates are not only applied to the transmitter camera tube system but also transmitted to the receiver along with the image signals. At the receiver these triggering pulses are separated from the signal components, which are then differentiated (horizontal synchronising pulses) and integrated (vertical synchronising pulses) to trigger saw-tooth wave generators for the application of saw-tooth voltage to horizontal and vertical deflection plates of picture tube respectively. Blanking pulses are also transmitted to cut off the electron beam during horizontal and vertical retrace time.

In fig. 20 a blanking pulse along with superimposed horizontal synchronising pulse and video signal for two successive horizon-

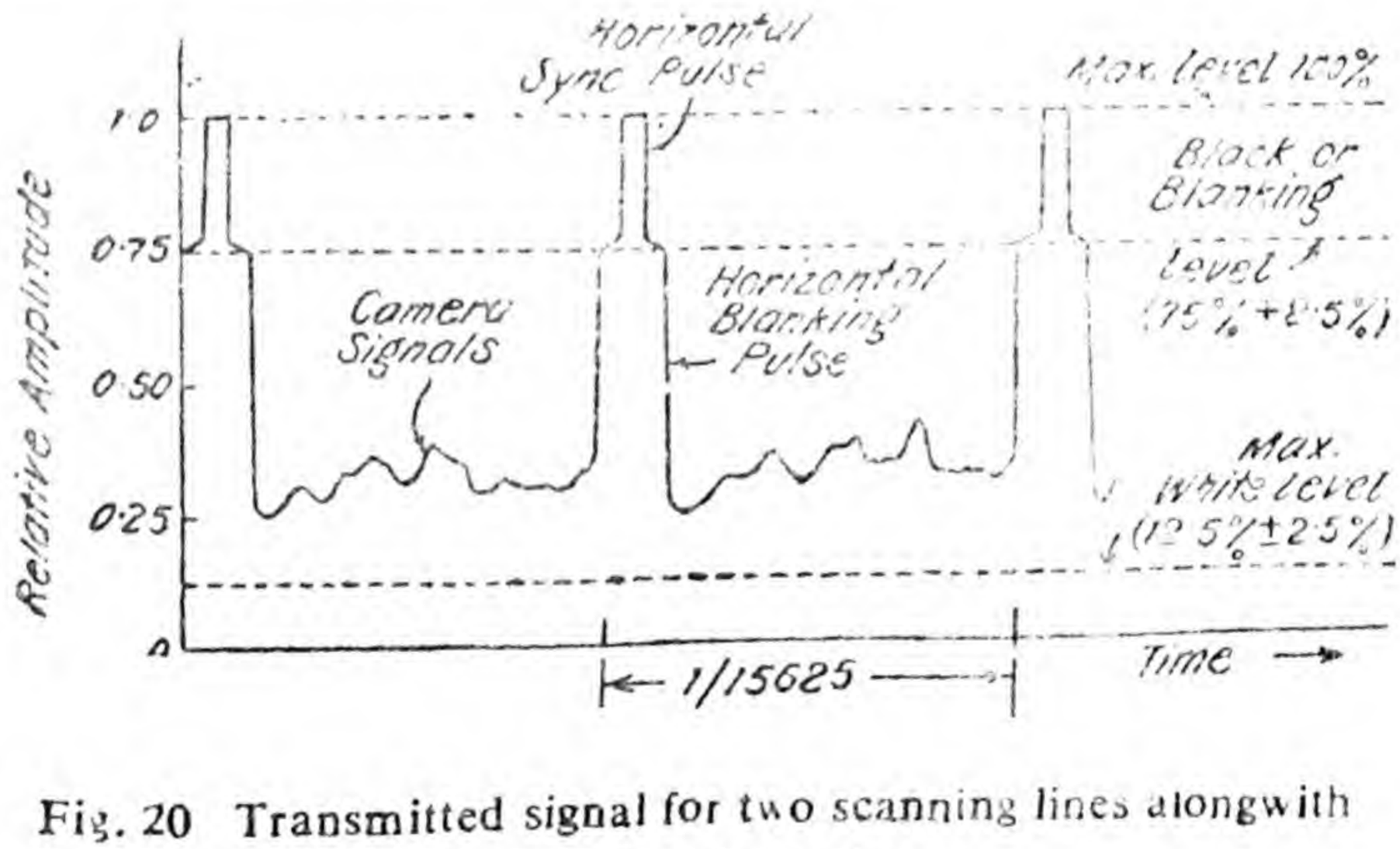


Fig. 20 Transmitted signal for two scanning lines alongwith blanking and synchronising pulses.

tal scanned lines are shown. *This set of pulses alongwith signal is called composite video signal.* The amplitude of blanking pulse is sufficient to cut off electron beam. The amplitude of synchronising pulse is greater than blanking pulse and therefore it can be separated from signal and blanking pulse when applied to a clipper

circuit which clips off all signals below the blanking level so that only the synchronising signals are permitted to pass. This is illustrated in fig. 21. The clipper circuit consists of a pentode

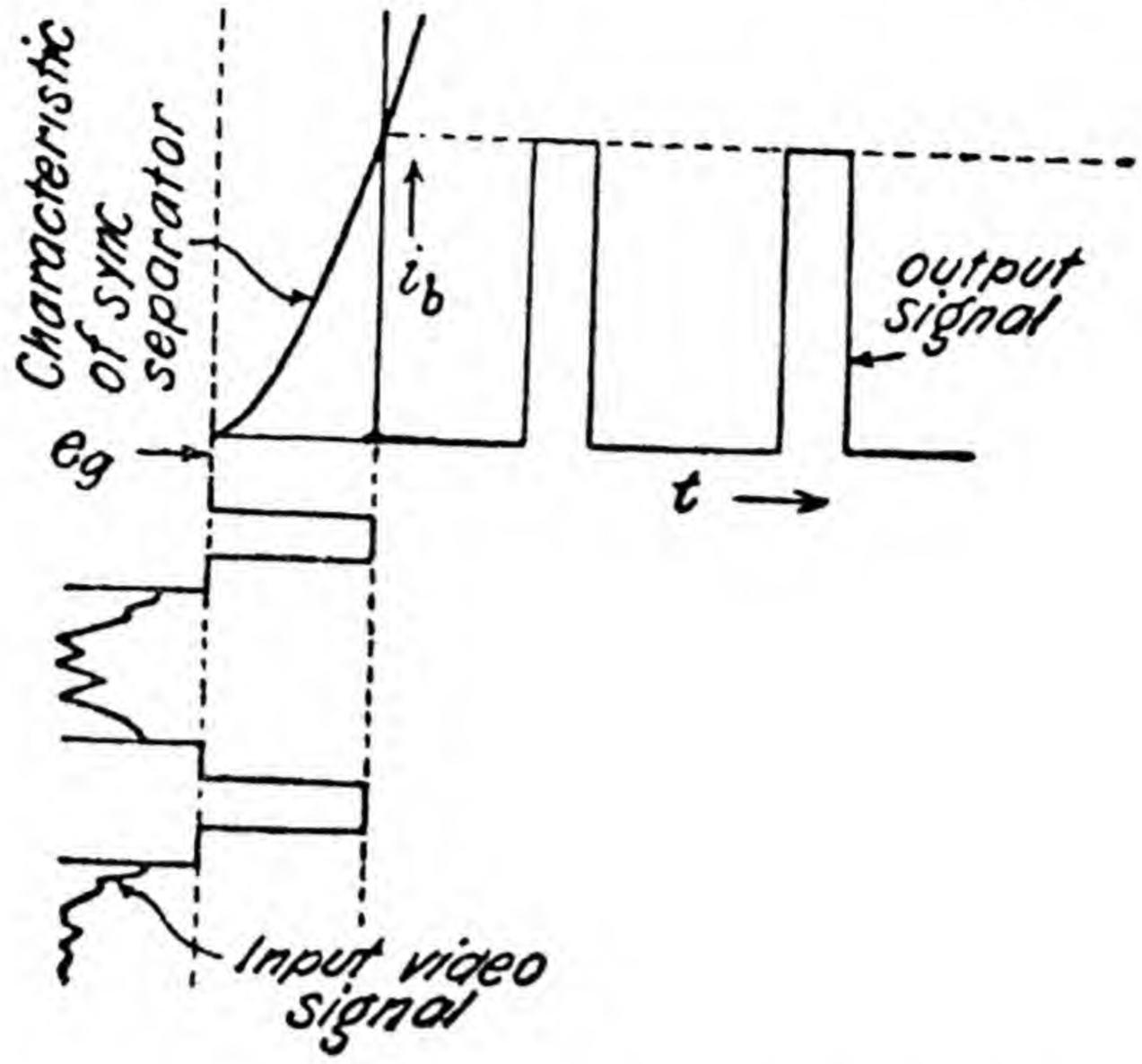


Fig. 21. Action of synch. separator.

tube with a low screen grid voltage. The composite signal, in which the *synchronising signals must be positive going*, is applied to the control grid. The circuit is designed for d.c. restoration at control grid so that synchronising level is approximately at zero volt. Because of low screen voltage, the grid bias is very limited and the picture signal are well beyond cut off. Thus plate current of pentode will correspond to synchronising signal only and gives isolated synchronising pulse. The isolated synchronising pulse is then applied to a horizontal deflection saw-tooth generator.

As the last line of a field is scanned, *a low blanking interval occurs in which a number of pulses are generated by the synchronising generator* as shown in figure 22.

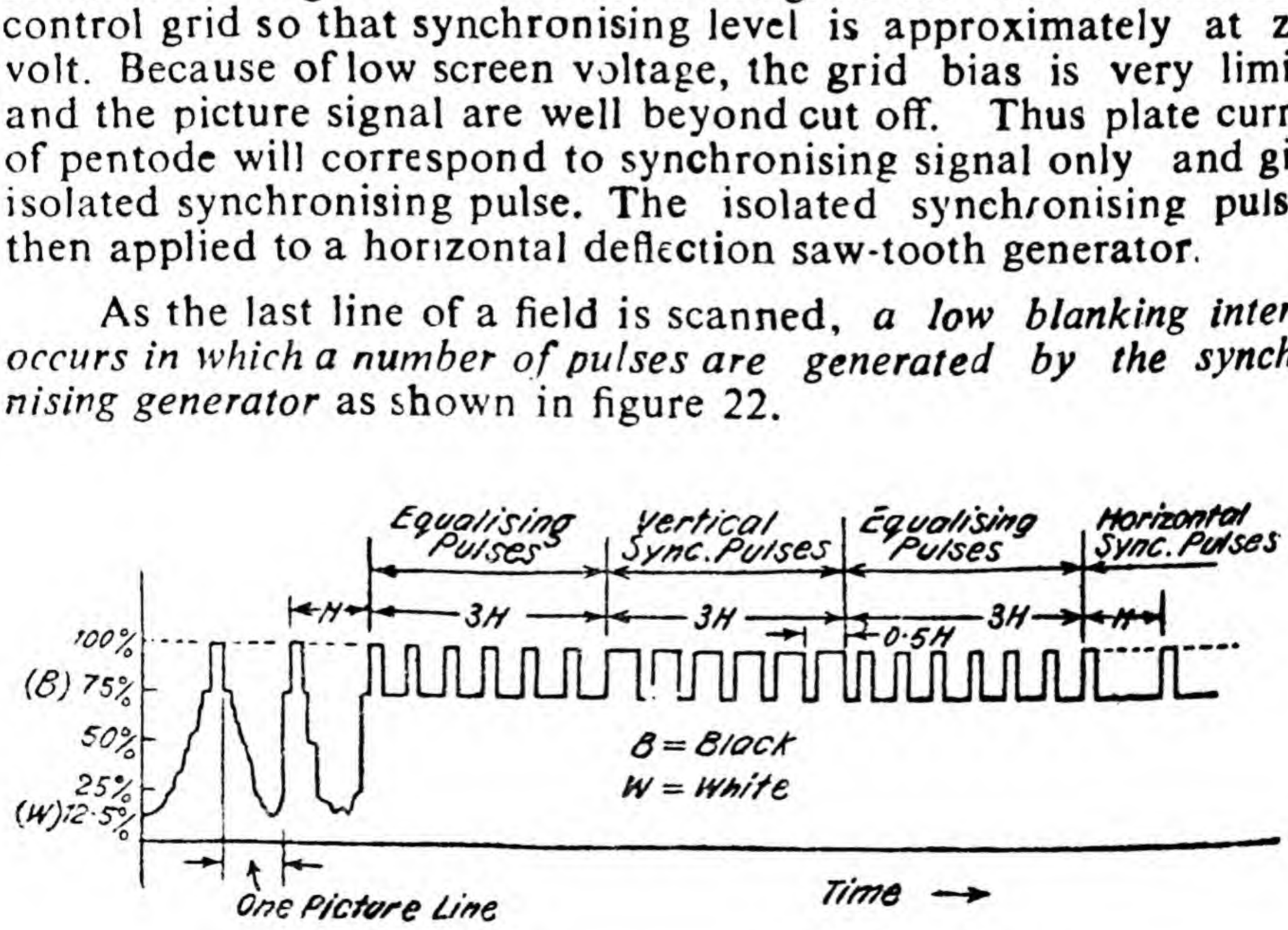


Fig. 22. Vertical and horizontal synchronising pulses.

It consists of

(*i*) horizontal synchronising pulses for the last three horizontal lines of field,

(*ii*) six equalising pulses somewhat smaller in width than horizontal pulses,

(*iii*) a long vertical pulse group divided into six individual pulses by the serrations,

(*iv*) Three more horizontal pulses for first three horizontal lines of the next field.

Each vertical pulse extends over a period of three horizontal lines, the serrations or notches being inserted at half line intervals. Due to the insertion of such serrations, it is possible to maintain horizontal synchron'sation while the vertical synchronisation is being accomplished. Following these vertical synchronising pulses are equalising pulses. As is obvious from scanning discussion that scanning of sequence *B* ends in the middle of the line while that of sequence *A* ends at the end of a line, both fields are out of phase by one half cycle. Thus, to secure required interlacing synchronising signals for alternate fields, these equalising pulses are provided.

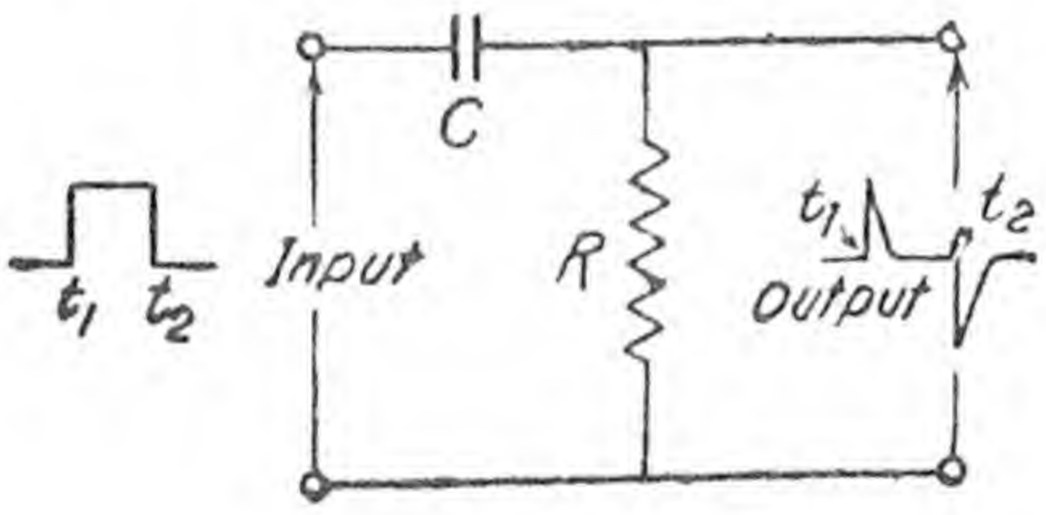


Fig. 23. (*a*)

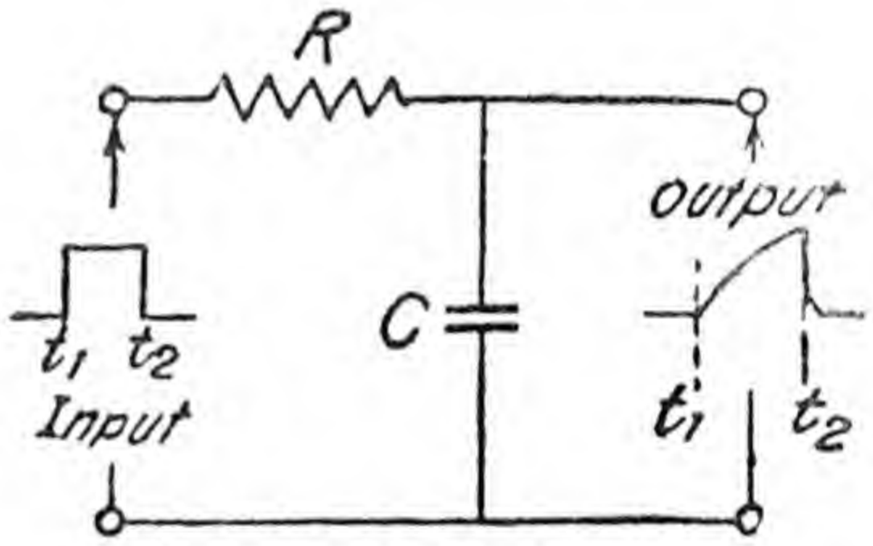


Fig. 23. (*b*)

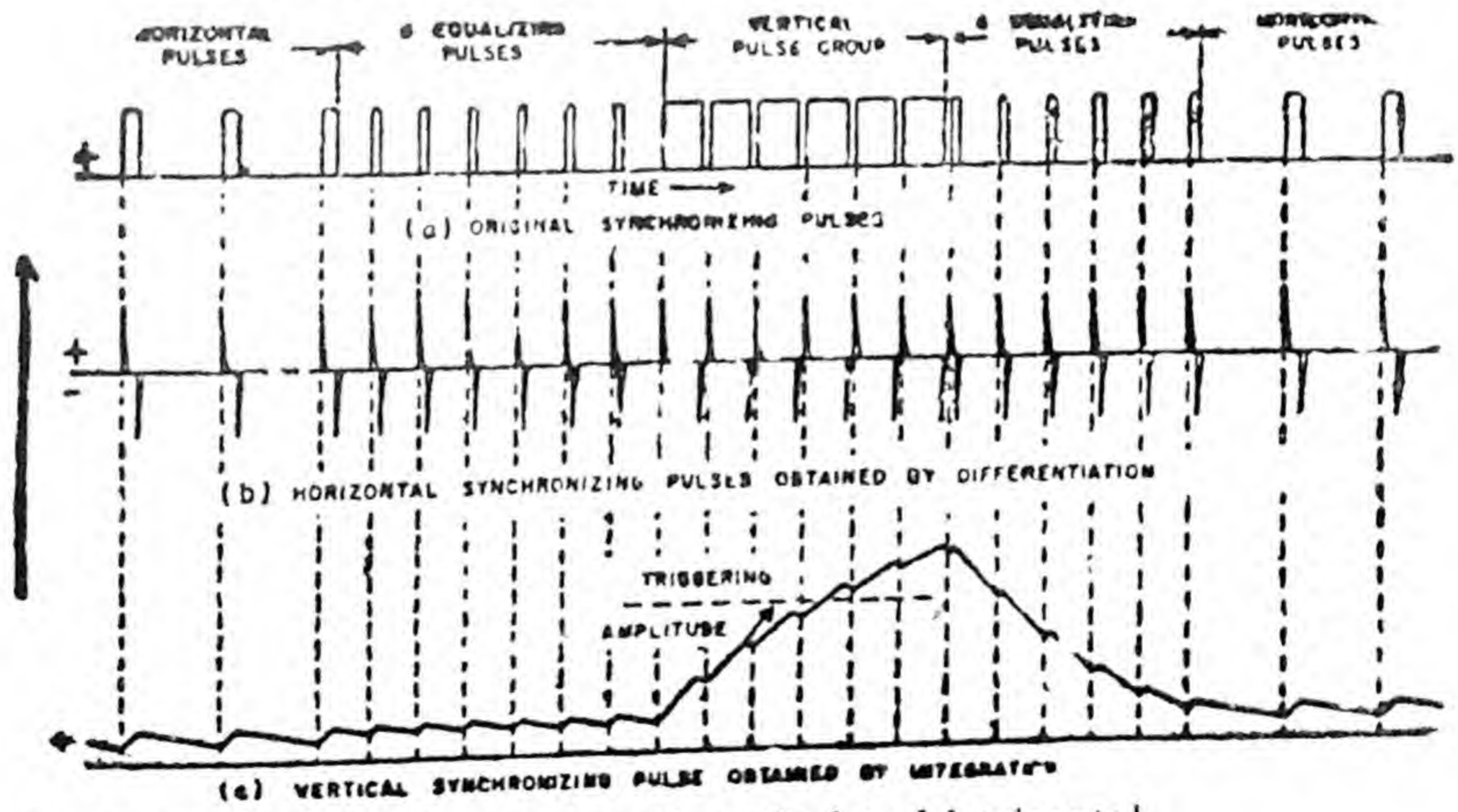


Fig. 24. (*a*), (*b*), (*c*) Separation of vertical and horizontal synchronising pulses.

At the receiver, separation between the vertical and horizontal synchronsing pulses is accomplished on the basis of the difference in their duration. Horizontal synchronising pulses are applied to a differentiating circuit to give a pulse shown in figure 23 (*a*) or 24 (*b*) while vertical synchronising pulses are applied to an integrating circuit to give a voltage variation like figure 23 (*b*) or 24 (*c*). The amplitude rises to a level which is above the synchronising control level for vertical deflection saw tooth generator only during the vertical synchronising pulse interval. The differentiated output is applied to trigger horizontal deflection saw-tooth generator. Equalising pulses associated with vertical synchronising pulses help to secure the same wave-form at the end of each field scanning.

MORE ABOUT COMPOSITE VIDEO SIGNAL :

As explained earlier video signal is composed of :

(*i*) camera signal carrying p icture information,

(*ii*) synchron'sing pulses to synchronise the scanning at transmitter and receiver, and

(*iii*) blanking pulses to make the retraces during scanning as invisible.

In fig. 20, we have depicted the compcsite signal. Note that amplitude of the video signal is divided in two parts : the lower $(75 \pm 2{\cdot}5)\%$ being used for camera signal and upper $(25 \pm 2{\cdot}5)\%$ for synchronising pulses. The camera signal is transmitted using a standard negative polarity of transmission which implies that white parts of the picture are represented by low amplitudes in the transmitted picture carrier signal whereas higher amplitudes correspond to darker picture information until black level is reached which is the fixed level at $(75 \pm 2\ 5)\%$ of maximum signal amplitude.

Black Reference Level : This level is constant at 72·5% amplitude and independent of picture information. When the image is reproduced, 72·5% of the video signal corresponds to the grid cut-off voltage of the picture tube and the absence of light. A black level is thus established. This level serves as the reference level in defining the brightness values of various shades of white and grey.

The Blacking Pulses : They are included in the composite video signal in order to make the retrace lines invisible by raising the signal amplitude to black level during the time the scanning circuits produce retraces. The horizontal blanking pulse blank the retrace from right to left in each horizontal scanning line. The repetition rate of these pulses is therefore, equal to the scanning frequency 15,625 c/s. The vertical blanking pulses blank the retrace when electron beam retraces vertically from bottom to top in each field. Therefore their frequency is 50 c/s.

(*a*) *Horizontal Blanking Time* : The interval between horizontal *scanning* lines is indicated by *H* (fig. 22) and is the time for scanning one complete line, including trace and retrace equal to 1/15,626 sec. or 64μ sec. Since horizontal *blanking* pulse has a

widtn about 18·5 to 19·2 percent of H, the horizontal blanking time is $0{\cdot}19 \times 64\mu$ sec. for H which equals $12{\cdot}16\mu$ sec. The time for visible scanning is $64 - 12{\cdot}16 = 51{\cdot}84\mu$ sec.

Horizontal sync. pulse is 0·07 to 0 077 or one half the average width for blanking pulse. This time equals 4·5 to $4{\cdot}9\mu$ sec.

For the remaining half of the blanking time, which is also 4·5 to $4{\cdot}9\mu$ sec., the signal is at the pedestal level. *The part of the pedestal level just before the sync. pulse is called the front porch and that which follows the sync. pulse is called back porch.* The front porch is 0·02 to 0·028 H or 1·28 to $1{\cdot}8\mu$ sec. and top back porch 0·06 to 0·84 H or 3·84 to $5{\cdot}4\mu$ sec. The back porch is three times longer than the front porch. Thus out of total blanking of $12{\cdot}16\mu$ sec., $1{\cdot}8\mu$ sec. is used for front porch, $4{\cdot}9\mu$ sec. for sync. pulse and $5{\cdot}4\mu$ sec. for the back porch.

(*b*) *Vertical Blancking time* : The vertical blanking pulses raise the video signal amplitude to black level so that the scanning beam is blanked out during vertical retraces. The width of the vertical blanking pulse is $[(18 \text{ to } 22)\ H + 12]\mu$ sec. *i.e.*, 1164 to 1420μ sec. The interval between scanning lines is 64μ sec. Therefore in this vertical blanking time of 1400μ sec., say. about 22 lines are blanked out in each field or 44 lines in the frame.

33·17. RESOLUTION AND BANDWIDTH IN T.V. SYSTEM :

We know that for transmission, T.V. picture is divided into a number of horizontal lines and each line is scanned by the electron beam of the camera tube. Each line may be further considered as consisting of a large number of picture elements or in other words, the whole picture may be thought of as consisting of a large number of picture elements. But there is some limitation to the faithful reproduction of picture elements which arises on account of the fact that any change in light intensity of the image in *vertical* direction occuring in a distance *less* than the width of the line cannot be reproduced and thus if there is to be as much detail in the horizontal direction as in the vertical, the picture elements must be of the same size as the distance between the lines. The most detailed pattern that can be transmitted is chequer board pattern of black and white *squares* as shown in fig. 25. With an aspect ratio (width/height) 4/3 and 625 lines every $\frac{1}{25}$ sec, the number of elements to be scanned is given by

$$n = 625 \times (\tfrac{4}{3} \times 625) \times 25.$$

Therefore the time for reversal from black to white in this case is

$$\frac{1}{n} = \frac{3}{625 \times 4 \times 625 \times 25} \text{ sec.}$$

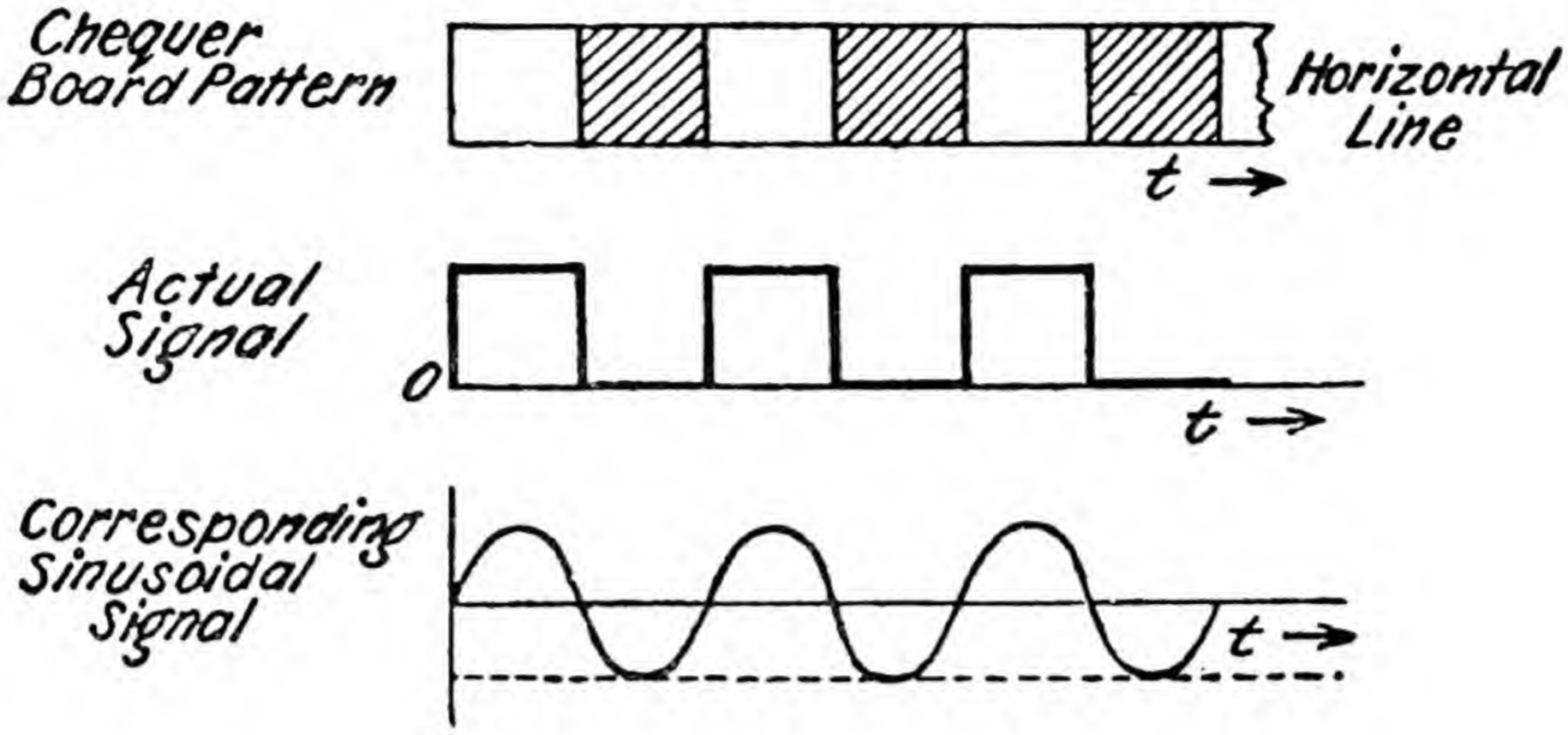


Fig. 25. Chequer board pattern of TV image signal and corresponding electrical signal.

and the time for a complete cycle, black-white-black, is thus

$$T = \frac{3 \times 2}{625 \times 4 \times 625 \times 25} \text{ sec.}$$

and hence the frequency band covered by T.V. video signal is

$$\text{Band Width (BW)} = \frac{1}{T} = \frac{625 \times 4 \times 625 \times 25}{3 \times 2}$$

$$= 6{\cdot}5 \text{ mc/s.}$$

In actual practice it is found that equal horizontal and vertical details can be obtained provided a bandwith about 70 to 80% of 6·5 mc/s is transmitted Therefore, in Indian TV system video bandwidth is kept at 5 mc/s.

33·18 TELEVISION TRANSMITTER :

Figure 26 shows the simplified block diagram of a television transmitter. The video signals obtained from camera tube are applied to a number of video amplifier stages. First stage is located in camera housing to increase weak signal voltage to such a level as to be transmitted over a coaxial cable to the succeeding amplifier stages.

Synchronising generator produces sets of pulses to operate the system at appropriate timings. This unit includes wave generating and shaping circuits, *e.g.* multivibrator circuit, blocking oscillator circuit, and clipping circuits etc. The repetition rates of these pulse trains are controlled by frequency stabilised master oscillator.

The horizontal synchronising pulses are applied to horizontal deflection saw-tooth generator ; vertical synchronising pulses are applied to vertical deflection saw-tooth generator ; two sets of blanking pulses are applied to control grid of camera tube to blank

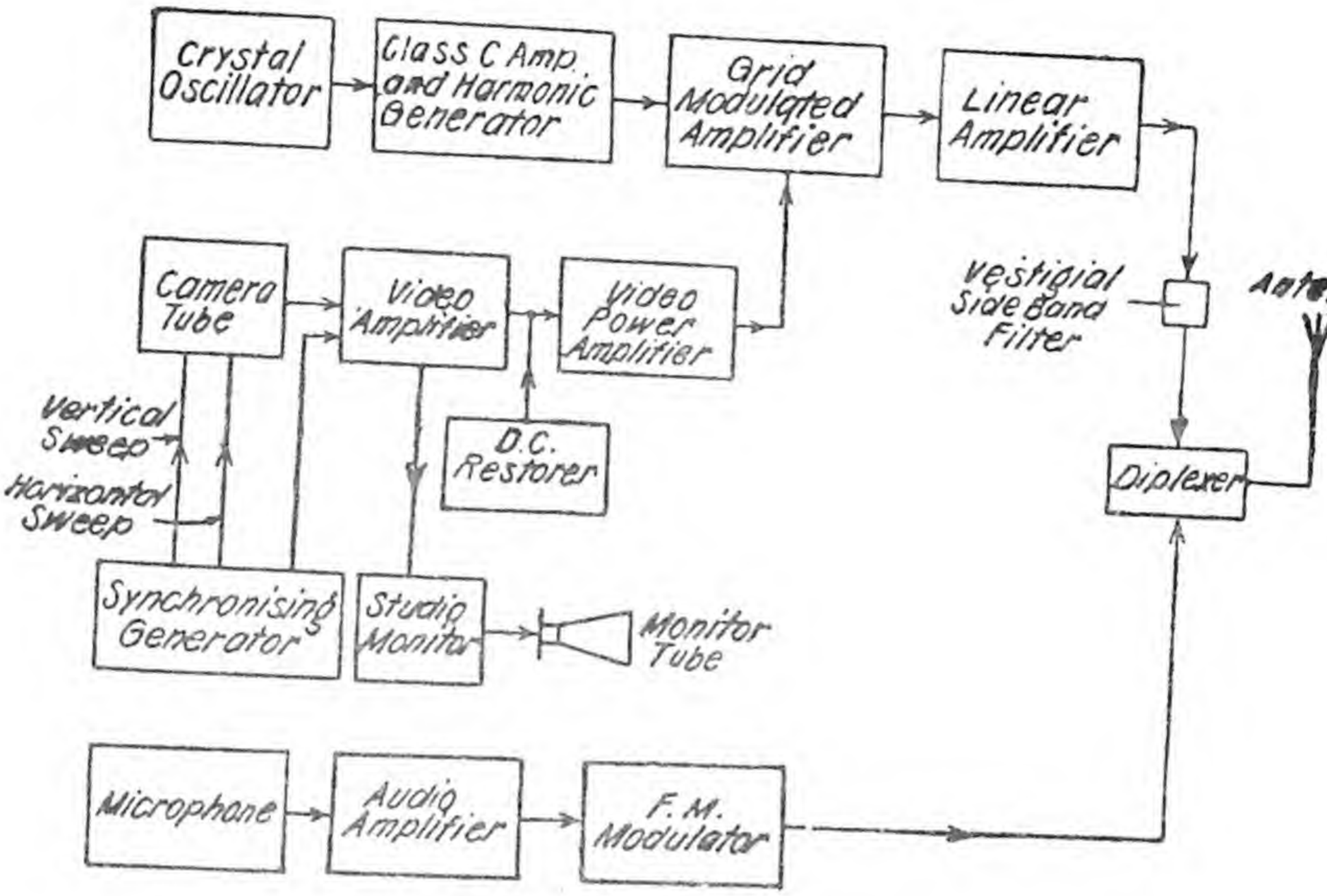


Fig 26 Simplified block diagram of a television transmitter.

it during vertical and horizontal retrace ; and a pulse train consisting all above puise groups is applied to video-amplifier channel for transmission to receiver.

The carrier frequency generated from a crystal controlled oscillator is passed through a number of frequency multiplier and amplifier stages. This results in the production of a carrier wave of desired frequency and energy content. The level of image signals, together with synchronising and blanking pulses, is raised to modulate this carrier frequency. A high level grid modulation is usually employed.

The carrier when amplitude modulated with video signal (BW = 5 mc/s) generates two sidebands and the total bandwidth required for TV channel would be 10 mc/s which is too large. Therefore vestigial sideband transmission in which one sideband (say upper) is transmitted in full along with reduced second sideband is used. For this purpose, output of the final r.f. amplifier is applied to a vestigial sideband filter which suppresses the undesired portion of the lower sideband of the modulated wave. The frequency spectrum of such a sideband transmission is shown in fig. 27.

The modulated r.f. energy is carried from transmitter to the transmitting antenna by means of a coaxial transmission line. The antenna elevation is kept high for larger transmission area.

A f.m. transmitter is used for the purpose of audio signal transmission. The carrier frequency used in audio modulation is 5·- mc/s. above that which is used in audio modulation. Both,

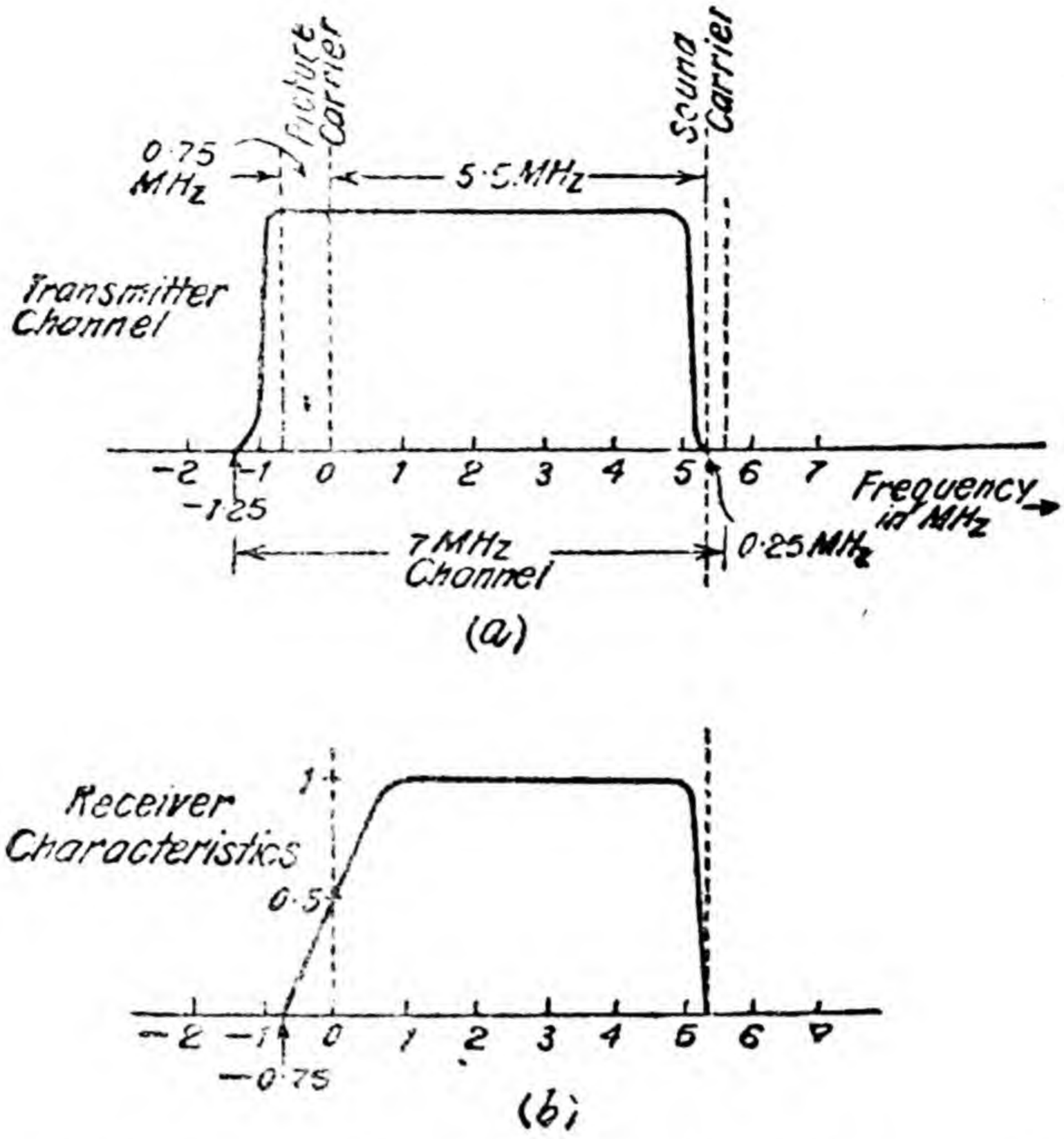


Fig 27. Frequency spectrum of vestigial sideband transmission with response characteristics for receivers used with such transmitters.

Sound and picture signals are transmitted by the same antenna by using a diplexer called picture-sound diplexer discussed below.

Picture-Sound Diplexer : The diplexer is shown in fig. 28.

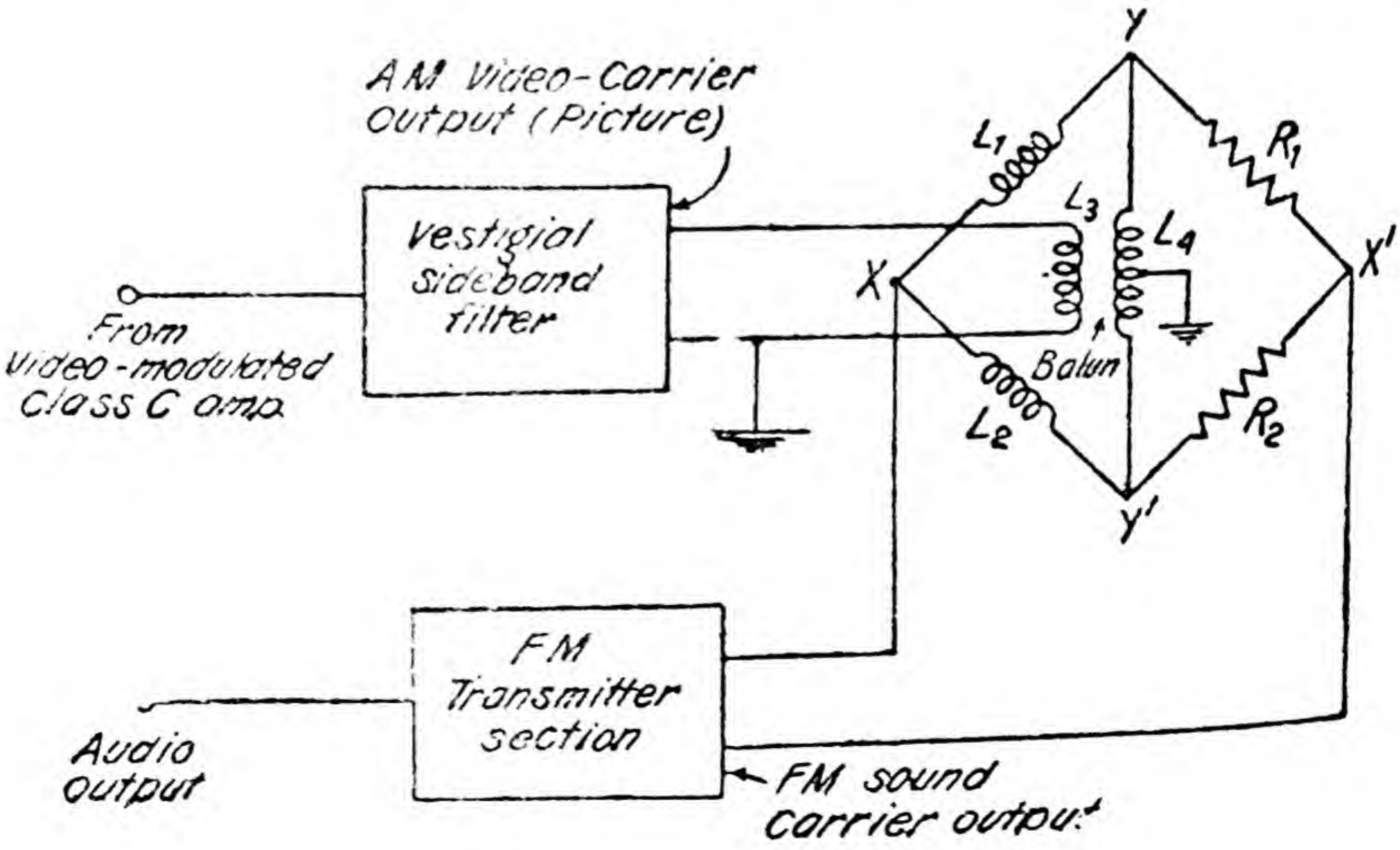


Fig. 28. Diplexer.

This system uses a balanced-bridge circuit to eliminate interaction between the video AM and the sound FM carriers when they are transmitted by a single antenna.

The output from vestigial sideband filter is applied to the transformer consisting coils L_3 and L_4. Since feeding to L_3 is through coaxial cable in which inner conductor is above ground while outer at ground, unbalanced line condition exists. The centre tapped L_4 converts this balanced input to a balanced line section and is called as *balun*.

Bridge circuit consists of two inductive reactances L_1 and L_2 which are coaxial-cable sections. Other two arms consist R_1 and R_2 which represent radiators of the transmitting antenna.

FM sound carrier is impressed across points X and X', while picture signal from the L_4 balun section appears at Y and Y'. With balanced bridge equal voltages exist on both the reactive as well as the resistive arms. Hence, sound modulated carrier is present in both antenna radiators R_1 and R_2. But across points Y and Y' there is zero voltage for sound modulated carrier because any potentials at these points will have the same polarity due to which no current flows through L_4 (from sound modulated carrier). Thus interaction between sound and video-modulated carriers does not occur. On the otherhand, video-modulated carrier, applied at Y and Y' also sets up equal voltages across the arms of the bridge but because of the balanced bridge system no voltage corresponding to this video modulated carrier will appear at X and X'. Therefore again interaction between sound and video modulated carrier is prevented.

In short, both modulated carriers develop equal voltages across arms of the bridge and at the same time no video modulated carrier is present at points at which sound modulated carrier appears and *vice versa*.

33·19. TELEVISION RECEIVER :

A simplified block diagram of a television receiver is shown in fig. 29. Both audio and video signals are subjected to superheterodyne action at the same time. At intermediate frequency section both signals are separated.

A folded dipole is used as a receiving antenna which offers a terminating impedance to the connected transmission line equal to the characteristic impedance to the latter. This possesses a better matching for maximum transference of the received energy.

The next stage is tuned r.f. amplifier with a broad bandpass characteristic. The stage helps in achieving a good noise figure and providing image rejection. The desired television channel is selected by tuned circuits.

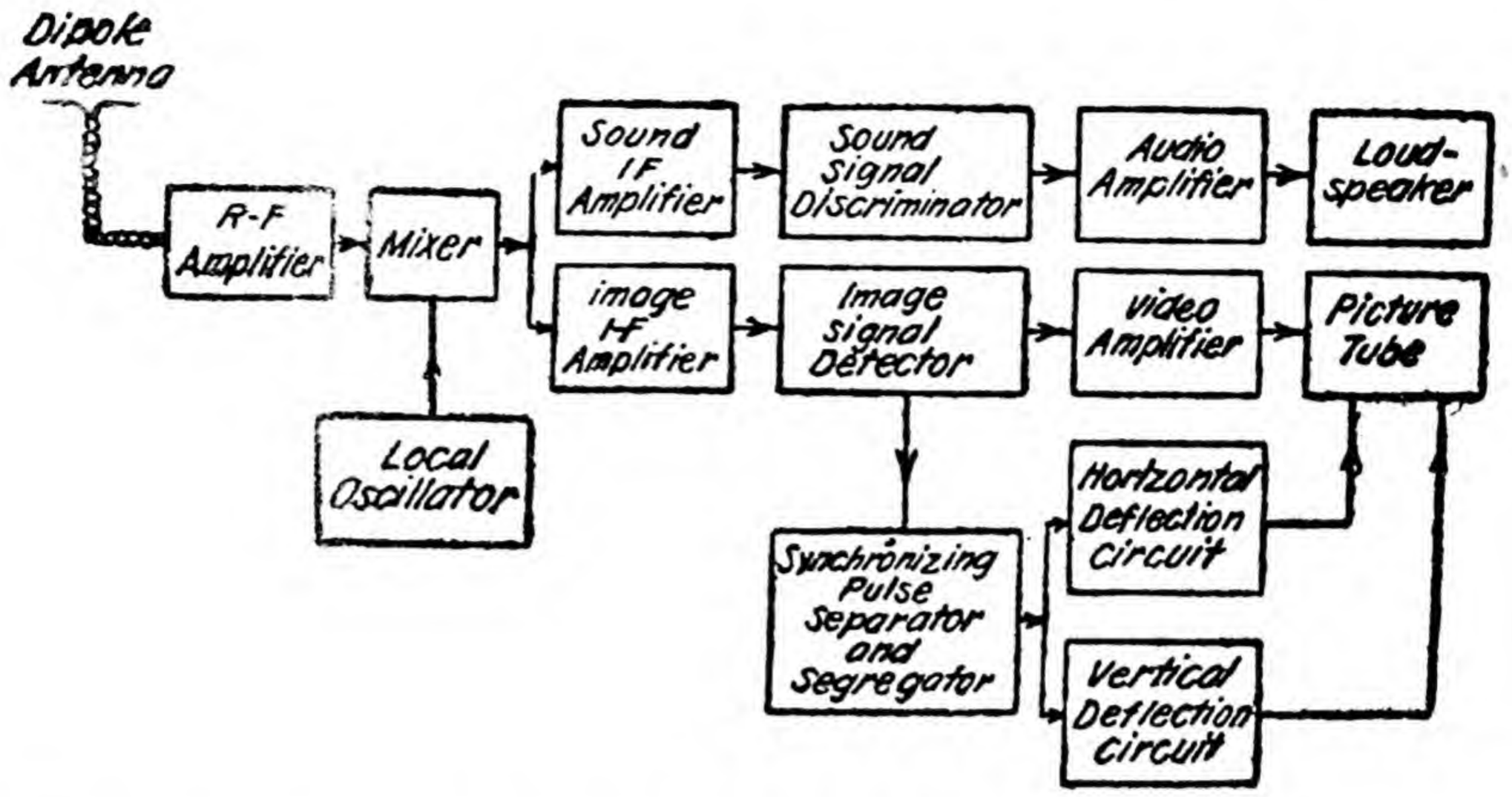


Fig. 29. Simplified block diagram of a television receiver.

The r.f. signal then heterodynes in the mixer with a locally generated frequency to produce a lower intermediate frequency equal to the difference between the two signals.

The output of this stage is applied to two sets of tuned intermediate frequency amplifier sections : one of these is tuned to sound intermediate frequency and the other to image intermediate frequency. The unwanted signals from each channel are removed by the application of rejection filters.

The sound from i.f. signal thus obtained is then recovered by employing the stage of an f.m. receiver. The sound channel consists of an i.f. amplifier, a limiter and discriminator or a ratio detector, one or two stages of audio amplifier and loudspeaker. Image signal detector consists of a diode detector, the output of which will contain video signal and control pulses. This demodulated composite video signal is fed to a 'sync' separator where the synchronising signals are seggregated and separated from the image signals. The image signal is directly applied to the picture-tube and synchronising pulses are applied to beam deflection circuits to keep the scanning of electronic beam at the picture tube in step with that at the camera tube.

We shall now give a description of sync. separator.

Sync. Separator : Sync. separator circuit is shown in fig. 30(*a*). The sync. tips necessary for synchronisation of the vertical and horizontal oscillators, both relaxation type, are extracted from the video signal and blanking levels. The *NPN* transistor, used in the circuit, has no forward bias applied and thus is at or near non-conducting state. If required, some reverse bias may be applied so that *only sync. tips* will have sufficient amplitude to cause conduction and develop the output.

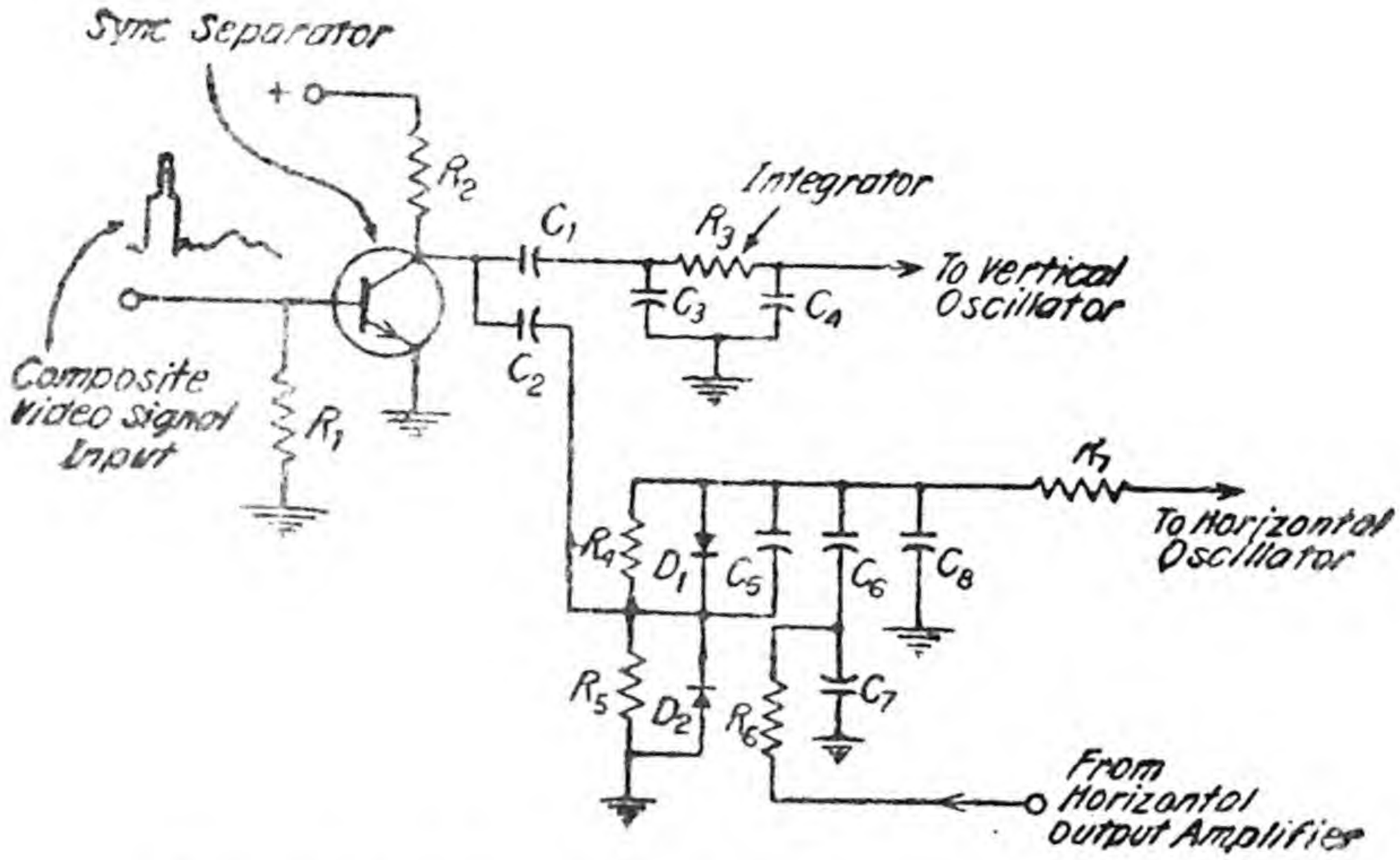


Fig. 30 (*a*). Sync. separatar and output circuitry.

The sync. pulses (equalising pulses, vertical sync. pulses and horizontal sync. pulses, fig. 22) are applied to respective processing networks required for synchronisation of the vertical and horizontal relaxation oscillators.

Now refer to fig. 30 (*b*), which represents sync. tips that are input to sync. separator. We note that both equalising pulses and vertical sync. pulses have the same repitition rate but their durations differ. Thus, while each type pulse will synchronise the horizontal oscillator, only long duration vertical sync. pulse block will trigger vertical oscillator and short duration equalising pulse will remain ineffective as described below :

The integrating circuit consisting R_3, C_3 and C_4 has a long time constant for vertical block pulses. When short duration equalising pulses arrive, then capacitors are charged slightly for each equalising pulse but because the interval between the pulses is longer than pulse duration, these capacitors atorce discharge. Thus this pulse group fails to build up a voltage (fig. 30 *b*) which may trigger the vertical oscillator. On the other hand, vertical sync. pulses have longer duration and smaller time interval between them so that in their case, capacitors do not find time to discharge much. The result is that each vertical sync. pulse will build up voltage so that the voltage built up by the vertical sync. pulse group (fig. 30 *b*) is sufficient to trigger the vertical oscillator. This voltage build up is removed when next set of equalising pulses is input to the separator because now capacitors find time to discharge. Thus integrator becomes ready again to receive the next vertical block.

For horizontal relaxation oscillator, the sync. pulses are applied to the junction of phase detector diodes D_1 and D_2, as shown, per-

mitting both diodes to conduct. A signal from the horizontal output amplifier circuitry is applied to the junction of capacitors C_6 and C_7. Thus a bridge is formed by the diodes and capacitors, and with in-phase voltages appearing at the diode junction and

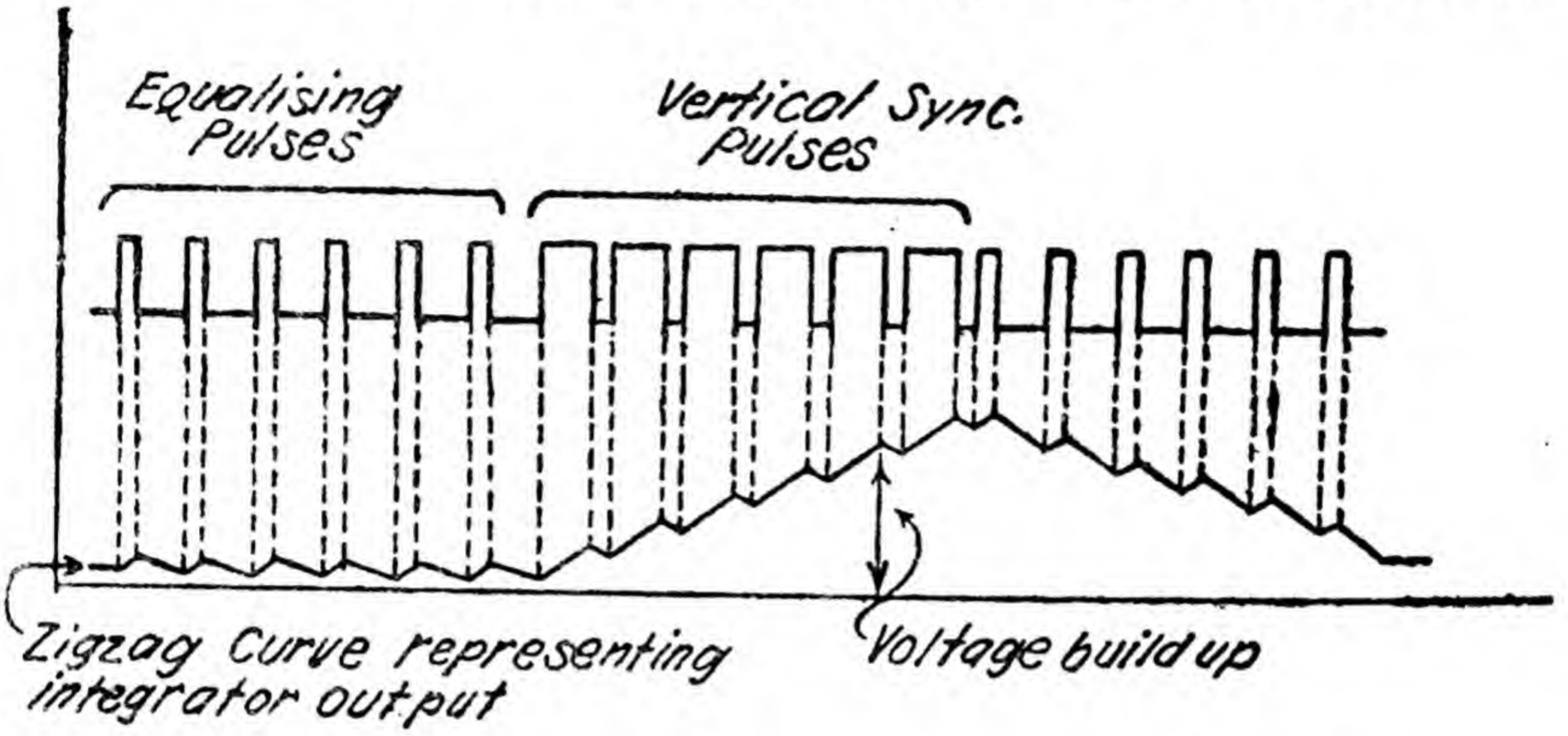
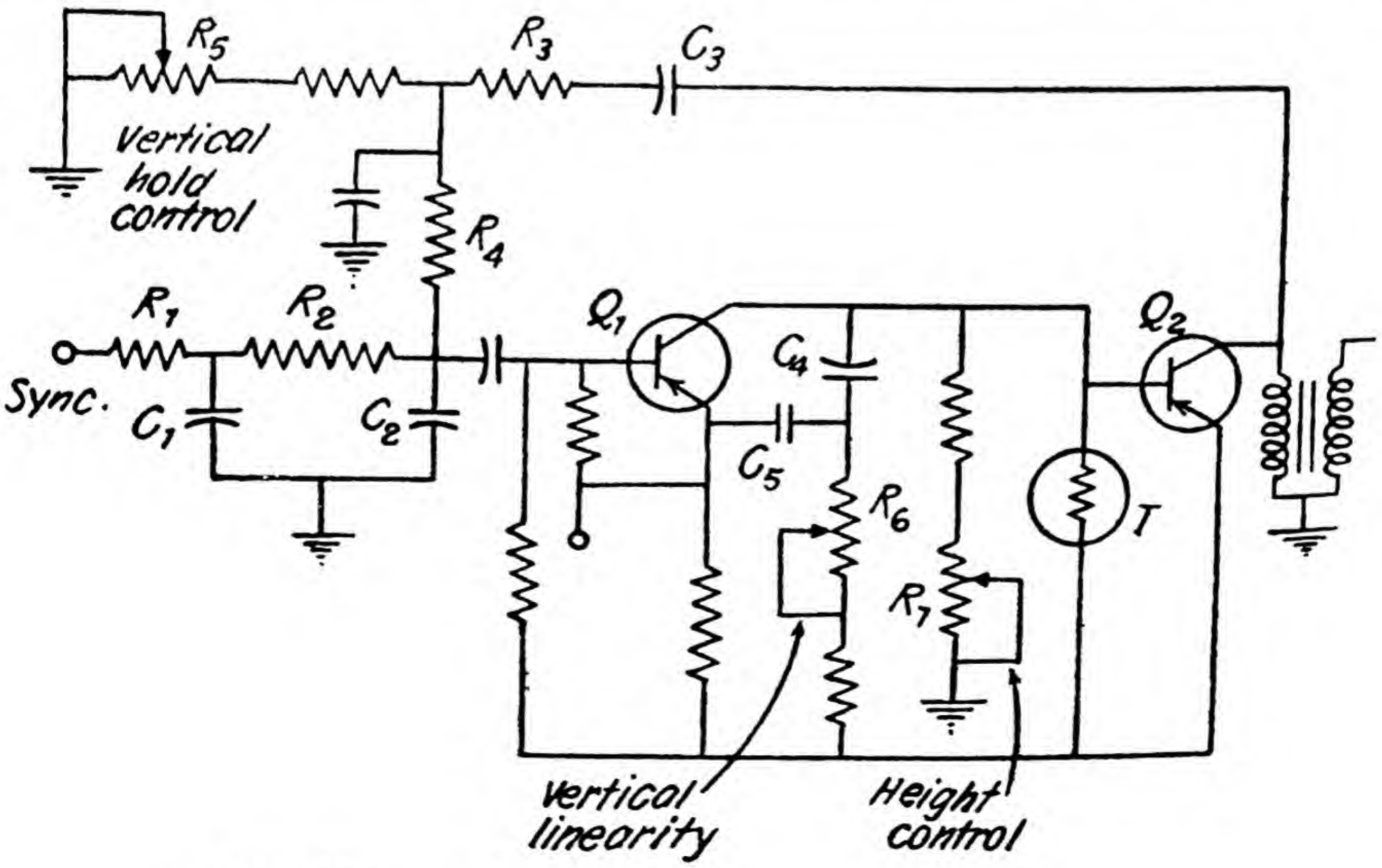


Fig. 30 (*b*). Output of the integrator and input plses.u

capacitor junction, a zero voltage appears across the top and bottom (ground) of the circuitry. This synchronises the horizontal oscillator which then generates the proper frequency. In case it does not generate the desired frequency then the phase of the signal fed back to the bridge phase detector will not coincide with the phase of the incoming sync. pulses and the bridge becomes unbalanced. Consequently, a voltage now appears at the output (across capacitor C_8 and at R_7), which is applied to the input of the horizontal oscillator and corrects drift in the frequency.

Multivibrator Vertical Sweep System : A multivibrator sweep



circuit is shown in fig. 30 (*c*). Synchronising pulses from sync. separator are first made to pass through integrating circuit consisting R_1, C_1, R_2 and C_2. Only vertical synchronising pulses, which are relatively long compared to horizontal synchronising pulses, are able to pass. These incoming vertical synchronising pulses trigger Q_1 and in this way control the frequency of operation of Q_1 and Q_2.

The oscillations of this circuit are maintained through the feedback pulses which travel from the collector of the output stage Q_2 to the base of Q_1 through C_3, R_3 and R_4. Across this feedback path, a resistance R_5, called vertical hold control, is shunted. By varying R_5, total resistance of the feedback path changes which, in turn, causes the frequency of the oscillator to vary within a fairly narrow range. Knob of R_5, is directly mounted on the front panel for ready use of user.

The height control, indicated in the figure, determines the amplitude of the voltage transferred from Q_1 to Q_2, and thus controls the vertical size of the image on the face of the picture tube.

Vertical deflection voltage is developed across C_4. By varying R_6, which is in series with C_4, the shape of the voltage reaching the base of Q_2 is varied, and hence the linearity of the image that is traced out vertically on the picture tube is controlled.

The purpose of thermistor, T, is to maintain the base voltage of Q_2 at the level amplitude. Due to significant current flowing towards output stage, the collector of Q_2 becomes hotter due to continued operation. This, in turn, causes an increase in the average current flowing through Q_2. But as the current tends to increase, the resistance of thermistor, T, will decrease, offsetting this increase in current and maintaining a fairly constant voltage across the thermistor element.

33·20. CHARACTERISTICS OF MONOCHROME TELEVISION SYSTEM ADOPTED IN INDIA

(system *B* of *CCIR**)

A. VIDEO CHARACTERISTICS :

(i) No. of lines per picture ..625

(ii) Field frequency...50 fields/sec.

(iii) Picture frequency...25 pictures/sec.

(iv) Line frequency...15,625 lines/sec. tolerance when operated $\pm 0{\cdot}1\%$ non-synchronously

(v) Aspect ratio (width/height)...4/3

*International Radio Consultative Committee

(vi) Scanning sequence (line)...Left to right
(vii) Scanning sequence (field)...Top to bottom
(viii) Nominal Video Band width...5 mc/s.

B. R. F. CHARACTERISTICS :

(i) Nominal radio frequency bandwidth...7 mc/s
(ii) Sound carrier relative to vision carrier...+5·5 mc/s
(iii) Sound carrier relative to nearest edge of the channel ...—0·25 mc/s
(iv) Nominal width of vestigial side band...0·75 mc/s
(v) Type of vision modulation...AM Negative
(vi) Synchronising level as a % of peak carrier...100%
(vii) Blanking level as a % of peak carrier...(72·5 to 77·5)%
(viii) Difference between black level and blanking level as a % of peak carrier...(0·3 to 6·5)%
(ix) Peak white level as a percent of peak carrier ...(10 to 12·5)%
(x) Type of sound modulation...FM
(xi) Ratio of effective radiated power of vision and sound ...5/1

C. DETAILS OF LINE SYNCHRONISING SIGNALS (Fig. 31 a)

(i) Line period (*H*)...64 μ sec.
(ii) Line blanking interval (a)...18·5 to 19·2% of *H* (*i.e.* of 64 μ sec.) (=11·8 to 12·3 μ sec.)
(iii) Front porch (*c*)...2 to 28% of *H* (=1·0 to 1·8 μ sec.)
(iv) Synchronising pulse (*d*)..,7 to 7·7% of *H* (=4·5 to 4·9 μ sec.)
(v) Build up time (10 to 90%) of edges of line blanking signal (*e*)...0·31 to 0·62% of *H* (=0·2 to 0·4 μ sec.)
(vi) Build up time (10 to 90%) of line sync. pulse (*f*) ...0·31 to 0·62% of *H* (=0·2 to 0·4 μ sec.)

D. DETAILS OF SYNCHRONISING WAVEFORMS :

(i) Field period (*V*)...20 *m* sec. (=20,000 μ sec.)
(ii) Line period (*H*)...64 μ sec.
(iii) Field blanking period (*j*)...[(18 to 22) *H*+12] μ sec.
(iv) Build up time (10 to 90%) of the edges of field blanking pulses...$\leqslant$ 6 μ sec.
(v) Duration of first equalising pulse sequence (*l*)...2·5*H*
(vi) Duration of synchronising pulse sequence (*m*)...2·5 *H*
(vii) Duration of second sequence of equalising pulses (*n*) ...2·5 *H*

(viii) Duration of equalising pulses (p)...3·4 to 3·75% of H (= 2·2 to 2·4 μ sec.)

(ix) Duration of sync. pulses *i.e.* interval between field sync. pulses (r)...7 to 7·7% of H (=4·5 to 4·9 μ sec.)

(x) Build up time (10 to 90%) of the edges of sync. signals (S).. 0·31 to 0·62% of H (=0·2 to 0·4 μ sec,)

E. IDEAL FREQUENCY CHARACTERISTICS OF VISION TRANSMITTER :

Intermediate Frequencies :

Vision—38·9 mc/s

Sound—33·4 mc/s

TV Channels in India :

Channel I	40—47 mc/s
Channel II	47—54 mc/s
Channel III	54—61 mc/s
Channel IV	61—68 mc/s

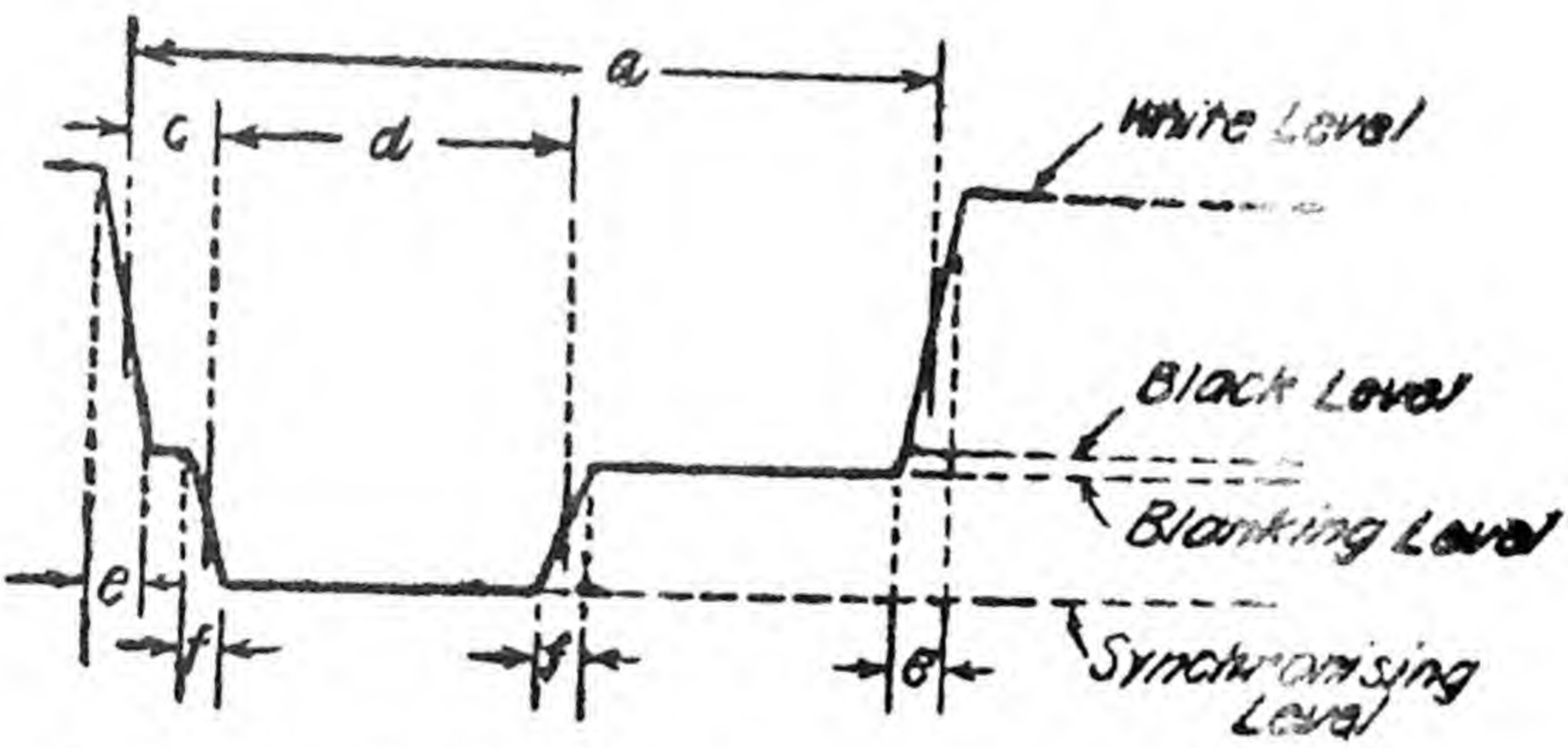


Fig. 31 (a). Details of line synchronising signals, H equals 64 μ sec.

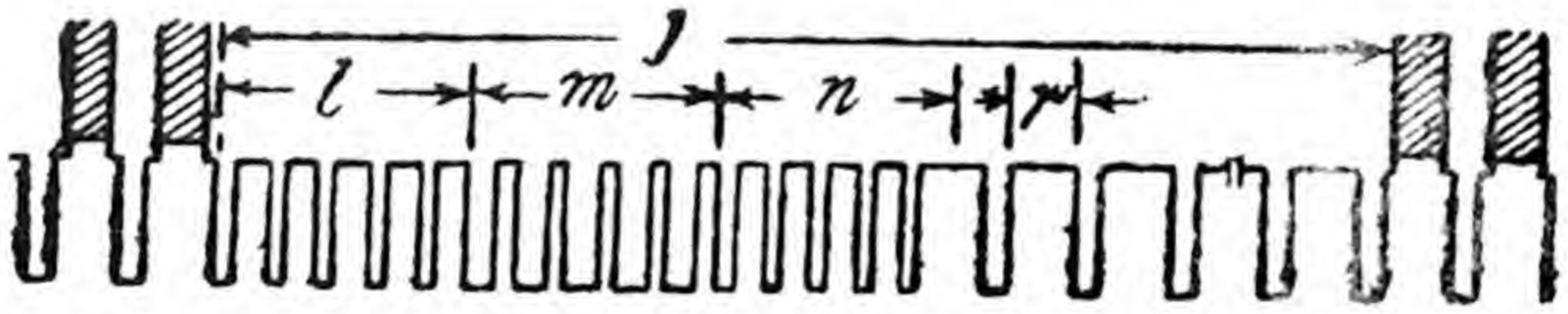


Fig. 31 (b). Signals at the end of even fields.

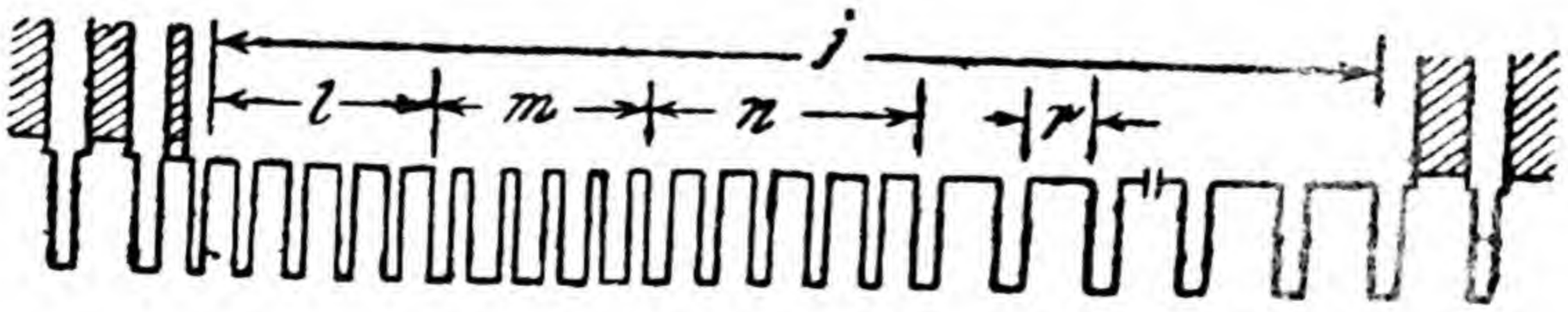


Fig. 31. (c) Signals at the end of odd fields.

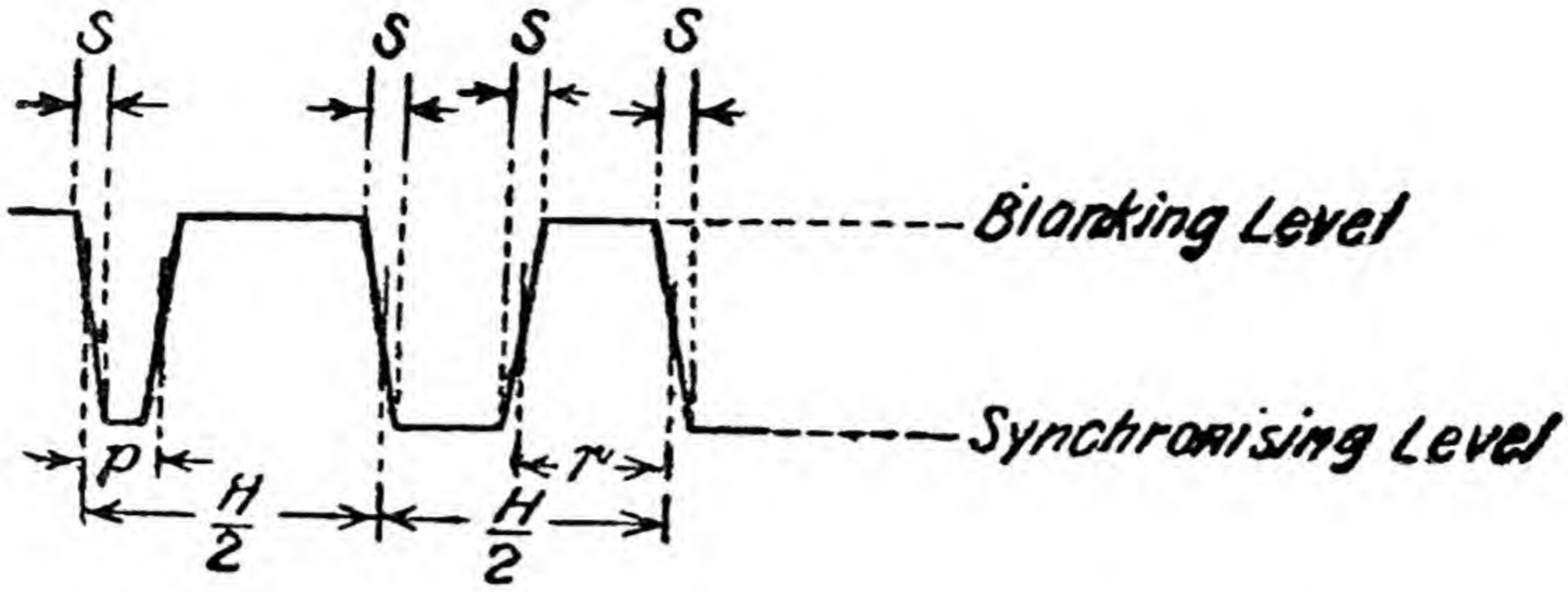


Fig. 31. (d) Details of equalising and sync. pulses.

COLOUR TELEVISION

If the reproduction in colour is desired, some means of transmitting information regarding the colour content of the original scene must be found. Satisfactory colour reproduction can be obtained by making use of three properly chosen primary colours of specified spectral composition. A number of different colours can be formed by the combination of three coloured lights. Experience has indicated that the colours red, blue, and green, when combined with each other in various proportions, will produce a wider range of colours than any other combination of colours. Therefore, red, blue, and green colours are chosen. In this section frequency and other related data refer to U.S.A., TV system.

33·21. COLOUR TV CAMERA :

It is essentially a combination of three basic cameras. Light from the televised scene passes through the standard camera lens system and strikes at blue reflecting dichroic mirror. The red and green components of the incident light will pass through, but the blue light components are deflected to the 'blue camera' where the blue light is converted to a video signal. Next, the main light strikes a red deflecting dichroic mirror, and the red-light components are directed to the 'red camera' and are converted into video signals. Meanwhile, the remainder of the main beam, which is now only the green-light component, passes into 'green camera' and is converted into video waves. Thus light is converted into video waves corresponding to the red, blue and green components of the original colour scene.

33·22. MATRIXING NETWORK :

These three video-signals are not used to modulate the transmitter directly because the bandwidth required for separate transmission of three colour signals would greatly exceed the FCC limi-

tation of 6 mc/s. To solve this problem, modulation system proposed by National Television System Committee (NTSC) was adopted. This system has been so designed that its signal occupies no more than 6 mc/s. (video and sound both signals) and it carries not only the full black and white (or *monochrome*) signal but, in addition, the colour information. Question is : how is all this information compressed into a 6 mc/s. band spread ? The answer is that since a video signal does not occupy every cycle of 4 mc/s, assigned to it, colour information can be inserted within 4 mc/s. Partial spectrum distribution of a monochrome signal is shown in fig. 32 (*a*). It is obvious that video signal (monochrome signal) appears in the form of 'clusters' of energy located near harmonics

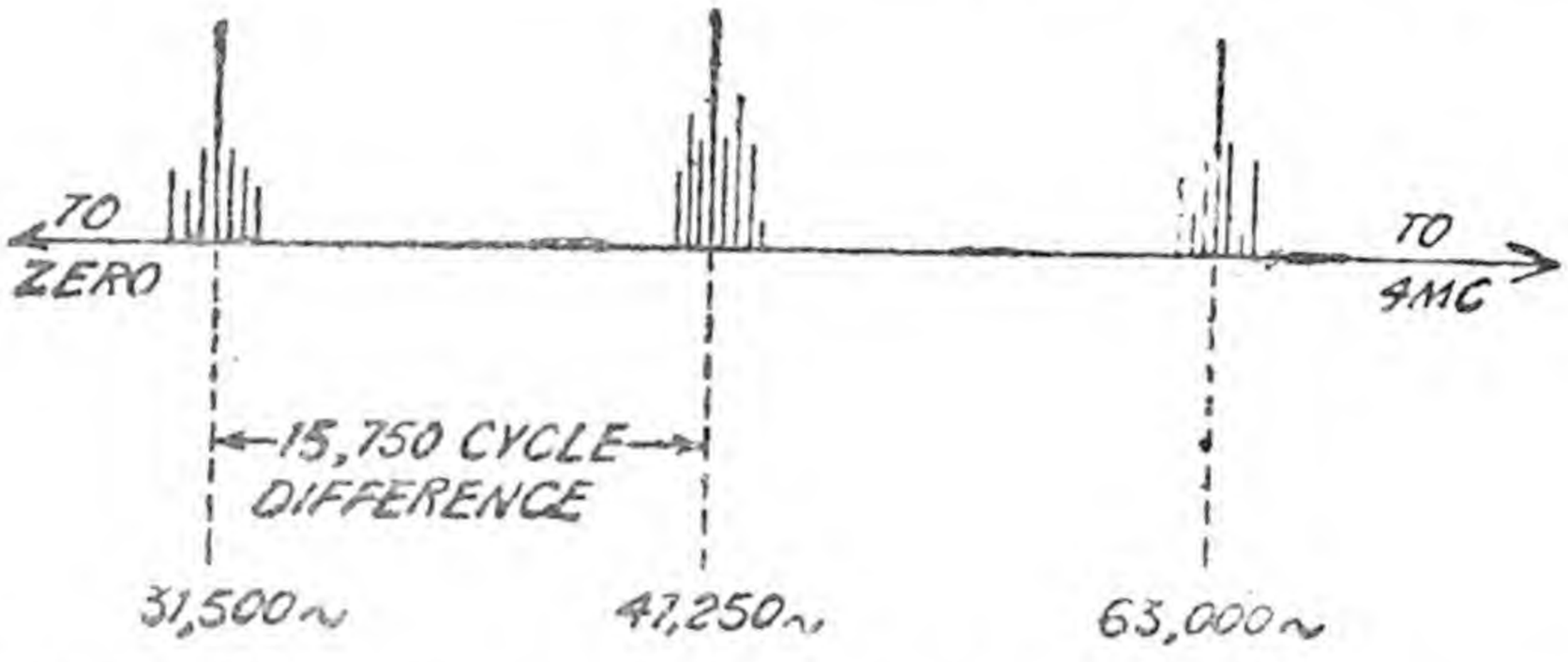


Fig. 32. (a) Partial spectrum distribution of a monochrome signal.

of the 15,750 cycle line. The energy is grounded around these points, which have relatively wide gaps between them. Since these empty spaces are not being used, colour information of NTSC colour television signal is placed in them fig. 32 (*b*). This process of placing the energy of one signal between the clusters of energy

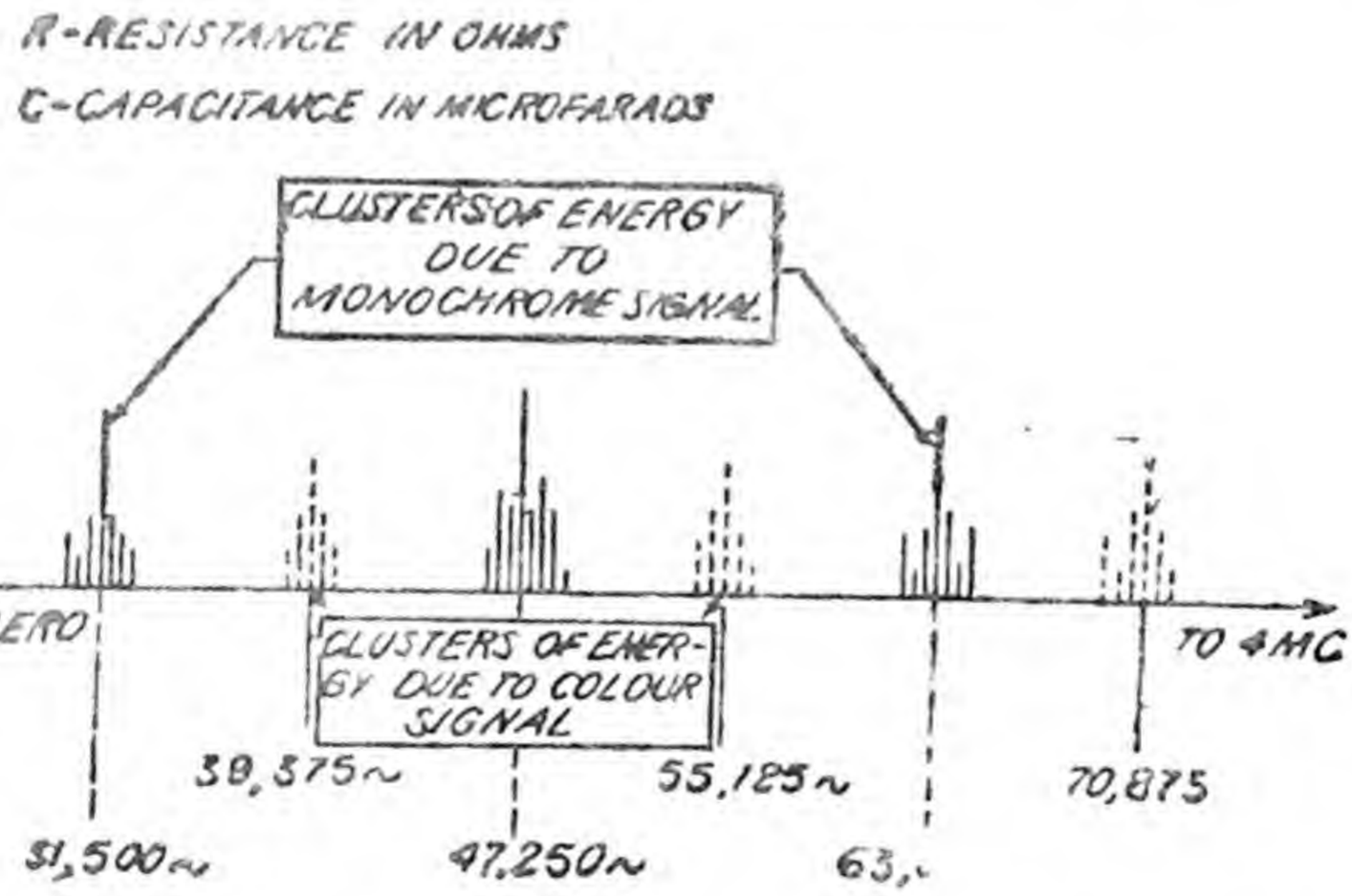


Fig. 32. (b) The energy of colour signal is inserted in the gaps between the energy of monochrome signal.

of another signal is called 'interleaving'. Thus a total colour signal consists of two components—a monochrome signal, and the signal which carries information concerning colour.

(a) The Monochrome Signal : The black-and white or monochrome portion of the total colour signal is equivalent in all respects to present black-and-white signals. It is formed by combining the red, green and blue signals from their respective colour cameras with the help of matrixing network. They are combined in the same proportion as the eye 'response'† to individual colour frequencies, or

$$Y=0{\cdot}59G+0{\cdot}30R+0{\cdot}11B,$$

where Y=a mathematical symbol representing the *monochrome signal*, also called *luminance* or *brightness signal.*

G=green signal, R=red signal, B=blue signal.

Thus monochrome signal is composed of 59 percent of the output of the green camera, 30 percent of output of red camera and 11 percent of the output of the blue camera.

Y signal (luminance signal) modulates the carrier directly. As in a monochrome receiver, this will produce an energy spectrum ranging from 1·25 mc/s. below the r.f. carrier to 4·2 mc/s. above it, as shown in fig. 32 (*a*).

(b) Colour Signal : Colour signal does not modulate the r.f. carrier directly as Y signal does but, instead, they are sent on a carrier that fall between the gaps of 'energy bursts' of r.f. carrier modulated by Y-signal fig. 32 (*b*). Therefore, first step is to select a sub-carrier frequency. Then two colour difference signals, I and Q, (to be explained later on) modulate this sub-carrier.

(*i*) *Choice of sub-carrier* : The sub-carrier frequency is approximately 3·58 mc/s. (the 455th multiple 15750/2=7875 c/s.) If we use an odd multiple (1, 3, 5 etc.) of 7875 c/s. as a sub-carrier, then this frequency will fall midway between the harmonics of 15,750 cycles. If we use even multiples of 7875 c/s. as sub-carrier frequency we would end up with 15,750 cycles or more of its harmonics, and this would place the colour signal at the same points throughout the band as those occupied by monochrome signal (Y), Therefore to avoid this overlapping between colour and Y-signals, we use an *odd* multiple of 7875 c/s. so that sub-carrier frequency which is modulated by two colour difference signals, falls in between the bundles of energy produced by the first signal (Y).

(*ii*) *Colour difference signals* : We have used above the word 'colour difference signals'. Let us explain its meaning now.

---

†If we look at each light separately, green will appear to be twice as bright as the red and from 7 to 10 times as bright as the blue. This is because eye is more sensitive to green than to red and more sensitive to red than to blue.

In fact, we do not modulate the sub carrier with three different voltages R, G and B but can do the same job by using only two quantities if we resort to the following modification : Take R, G and B voltages, and combine each with a portion of the monochrome signal Y, after the later has been inverted to 180 deg. This produces R-Y, G-Y and B-Y signals. The circuit to achieve it is shown in figure 33. A portion of Y-signal is taken, it is then passed to a low pass filter so that low frequency component are allowed to pass through. This is desired because colour signals are also con-

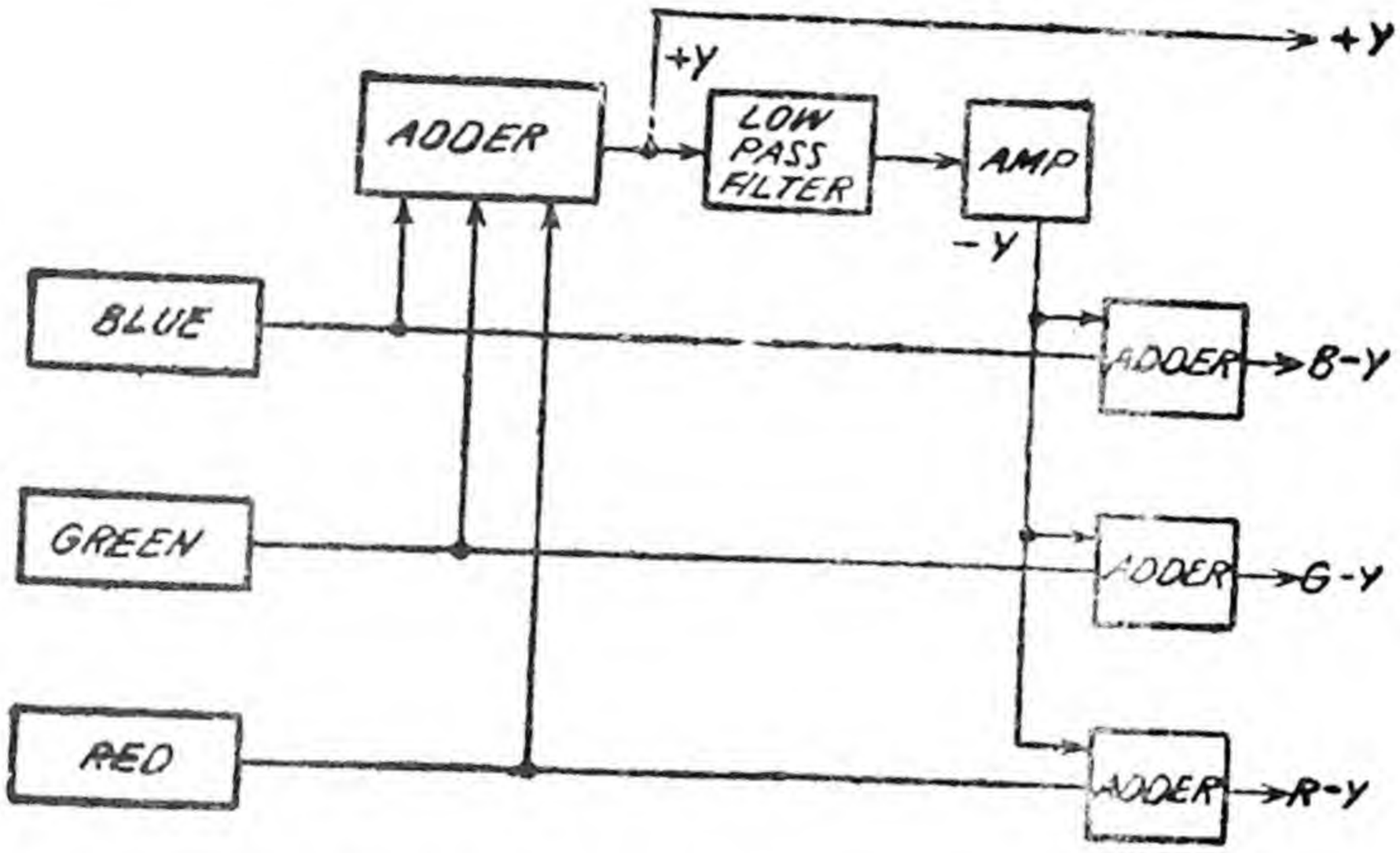


Fig. 33. Circuit to obtain colour difference signals.

cerned only with low frequencies. Low frequency components of Y-signal are then applied to an amplifier. At the output of amplifier Y signal will be inverted to give—Y (low frequency components), $-Y$ is then added to each of the three colours G, R, and B to give G, Y, a R-Y, and B-Y signal. Each of them is called a *colour difference signal* (At the receiver original R, G and B can be obtained by adding Y to G-Y, R-Y and B-Y). Once this is done, it turns out that instead of requiring the three colour difference signals only two, say R-Y and B-Y, can serve the purpose of three. This is because G information is already present in brightness signal, Y, since $Y=0{\cdot}59G+0{\cdot}30R+0{\cdot}11B$. Hence if we send along only R-Y and B-Y in the colour signal to the receiver, we can use these two to obtain the third G-Y information.

(*iii*) *Modulation by colour difference signals* : These two colour difference signals R-Y and B-Y, then modulate the 3·58 mc/s. subcarrier separately R-Y and 3·58 mc/s. carrier are applied to one modulator circuit; whereas B-Y and 3·58 mc/s carrier to another modulator circuit. The output of each modulator consists of 3·58 mc/s. carrier and a series of side bands. The modulation intelligence is contained in the series of side bands that extend above and below 3·85 mc/s. The extent by which side band extend above and below 3·85 mc/s. depends on the band of frequencies

contained in *R-Y* and *B-Y* modulating voltages. It was discovered that the eye is satisfied by the colour image produced if we include colour information upto 1·5 mc/s., while the portion of image from 1·5 mc/s. to 4·0 mc/s is reduced in black and white. Hence side band frequencies of *R-Y* and *B-Y* modulating voltages need extend only from 0 to 1·5 mc/s. Since the frequency range of side

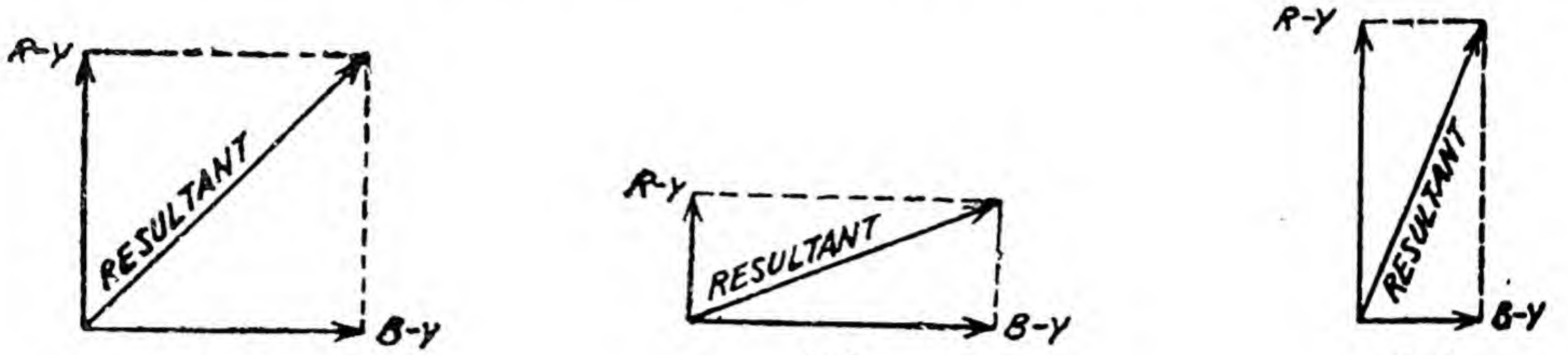


Fig. 24. Resultant carrier for various amplitudes of *B-Y* and *R-Y*.

bands due to two modulating voltages is the same, means should be found to prevent this interference. For this, subcarrier (3·58 mc/s) for one of the signals is first shifted to 90° before modulation. Therefore, side bands of *B-Y* and *R-Y* would then be 90° apart and when they add vectorially, they produce a resultant that will shift towards the stronger signal as illustrated in fig. 24. Thus the phase angle of the resultant will be governed by the colouring or hue* of the picture where as the amplitude (or length) of the vector determine saturation† of the colours.

It has been investigated that human eye sees a full colour range only when the area of object is relatively large. The capability of eye to distinguish between different colours decreases as the size of the object decreases and consequently, only two colours will do well. That is, whereas for large objects eye required three colours in medium size areas eye would need only a limited range of colours which can be achieved by two colours and their various combinations. In view of this property of eye, we now revise our previous statement that side band frequencies of colour modulating voltages *R-Y* and *B-Y* need extend only from 0 to 1·5 mc/s. Because for large objects or areas, say which are produced by video frequencies upto 0·5 mc/s. three primary colours are required and for medium size objects, say, those produced by video frequencies from 0·5 to 1·5 mc/s. only two primary colours need be employed. Taking the advantage of this situation, we need two colour signals: one which has a band pass only upto 0·5 mc/s and other which has

---

*Hue represents colour, such as red, green and orange. The term is associated with colour wavelength and when we call a certain colour green, or orange, or red, we are specifying its hue.

†Saturation tells us how 'deep' colour is. If the colour is highly saturated, we say that it is a deep colour.

a band pass from 0 to 1·5 mc/s. We shall choose them as Q and I colour difference signals as described below.

(*iv*) *Composition of colours : Q and I colour difference signals :* We have already stated that by the composition of the colours, other colours can be obtained. Therefore, the next problem is to determine the composition of these two colour signals R-Y and B-Y. We know that

$$Y = 0{\cdot}59G + 0{\cdot}30R + 0{\cdot}11B$$

so that
$$R\text{-}Y = R - (0{\cdot}59G + 0{\cdot}30R + 0{\cdot}11B)$$
$$= 0{\cdot}70R - 0{\cdot}59G - 0{\cdot}11B$$

and
$$B\text{-}Y = B - (0{\cdot}59G + 0{\cdot}30R + 0{\cdot}11B)$$
$$= 0{\cdot}89B - 0{\cdot}59G - 0{\cdot}30R,$$

which show that R-Y and B-Y vector contain R, G and B voltages in proportions shown. Suppose the colour camera is scanning a scene containing only red. Then no green and blue voltages will be present and R-Y signal becomes $0{\cdot}70R$, while B-Y signal is reduced to $-0{\cdot}30R$. Then the position of resultant vector will be as shown in fig. 35. Similarly, we can obtain the position that the resultant vector occupies when only green or blue or any other colour formed by combining the three primary colours, is being sent. Various colours are shown in fig. 36, which also predicts that how the phase of the colour subcarrier depends upon the colour to be sent (transmitted). As mentioned previously, hue and saturation determine the phase angle and amplitude (length) of the resultant signal.

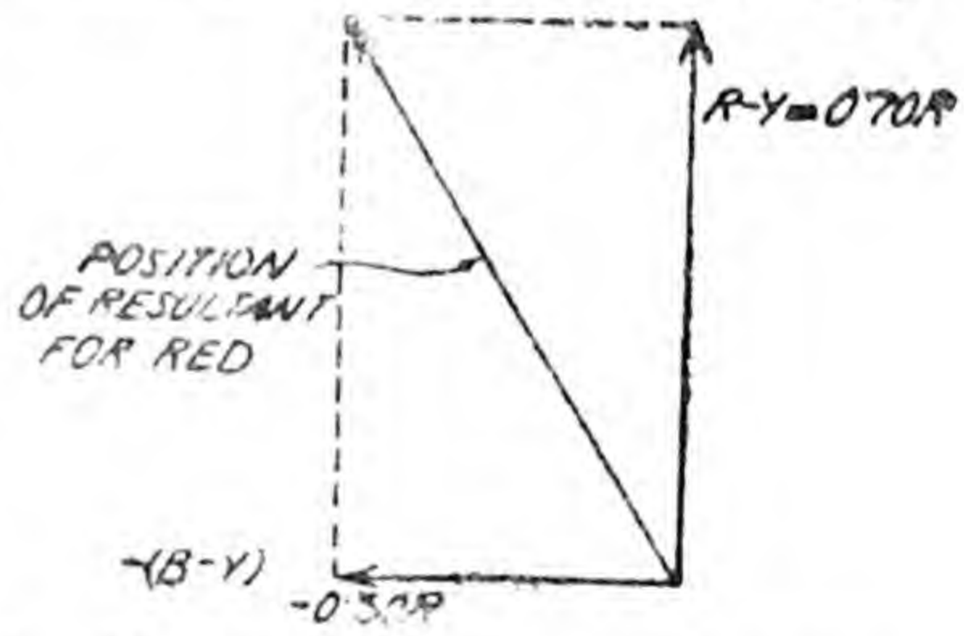


Fig. 35. Position of signal vector when only red field is being scanned.

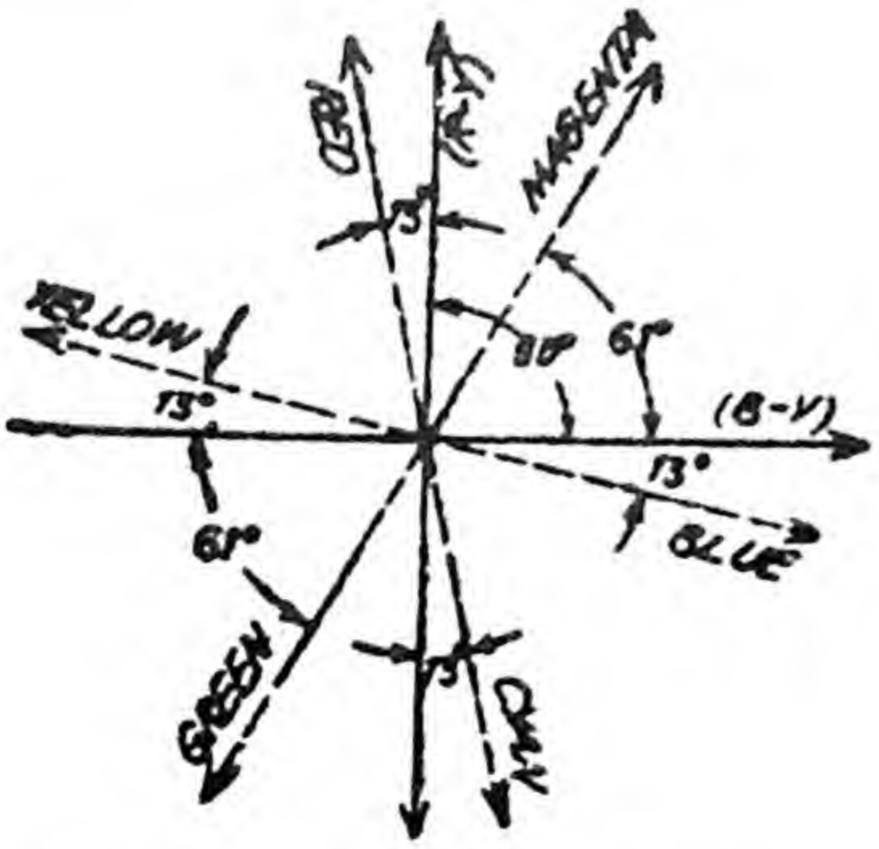


Fig 36. Phase of colour sub-carrier.

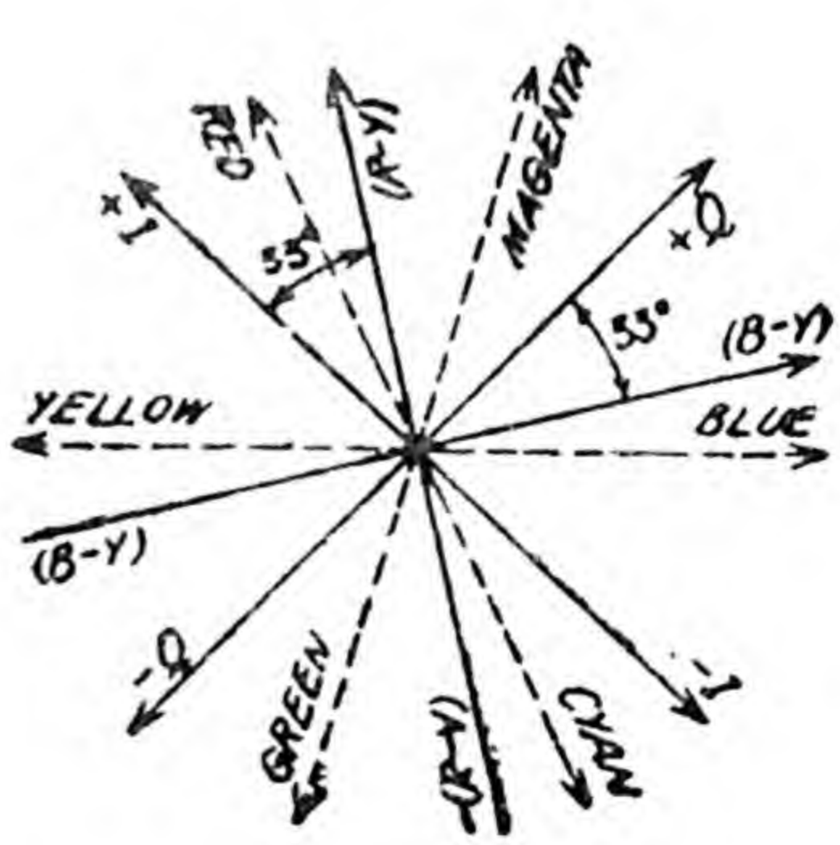


Fig. 37. Position of I and Q signals with respect to R-Y and B-Y

The designers of NTSC system found that operation of the system is improved if they choose two other signals labelled as *I* and *Q* (which are situated not far from *R-Y* and *B-Y* signals, see fig. 37) to modulate 3·58 mc/s subcarrier instead of *R-Y* and *B-Y* signals. Furthermore, *Q* signal possesses frequencies upto 0·5 mc/s, whereas the *I* signal is permitted to have sidebands upto 1·5 mc/s.

(*v*) *Advantage of using I and Q signals instead of R-Y and B-Y :*

Upto 0·5 mc/s both *I* and *Q* are active (*I* possesses frequencies from 0 to 1·5 mc and *Q* upto 0·5 mc) and as they are 90° apart (after modulating subcarrier 3·58 mc), they will produce, in combination with each other, all colours (shown in fig. 37) exactly as *R-Y* and *B-Y* acted (fig. 34). But the situation is different for sideband frequencies from 0·5 mc to 1·5 mc., because for this range only *I* signal is left to produce colour on the picture tube screen. Refer to fig. 37, we find that the positive value of the *I* signal will produce colour between yellow and red, or actually a reddish-orange, and negative values of *I* will produce colours between blue and cyan, or, in general the bluish-green range. Hence *I* signal alone is capable of producing colours from reddish-orange to bluish green. Since eye is sensitive, for medium sized objects (those produced by video signals from 0·5 mc. to 1·5 mc.), to bluish-green or reddish orange, I signal thus obviously takes the advantage of this property of eye.

33·23. SIMPLIFIED COLOUR CHANNEL OF A TV TRANSMITTER :

A simplified colour portion of a TV Transmitter is shown in figure 38.

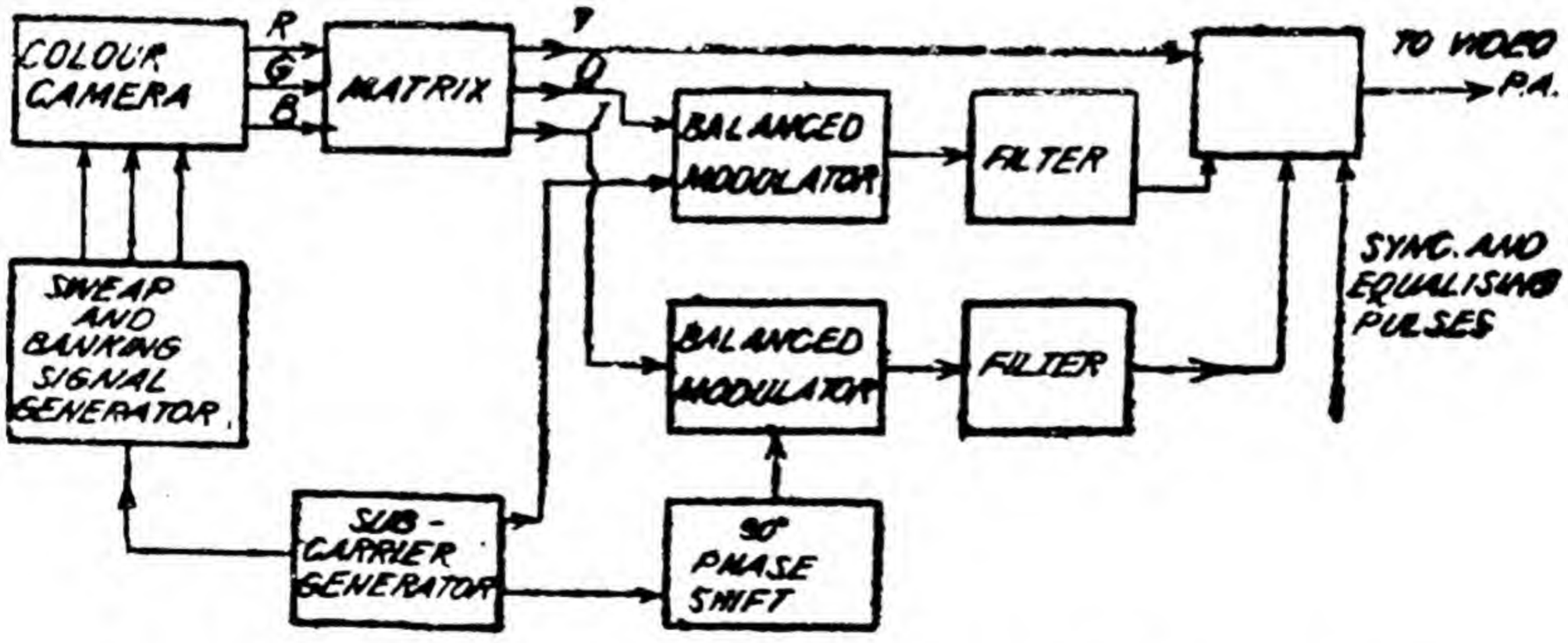


Fig. 38. Simplified colour portion of a TV transmitter.

There are following operations :

(*i*) Monochrome signal, *Y*, directly modulates the carrier; the components extend from 0 to 4 mc/s.

(*ii*) The two colour difference signals *I* and *Q* modulate a subcarrier whose frequency is set at 3·58 mc/s.

Q signal has colour frequencies that extend from 0 to 0·5 mc. This means that when it modulates the subcarrier, upper sideband

extends from 3·58 mc. upto 3·58+0·5, or 4·08 mc. The lower *Q* sideband goes from 3·58 mc. down to 3·58—0·5 or 3·08 mc.

I signal has colour frequencis from 0 to 1·5 mc. so that, when it modulates the subcarrier, lower sideband would extend from 3·58 mc. down to 3·58—1·5, or 2·08 mc. Upper sideband should extend from 3·58 mc. to 3·58+1·5 or 5·08 mc. But we know that band spread for both signals—video and audio in NTSC system is 6 mc. This means upper side band of *I* signal at 5·08 mc., alongwith audio signal would need a band spread of more than 6 mc. Therefore, to keep upper side band within the limits of 6 mc., the *I* signal colour frequency is limited to 0·6 mc. so that, after modulation, upper sideband would extend upto 3·58+0·6 or 4 .2 mc. The video side passband then ends sharply at 4·5 mc. (fig. 39)

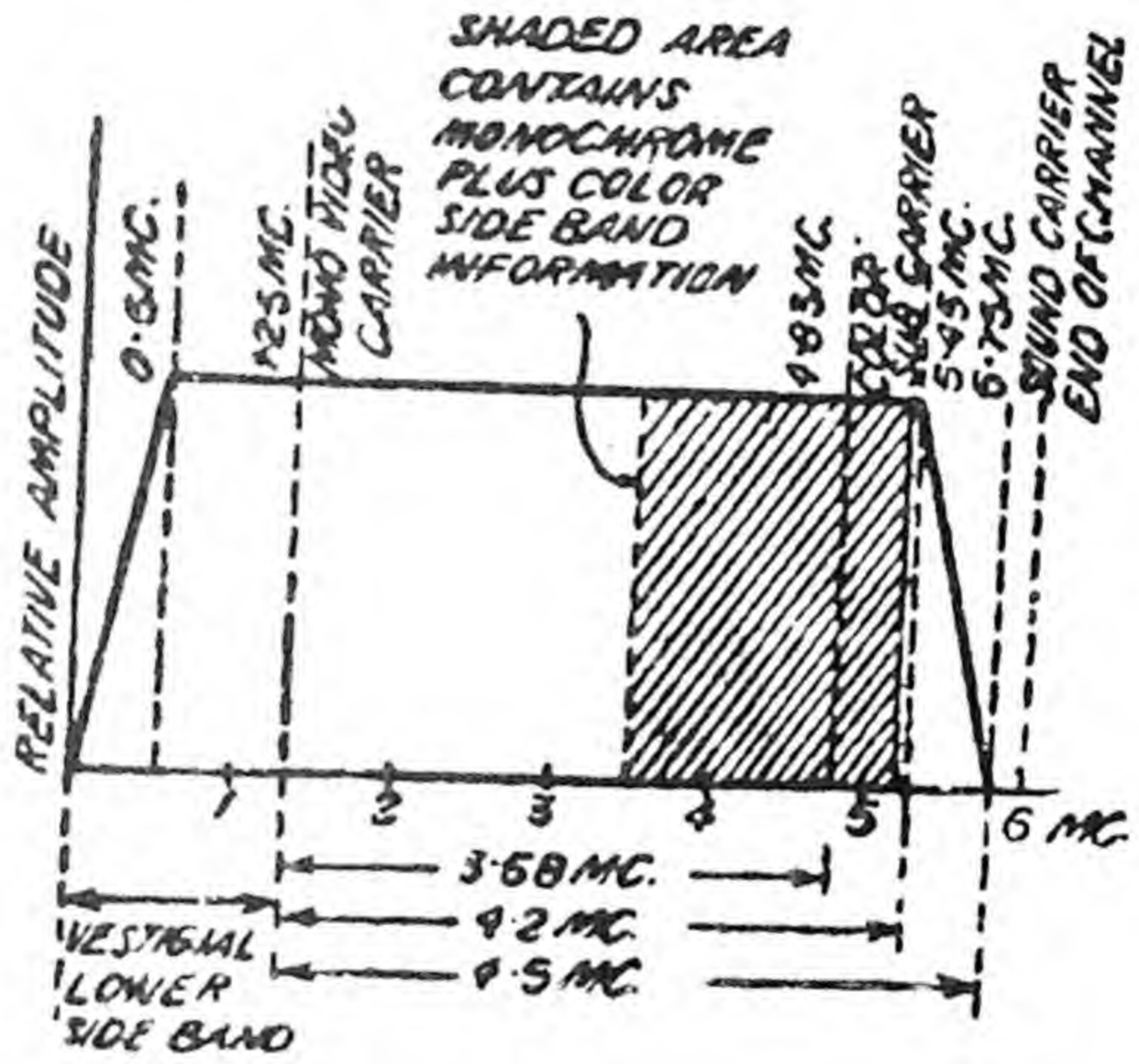


Fig. 39. Distribution of full colour signal within its alloted band.

(*iii*) The process of modulation of subcarrier by *I* and *Q* signals is performed by balanced modulators. Subcarrier to one of the balanced modulators is applied after shifting its phase by 90° which is necessary as explained earlier. Balanced modulator suppresses the colour subcarrier and only sidebands are sent through. Subcarrier suppression has following advantages and disadvantages :

Advantages :

(*a*) It reduces the formation of a 0·92 mc. beat note between the subcarrier (3·58 mc.) and 4·5 mc. sound carrier. This beat note, if present, appear as a series of interference lines on the face of the picture tube.

(*b*) It leads to an automatic removal of the entire colour signal when the televised scene is to be sent wholly as a black-and-

white signal. When this occurs, *I* and *Q* decrease to zero, and since the balanced modulators suppress the carrier, no colour signal at all is developed.

Disadvantages :

To demodulate, *I* and *Q* sidebands at the receiver, it is essential to reinsert subcarrier. Consequently, a 3·58 mc. generator is essential at the receiver stage which would supply subcarrier to *I* and *Q* demodulators. Note that the subcarrier is shifted to 90° before it is applied to the *I* demodulator. This is done to compensate for the original 90° shift of subcarrier when applied to *I* modulator at the transmitter.

(*iv*) To reproduce the subcarrier at the receiver with exact frequency and phase, information is sent as a 'colour burst' along with the signal. Burst contains atleast eight cycles of subcarrier and is phased in step with the colour subcarrier used at the station. This burst is superimposed on the back porch of each horizontal synchronising pulse. At the receiver this burst is used to lock in the frequency and phase of a 3·58 mc. generator, and thus ensures the exact frequency and phase of the subcarrier.

33·24. COLOUR TV RECEIVER :

A simplified diagram of a receiver video and chroma section is shown in figure 40. We shall describe few points regarding the operation in brief as under :

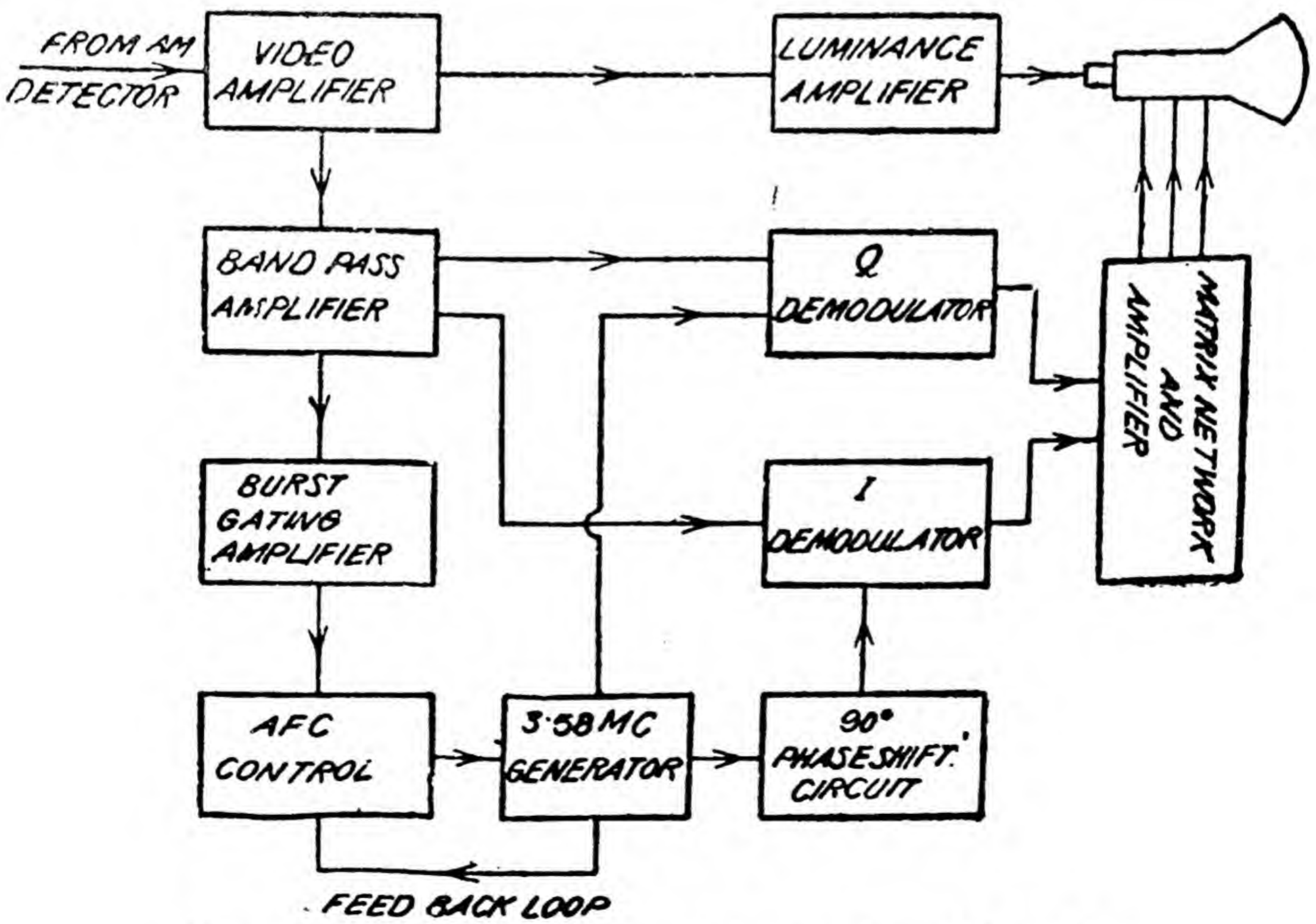


Fig. 40. Simplified diagram of a receiver video and chroma section.

(*i*) The sound section (IF amplifier, FM detector, and audio amplifier) is identical to that of a black-and-white receiver except, inorder to reduce possible interference, sound-channel take-off is affected at the output of the composite IF amplifier and a separate diode mixer is used to produce the 4·5 mc/s sound *IF*.

(*ii*) Horizontal and vertical deflection circuits are also quite similar in the two receivers.

(*iii*) The main difference in colour receiver and black-and-white receiver is that the video amplifier section is expanded to include the chroma circuitry.

As shown in figure, after video amplifier, *Y* signal (luminance signal or composite signal or brightness signal) is separated from colour signals *I* and *Q*. Let us describe it in two separate points :

(*a*) *Y* signal : *Y* signal is passed into the luminance block for further amplification and is finally fed to the three cathodes of a three-gun cathode ray tube (fig 41). This signal, of itself, will produce a complete black-and-white picture.

(*b*) *I* and *Q* signals : As shown in fig. 40, the output of first video block is also fed to a band pass filter and amplifier. Through this stage, only sidebands (of modulated subcarrier by *I* and *Q* signals) and subcarrier bursts (sent to keep phase and frequency of subcarrier generated at the receiver by 3·58 mc. oscillator exactly the same as that of the subcarrier at the transmitter) will pass. These signals are fed simultaneously to the *Q* demodulator, *I* demodulator, and the burst gating amplifier.

Q and *I* demodulators produce *Q* and *I* video signals. Note that subcarrier to *I* demodulator is shifted in phase by 90° as compared to that applied to *Q* demodulator. This is to compensate for the original 90° phase shift as the transmitter. The matrix network converts these *I* and *Q* video signals inputs into *R-Y*, *G-Y* and *B-Y* signals which are then fed to the separate grids of three-gun tube (fig. 41). The combination of *Y* signals on the cathodes, and *R-Y*,

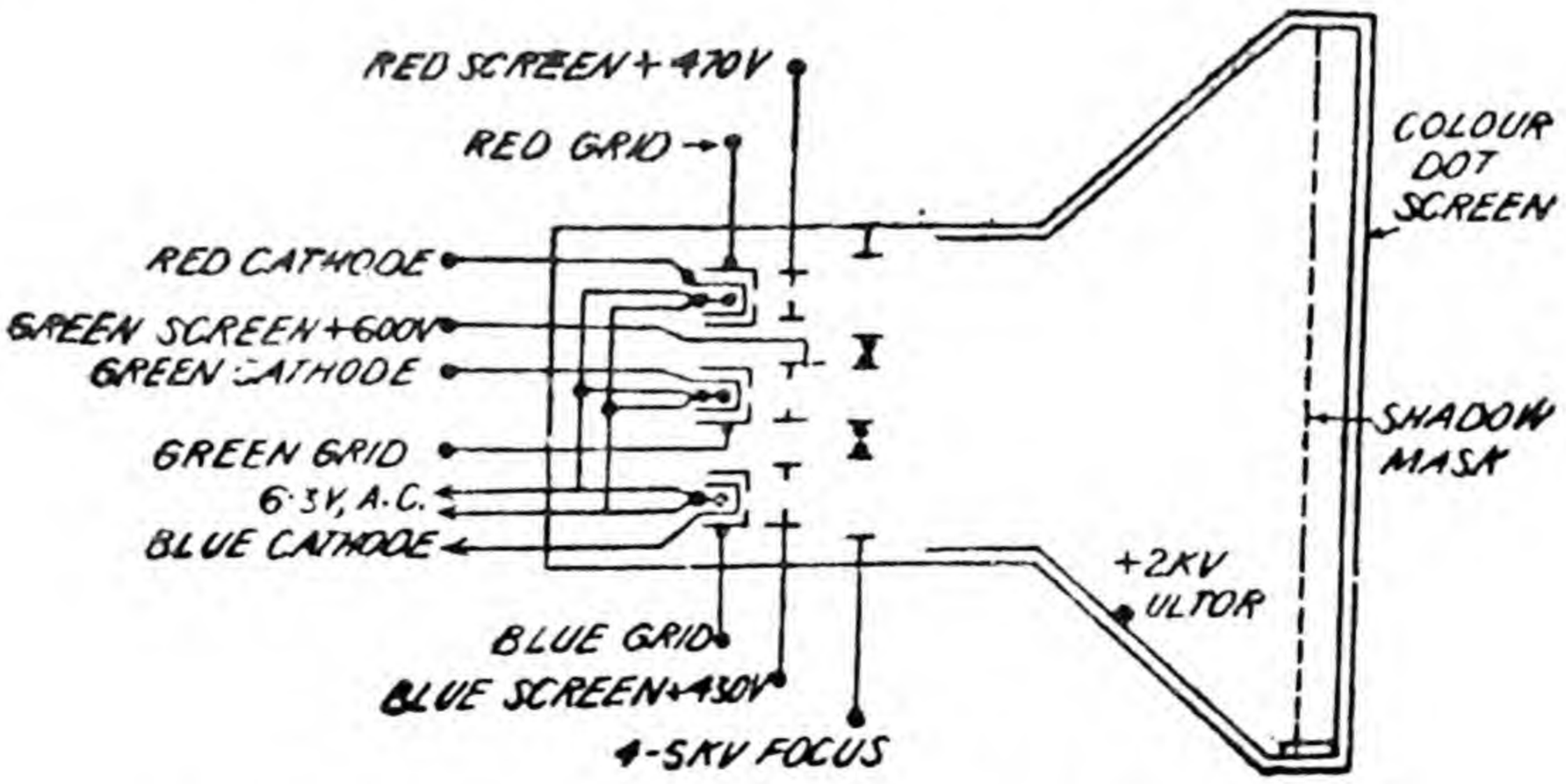


Fig. 41. Basic elements of a colour picture tube.

G-Y and *B-Y* signals on the grids produce three electron beams that are dependent on the amplitude of the red, green and blue components of the televised scene.

Burst gating amplifier conducts only during the burst intervals and its output is used (through AFC and feed back circuits) to maintain correct phase and frequency of 3·58 mc. subcarrier generator.

(*iv*) *Picture tube* : The screen of a colour picture tube consists of a dot pattern of three types of phosphorous (fig. 42) that emit green, blue, and red light, respectively, when excited by the electron beam. For each phosphor colour there is a separate electron beam, originated and controlled by its respective electron gun. All the three electrons beams are focussed and accelerated together and they are deflected by a single magnetic deflection yoke.

These phosphorescent dots which produce coloured light are arranged on the screen in an orderly array of small triangular groups, each group containing a green-emitting dot, a red emitting dot and a blue emitting dot. The actual number of such dots, for a 21 inch screen is about 1,071,000. With such a number of dots, there are 357,000 trios. Each dot has a diameter of approximately 16 mils. If all the three dots in a group are bombarded at the same time, the combined red, green, and blue-light output will present one colour to the observer's eye.

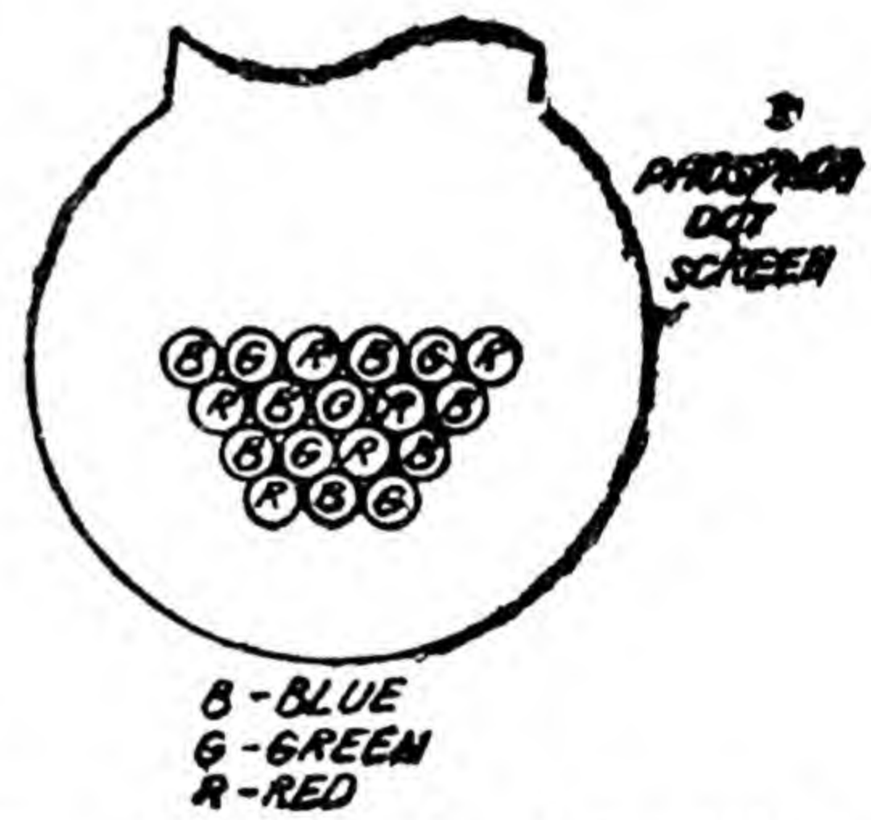


Fig. 42. Phosphor dot screen.

EXERCISES

1. Discuss how the azimuth and range of a target is measured with the help of *PPI*. Calculate the maximum range of a radar set.

2. Describe in full the radar transmitting and receiving systems with reference to radar antennas.

3. Explain the principle of picture transmission. Describe the transmitting and receiving components and their functions in a TV set.

4. Describe, in detail, the mechanism of colour picture transmission. Explain and sketch the block diagram of a colour TV.

5. Explain what do you understand by blanking and synchronising pulses? What are the functions of these pulses in TV? What is a front porch? Give briefly a method of arranging a front porch of 1·25 μ sec. Draw a neat sketch of the composite video signal.

6. Describe the construction of a TV picture tube. How is magnetic deflection obtained in such a tube? Discuss the various controls provided, giving necessary circuit diagrams.

7. Describe, sketching all relevent details, the composite video signal for TV transmission. Explain how the level of background brightness is maintained at the same level for each segment of the signal.

8. Describe, by a block diagram, the general working of a TV receiver. Explain, showing a special circuit, the operation of the sweep circuit and enumerate its special features.

34 ELEMENTS OF MASERS AND LASERS

Maser is acronym for *microwave amplification by stimulated emission of radiation* and laser is acronym for *light amplification by stimulated emission of radiation*. In the laser, the maser principle has been extended to optical or visible light frequencies (of the order of 10^{14} or 10^{15} cycles per second) and is, therefore, also termed as optical maser. As shown in fig. 1, when stimulated

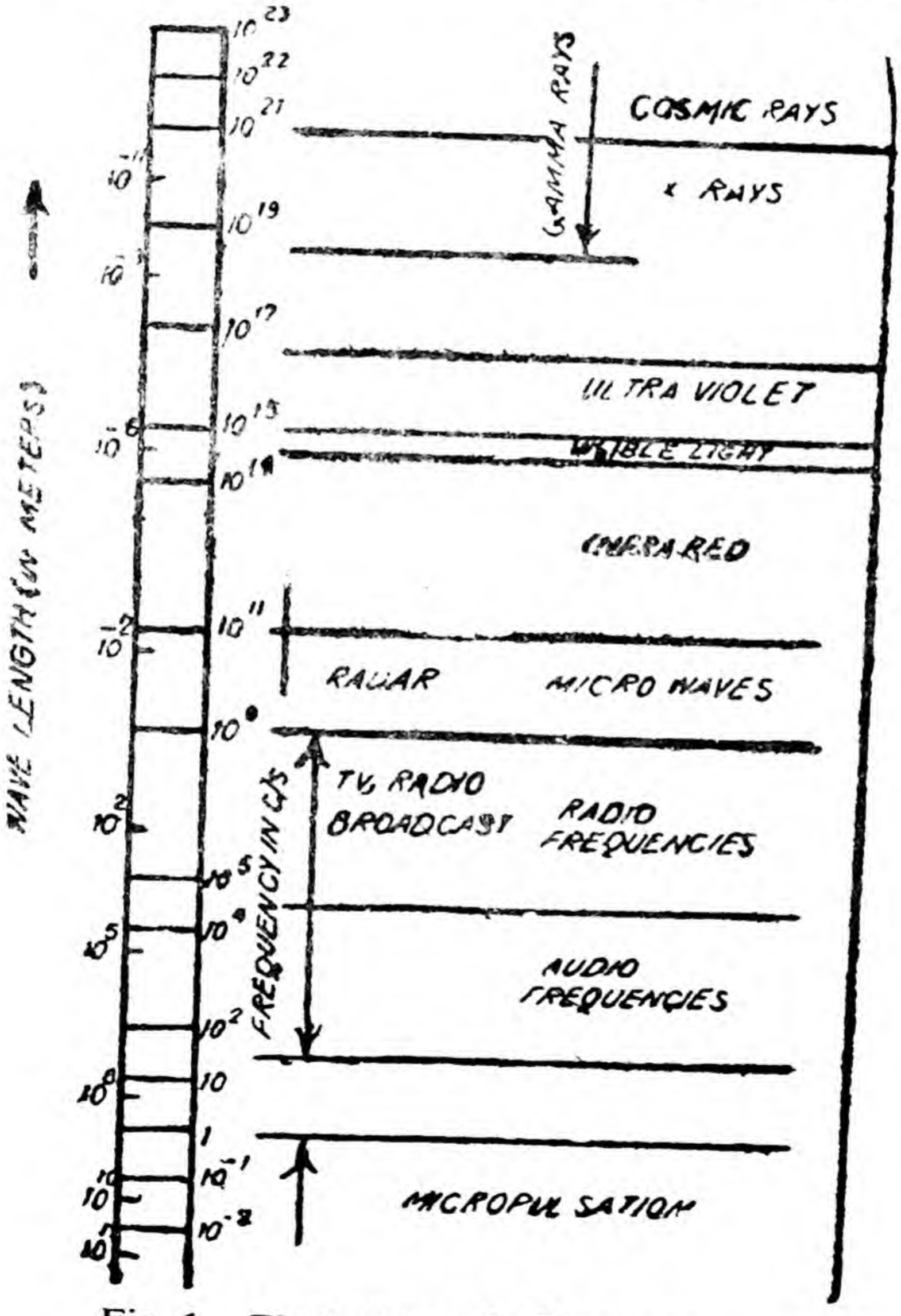


Fig. 1. Electromagnetic Spectrum.

emission occurs below the infra red portion of the electromagnetic spectrum, the term maser will be employed and when stimulated emission occurs in the infra red, visible and ultra violet portion of the spectrum, the term laser or optical maser is used.

34·1. CONCEPT THAT LED TO THE ORIGIN OF MASERS :

Dr. Charles H. Townes, in 1951, was asked by the United States Navy to help find ways to extend to still higher frequencies the range of microwaves that could be used in communication work. Dr. Townes then was professor of physics at Columbia University in New York City.

Dr. Townes was aware that the practical upper limit of frequencies that could be generated with the help of tubes and cavities was about 3.10^{11} c/s and it was hopeless to consider a further reduction in the size of cavities to fit the still more higher frequency (microwaves) generation and therefore he thought :

'Well, then, if man made structures would not do, why not try some of nature's smallest structures ? Molecules, for instance ?'

'Could a great number of gas molecules be energized and then in response to weak electromagnetic waves of suitable frequency' be stimulated to emit additional radiation at the same frequency ? This became the basis of maser operation and Townes led to the development of devices that used energy levels within atomic systems to produce wavelengths too short to be produced in any other manner.

34·2. BASIC PRINCIPLE OF MASER OPERATION :

In explaining the principle, we shall use the term coherent stimulated emission, population inversion and pumping. Therefore, it is essential first to write few lines about each of these terms.

(i) **Coherence :** The qualification of short wave transmitter demands that it should be

(*a*) powerful,
(*b*) pure and precise, and
(*c*) free from noise.

All these factors require that full radiated power should be concentrated in a narrow frequency band as narrow as 1 mc/s. This requirement can be fulfilled only when the radiated beam is coherent as explained below :

Though sun radiations are most intense but they cover a wide range of frequencies from deepdown in infra red, through light, to far into ultra violet. Therefore, power of radiation with a bandwidth of 1 mc/s is found to be quite low. Similarly, radiations from ordinary light sources *e.g.* tungsten filament, neon and fluorescent tube lights, cover wide range of frequencies, *i.e.*, their output is spread over a wide range of frequencies without supplying much power at any particular frequency. In other words, output of such light sources is irregular, mixed, composed of bits and pieces, differing in their frequencies, directions and durations or in short it is *incoherent*. The incoherent beam is thus noisy and can not be intensely concentrated into space whereas coherent beam is uniform in frequency, in amplitude and continuous. For long range

communication we require intense and noise free radiation, for the production of which coherenec is an essential condition because only coherent light :

(*a*) can be concentrated into a narrow pencil for it can be held parallel as it travels.

(*b*) avoids dilution of energy as it travels through space. This is because coherent radiation avoids great spreading.

We have thus come to the conclusion that coherent light beam is intense and can be used to shoot at the pin point in space. We shall find that process of stimulated emission leads to the generation of coherent radiations.

(ii) **Stimulated Emission :** Between radiation (photon) and matter (atoms) there are three possible interactions :

(*a*) *Spontaneous emission* : In this type of interaction, excited atoms emit photons that carry corresponding amount of energy into the surrounding electromagnetic field. The atoms then drop to the lower levels. Refer to fig. 2 (*a*).

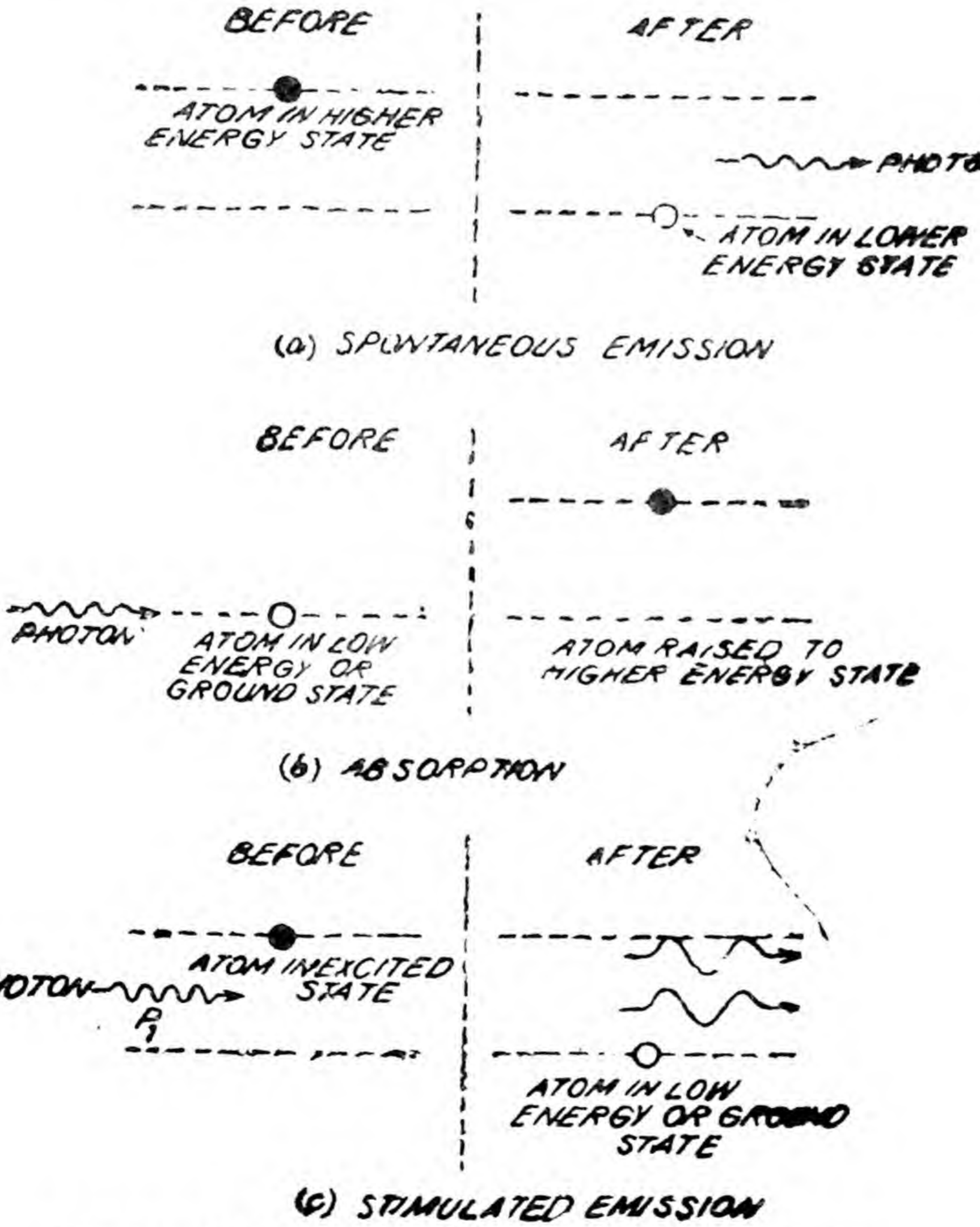


Fig. 2. Possible interactions in light and matter.

(*b*) *Absorption* : Photons of suitable size shower down on un-excited atoms. Photons are absorbed by the atoms. The energy of photon is converted into internal energy of the atom. The atom is then raised to an excited energy state. At some later instant, the excited atom may radiate this energy spontaneously emitting a photon and reverting to ground state or some intermediate state. Refer to fig. 2 (*b*).

(*c*) *Stimulated emission* : Photon of suitable energy size shower down on *excited* atom and each atom emits a photon of size identical to the stimulating (incident) photon. Thus it results in the appearance of one additional photon. Refer to fig. 2 (*c*).

In spontaneous emission, the atoms emit the photons separately and at random and we can not predict which atom among the excited atoms will emit next, *i.e.*, one can not be sure of the direction or timing of emission. Thus spontaneous emission is incoherent, whereas stimulated emission takes place only when a photon of the same size as that which the excited atom is ready to emit, is showered down on the latter. The timing of emission of additional photon is thus subject to the incidence of stimulating photon and therefore, exact. Both photons (stimulating and stimulated) fly off together in identical direction. Thus one can always be sure of the direction and timing of emission or in otherwords stimulated emission is coherent with stimulating emission.

(iii) **Population inversion and Pumping** : Most atoms have many possible energy levels. Let for the present discussion, we consider that each atom has four *possible* energy levels and the body of matter constituted by such atom is in a state of thermal equilibrium, *i.e.*, its temperature is steady. We take that the temperature of the body is quite low and predict the number of atoms in each of the four possible energy levels, called the population at

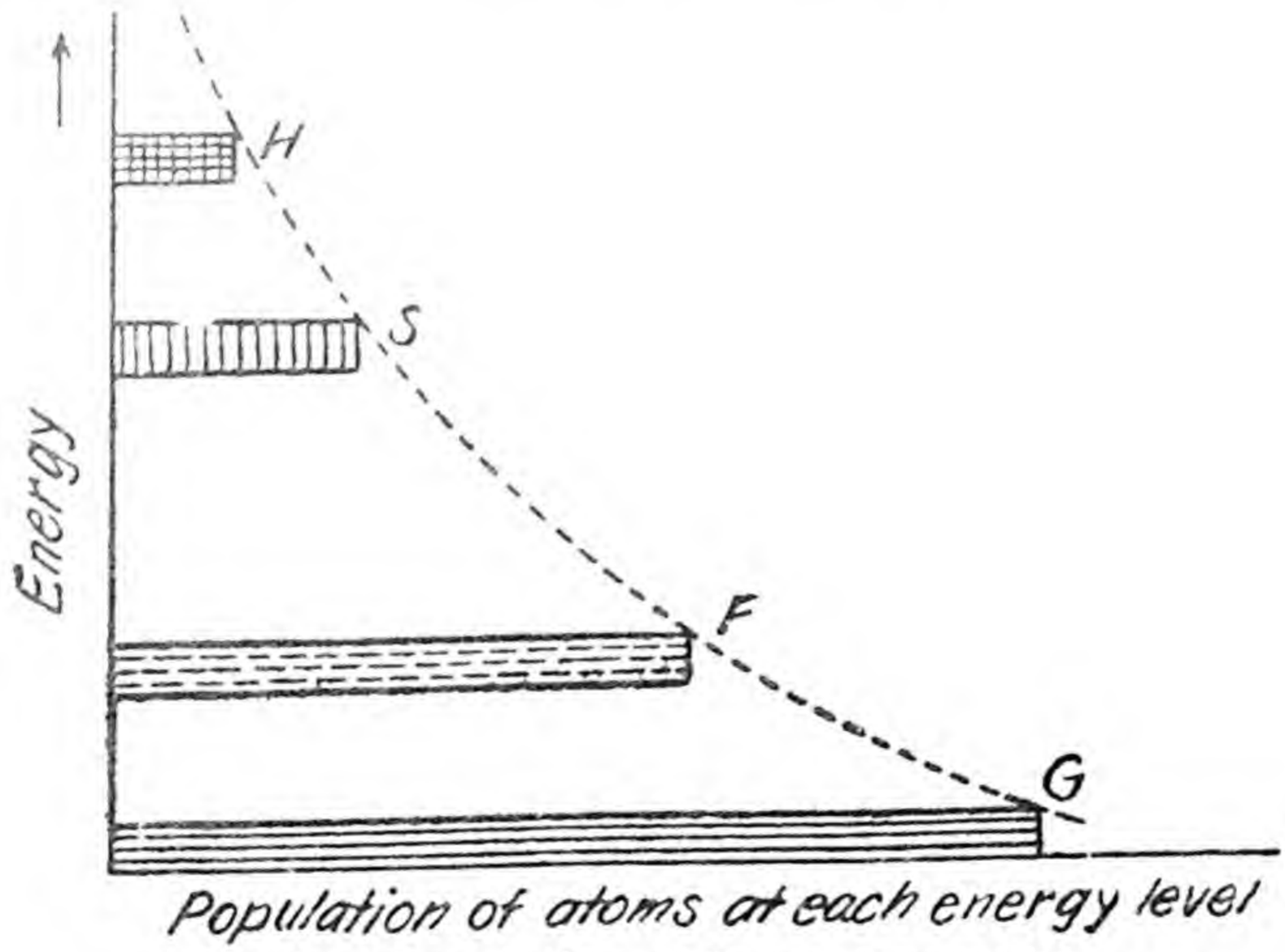


Fig. 3. Population at low temperature.

that level by fig. 3. The largest population is in the ground state—the lowest possible level.

Now photons of size that these atoms can absorb are made to incident upon this body. Temperature of body inreases but remains steady. At this new temperature, population chart is shown in fig. 4 (*a*). We note that relative differences in population are less though biggest population is still at the lowest energy level. As the temperature is further increased by showering over the body the photons of suitable energy size., population differences are further decreased as shown in fig. 4 (*b*).

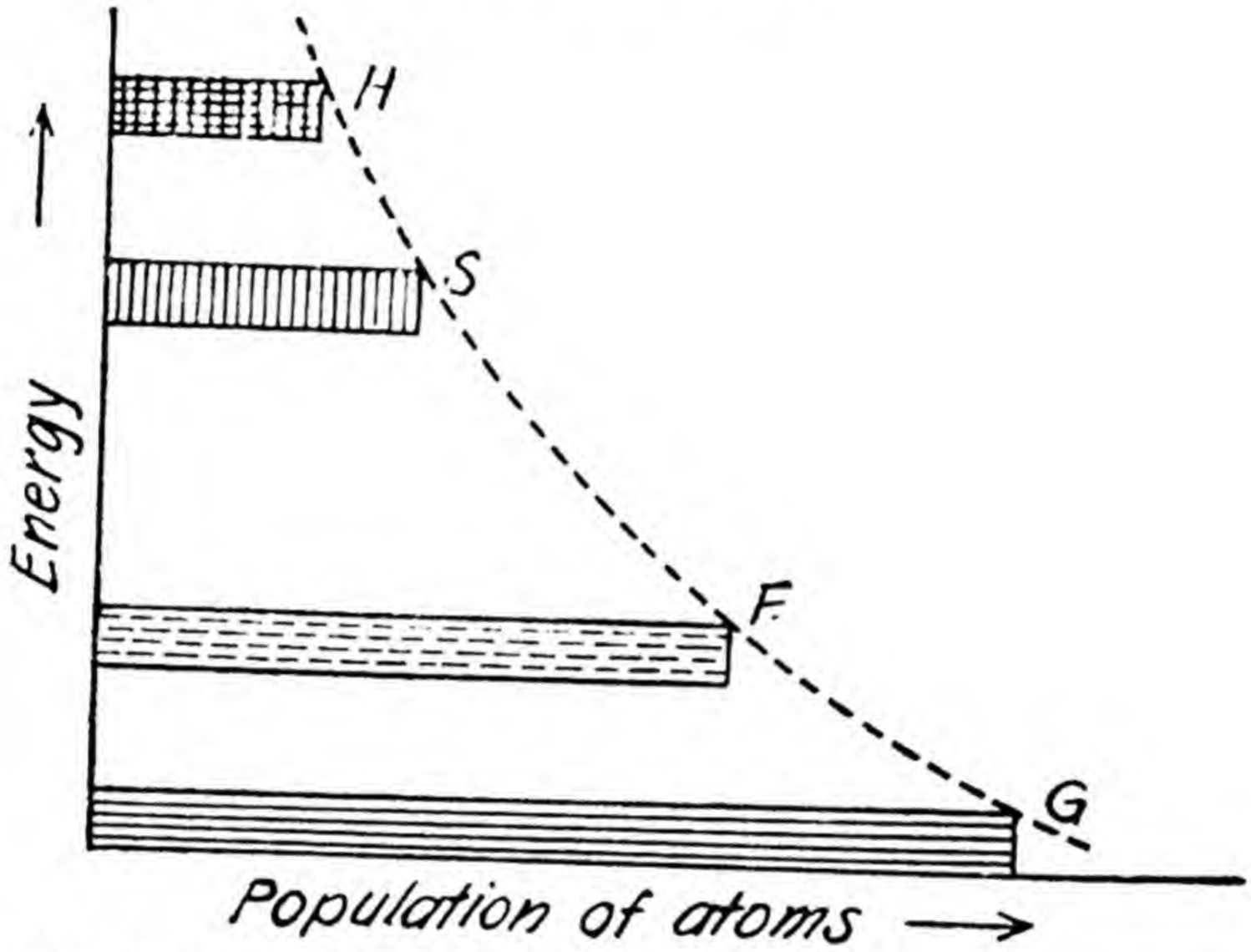


Fig. 4. (*a*) Population at different levels in a body at a higher temperature.

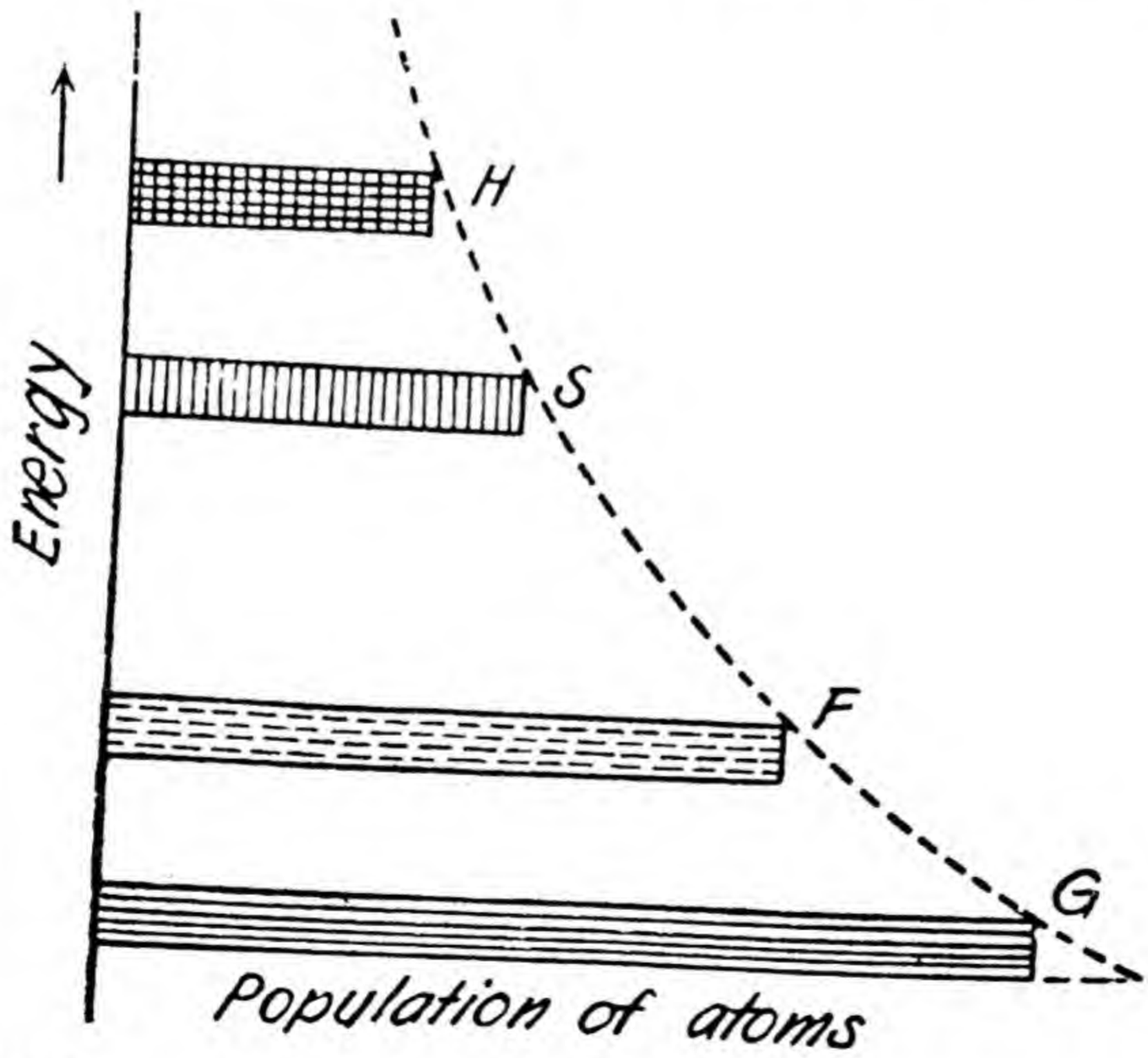


Fig. 4. (*b*) Population at different levels in a body at a still higher temperature.

When by some means it is arranged that population at higher energy levels is more than that in the lower energy levels—process

called *population inversion or reversal,* then a chain reaction of emission will start. Some atoms in *H* level will start the chain to *G* level, emitting a photon. Each photon will start the chain reaction of stimulated emission and thus most of the photons (born as a result of simulated emission) emitted are completely coherent and uniform with one another. It is thus obvious that population reversal is essential to produce the stimulated emission.

The method by which such a population inversion is affected is called 'pumping'. In this process atoms are raised to an excited state by injecting into the system electromagnetic energy at a wavelength *different* from the stimulating wavelength. The process will be described under the separate head of solid state maser.

In short, maser principle can be stated as follows :

Population of atoms in higher energy level as compared to lowest energy level is increased by pumping the matter by photons of appropriate size. Thus atoms are excited to higher energy states. During the period in which the atom is excited, it can be stimulated to emit a photon if it is struck by an outside photon having precisely the energy of the photon that would have otherwise been emitted in spontaneous emission. Then stimulated photon (emitted wave) falls precisely in phase with the stimulating photon (incident wave). For stimulated emission to predominate over absorption, it is essential that excited atoms should be in excess in the active medium.

34·3 DESIGN OF A MASER :

The first problem encountered in the design of the maser is to prepare active medium in which most of the atoms can be placed in an excited state. Then an electromagnetic wave of suitable frequency is made to incident upon this active medium that results in the stimulated emission of cascade of photons. As told previously, excited atoms must outnumber the normal atoms for the stimulated emission to predominate over absorption.

To prepare the active medium, electromagnetic energy is pumped into the system. After preparing the active medium, it can be enclosed into a reflecting box or cavity resonator of appropriate dimensions.

When the medium is stimulated by the electromagnetic wave of right frequency, emission starts. The stimulated waves (photons) will all have the same frequency. When enclosed in a box, the reflections will occur at the two reflecting walls. These reflections will amount to additional stimulated emission—the so called amplification of stimulated emission. The box can be so designed as to allow the building up of waves with only one mode of oscillation (and hence of single frequency)—the mode of stimulated radiation. Thus with appropriate dimensions of the box, a steady wave is built up in the box.

34·4. THE FIRST MASER (Gas Maser) :

In 1954, C. H. Townes, with James P. Gordon and Herbert J. Zeiger, constructed the first maser. This is also called as the Ammonia Gas Maser. The scheme is shown in fig. 5.

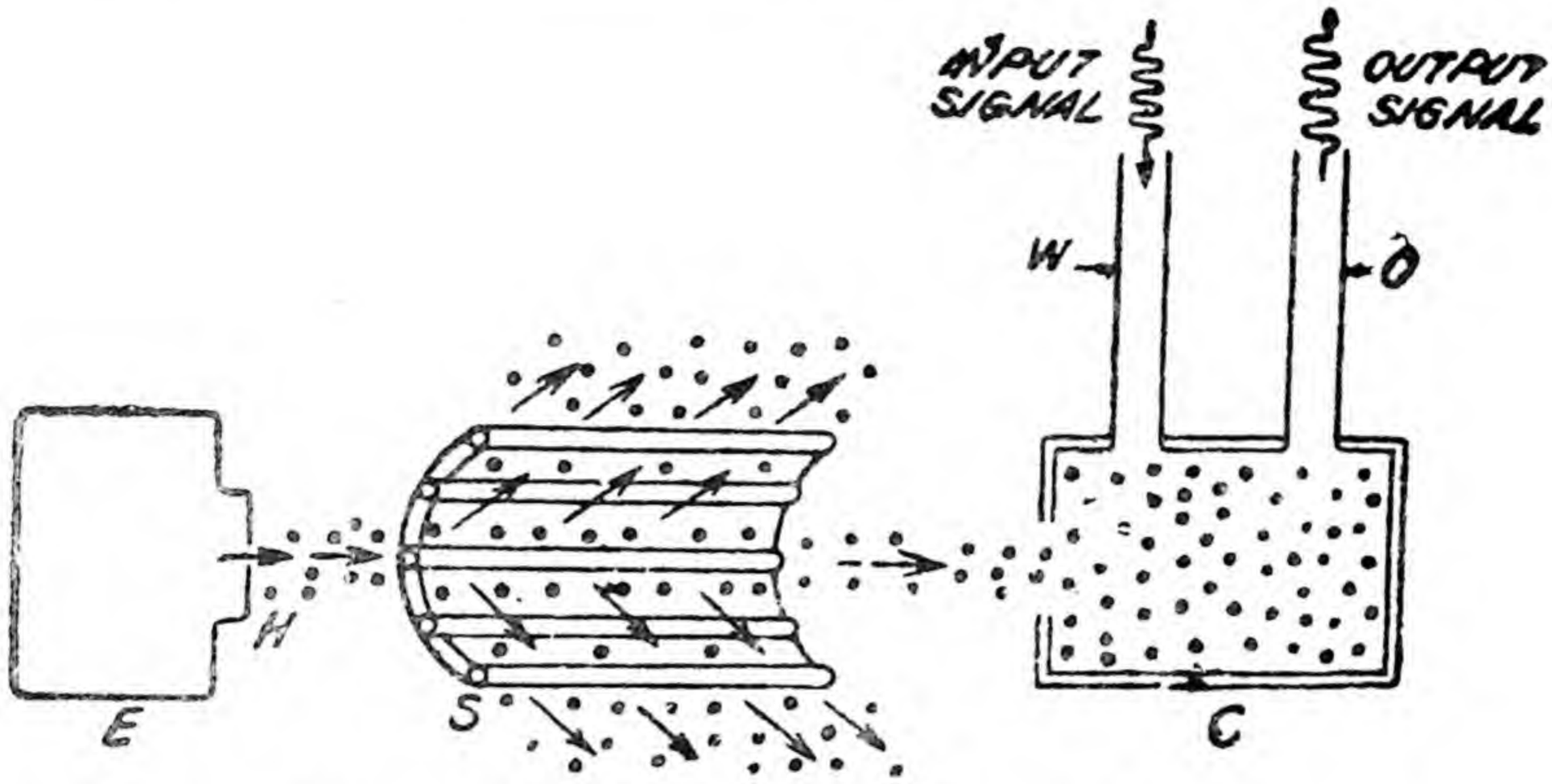


Fig. 5. Essentials of a gas beam maser.

Townes based the use of ammonia on the findings that an ammonia molecule, consisting three hydrogen atoms and one nitrogen atom, has two energy states (levels) with an energy difference equal to the energy of a photon whose frequency is 24,000 mc/s. Further, excited ammonia molecules are repelled by an electric field while unexcited molecules are attracted by the electric field.

Operation : Ammonia molecules are excited in the oven *E*. They stream out of the outlet *H*. This stream of ammonia molecules is then passed through the separator *S* which consists of charged rods. The electric field due to the charged rods attracts the unexcited molecules while excited molecules are repelled. But since the separator is cylindrical in the shape, repulsion from all sides will be balanced so that excited molecules continue to move along the axis of separator. Thus unexcited molecules are removed from the molecular beam and only excited molecules reach the resonant cavity *C*. Then these excited molecules are subjected to stimulated emission by a microwave beam (a beam of photon) of frequency 24,000 mc/s that enters the cavity through the waveguide *W*.

Output : Stimulated photons (photons given out by the excited molecules due to the incidence of microwave beam or stimulated wave) suffer reflection from the wall of the cavity, colliding with molecules and producing additional photons. The cavity is designed so as to build up the stimulated photons of frequency 24,000 mc/s, to usable strength. The frequency deviate only 3 to 5 kc/s from its centre frequency of 20,000 mc/s, *i.e.* bandwidth is narrow.

34·5. SOLID STATE MASER :

In gas, the molecules are not so closely packed as they are in a solid. Due to this reason the power output of the gas beam masers, described earlier, remain low. There are two types of solid state masers as discussed below :

(a) Inversion Maser : In most atoms, the electrons pair off so that the magnetic field from one cancels the magnetic field of other and no magnetic field is associated with these atoms. But in the atoms in which unpaired electron exists, a resultant magnetic field is always associated with such atoms. These materials are called paramagnetic substances and unpaired electron can exist at a number of energy levels. Example is phosphorous, chromium. Thus each atom of chromium has unpaired electron and therefore, it behaves like an ion rather than like a complete normal atom.

We consider a ruby crystal as active material. It is a synthetic crystal in which a small percentage of aluminium atoms is replaced by chromium atoms, —the latter being called as impurity or guest atoms. Percentage of chromium atoms is kept low particularly because they work better when separated from each other by a large number of aluminium atoms. The difference between the energy levels of the active materials (crystal) can be varied by the application of magnetic field of varying strength. This means chromium atoms can be tuned magnetically to certain transition frequencies—stronger the magnetic field, higher the transition frequencies.

Operation : We proceed in the following steps :

(*i*) First the ruby crystal is subjected to a magnetic field that will fit a suitable energy gap between the lower energy state L and higher energy state H of the crystal. More chromium atoms will be found in L state.

(*ii*) Crystal is then cooled by placing it in liquid helium to remove the energy of random atomic motion. Due to this cooling, most of chromium atoms fall to lower energy state L and the difference in the population of the chromium atoms in the two energy states will be sufficient.

(*iii*) The crystal is then subjected to high power microwave signal. The frequency of this microwave signal is varied and just as the frequency reaches and passes the frequency which matches the energy difference between the chromium atoms in L state and those in H state, population reversal takes place *i.e.*, most of the chromimum atoms are raised to high energy state, H.

(*iv*) Due to this population inversion, chromium atoms that are excited fall to lower energy state L giving off radiations (photons). Obviously, stimulated emission start at the input microwave signal frequency which matches the energy difference bet-

ween chromium atoms in *L* and *H* states. Thus emission starts at the desired transition frequency to which the energy gap between *L* and *H* states was adjusted by the application of magnetic field of appropriate strength.

(b) Multilevel Maser : George Feher, **H. E. D. Scovil** and H, Seidel of Bell Telephone Laboratories operated a three level maser.

We choose a crystal containing as an active material a group of atoms which have three energy level—marked as *L*, the ground state ; *M* an in between state ; and *H* the highest of three.

Operation :

(*i*) Crystal is cooled so that largest possible difference in the size of population of three levels occurs as shown in fig. 6.

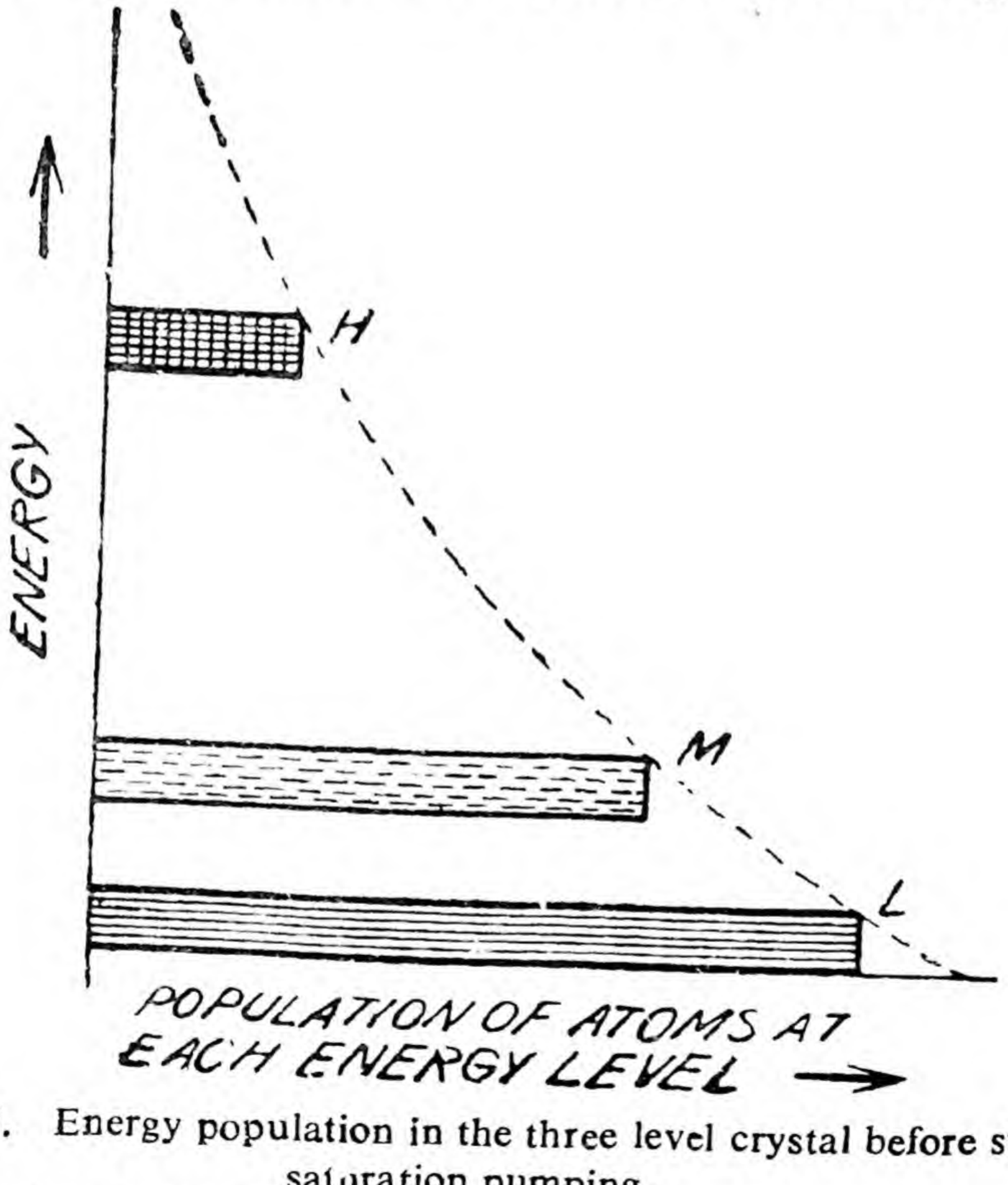


Fig. 6. Energy population in the three level crystal before start of saturation pumping.

(*ii*) The crystal is then subjected to a microwave signal of such a frequency as to raise atoms from *L* to *H* level, while population at *M* level remains unaffected. Energy population is now shown by fig. 7.

Population reversal occurs and *H* as well as *M* levels have higher energy population than *L* level.

(*iii*) Now the stimulating wave (radiations or photons) of frequency equivalent to energy difference between *M* and *L* levels is

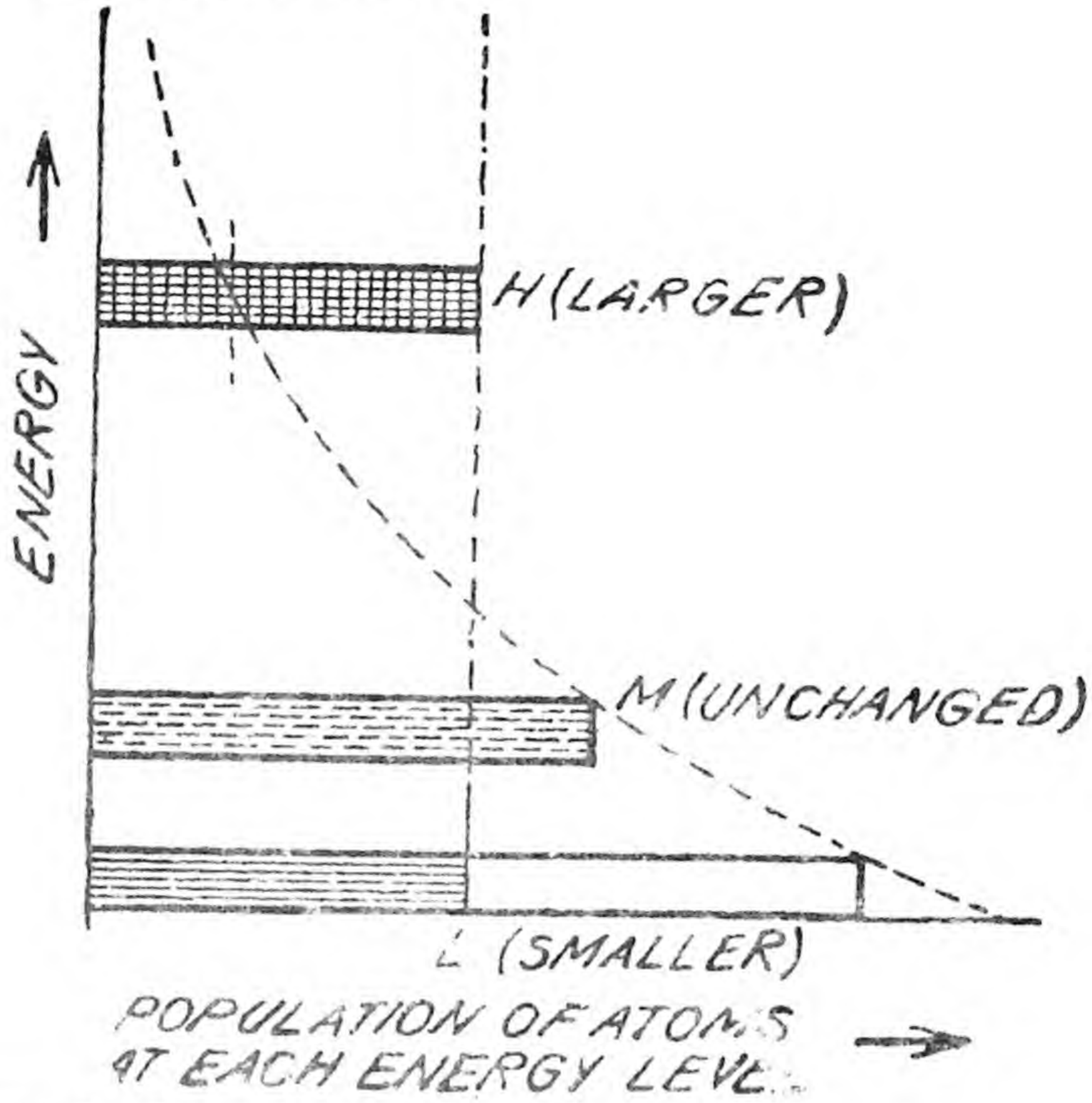


Fig. 7. Population reversal after saturation pumping : population of level, *M*, is unaffected.

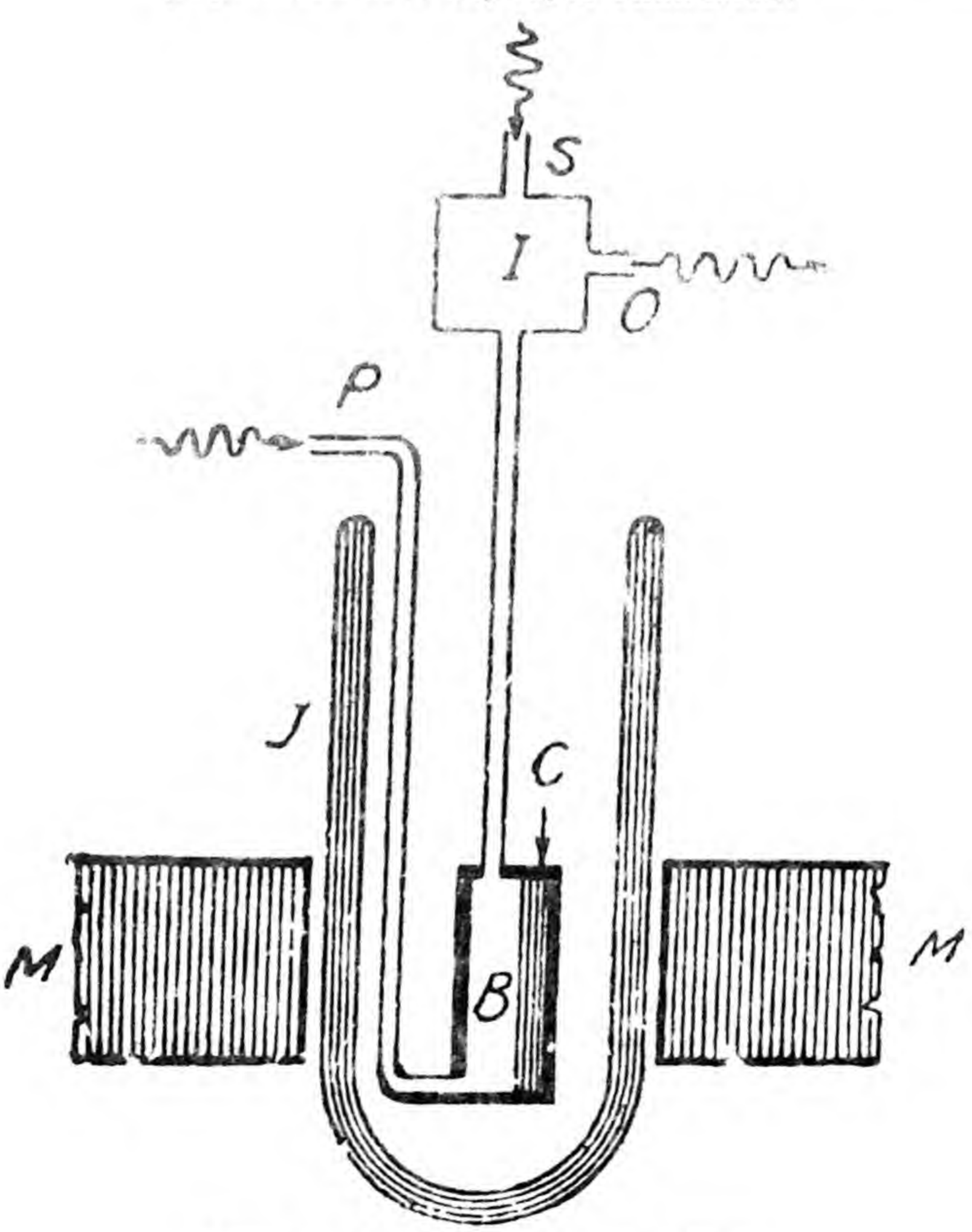


Fig. 8. Multilevel Maser parts.

made to fall over the crystal. The maser action will start atonce. Excited atoms from *M* level will drop to *L* level. Thus maser action will emit the frequency which corresponds to the transition from *M* to *L*. This is a smaller energy jump, hence corresponds to a lower frequency, than that from *L* to *H* state. The pump frequency (frequency of the wave used to excite the atoms) which corresponds to the energy gap between *L* and *H* states is obviously higher than emission or maser frequency (from *M* to *L* state).

Arrangement : Multilevel masers have many parts as shown in fig. 8 *C* is the crystal; *B* the resonant cavity ; *P* the waveguide through which pumping energy enters; waveguide pipe, *S*, through which the stimulating radiation or signal to be amplified (by the addition of the photons of the same frequency as that of the signal emitted by excited atoms through the process of stimulated emission) is applied to the crystal : *I* is the isolator : *O* is the out-let : *J* is cold flask containing liquid helium and *M* is the magnet. Magnetic field is applied to adjust the energy gaps between the different levels to the desired value—so as to amplify the signal of desired frequency.

34·6. EXTENSION OF MASER PRINCIPLE TO OPTICAL RANGE ; LASER :

In 1958, A.E. Schawlow and C.H. Townes proposed a method of constructing a maser for optical wavelengths by using a resonant cavity whose dimensions were millions of times the wavelength of light. Fig. 9 shows the diagram of their method. Active material is shaped into a cylinder with parallel transparent sides as reflectors at both ends.

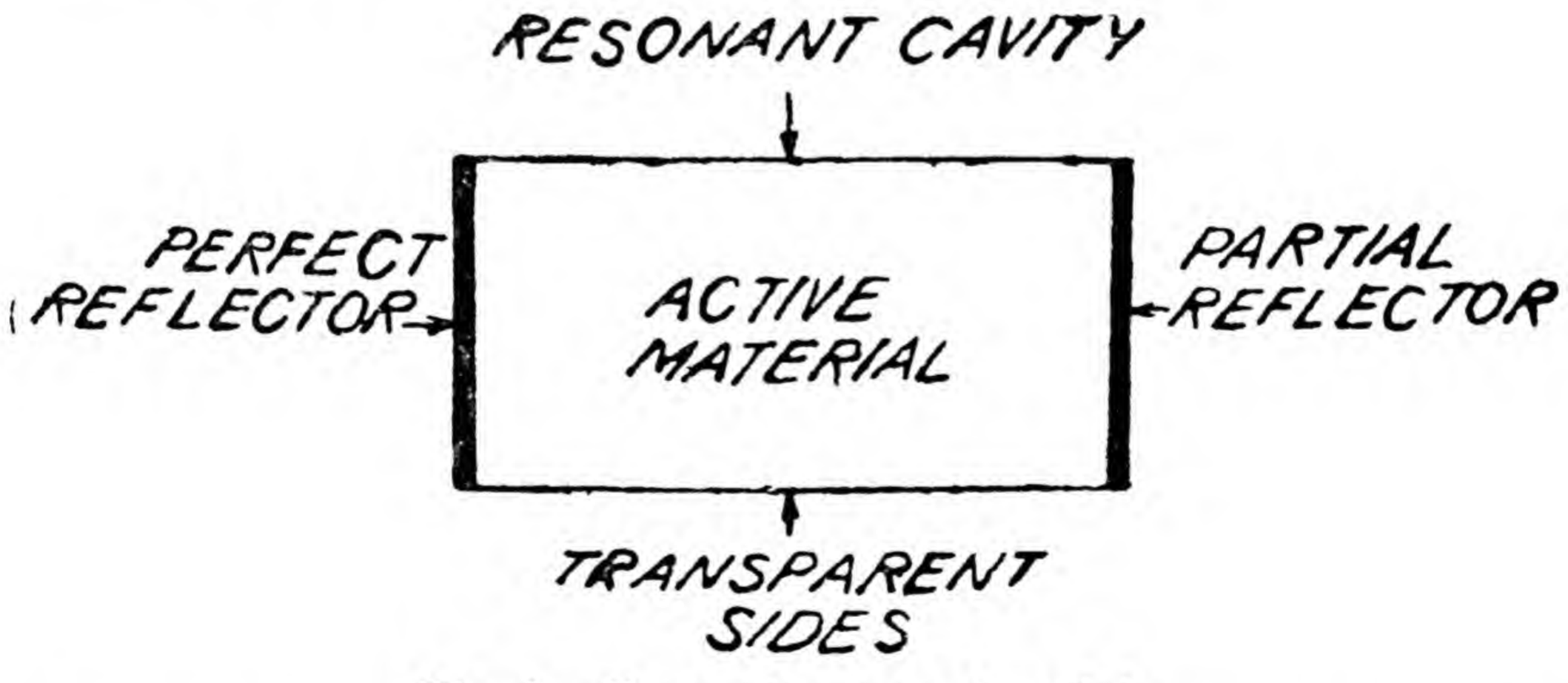


Fig. 9. Townes proposed model.

Active material is energized by the pulses of light from flash tube. When emission begins, photons that do not travel parallel to the main axis of the cylinder will leave the cylinder through the transparent sides either immediately or after one or two off angle reflections. Thus such photons will not collide with other excited atoms of the active material.

Photons moving parallel to the axis of the tube will suffer several reflections between the two end reflectors. During the

reflections, in their way, they interact with other excited atoms of the active material and thus produce more photons. Due to the chain reaction of stimulated emission, intensity of photon beam parallel to the axis of tube grows on. If one of the reflectors is partially transparent (say 10% transmitting) then some of the coherent light (stimulated photon beam which is parallel to the axis of the tube) would emerge through this reflector and serves as output laser beam.

34·7. RUBY LASER : (Solid State Laser) :

In 1960 T.H. Maiman constructed the first laser using a ruby

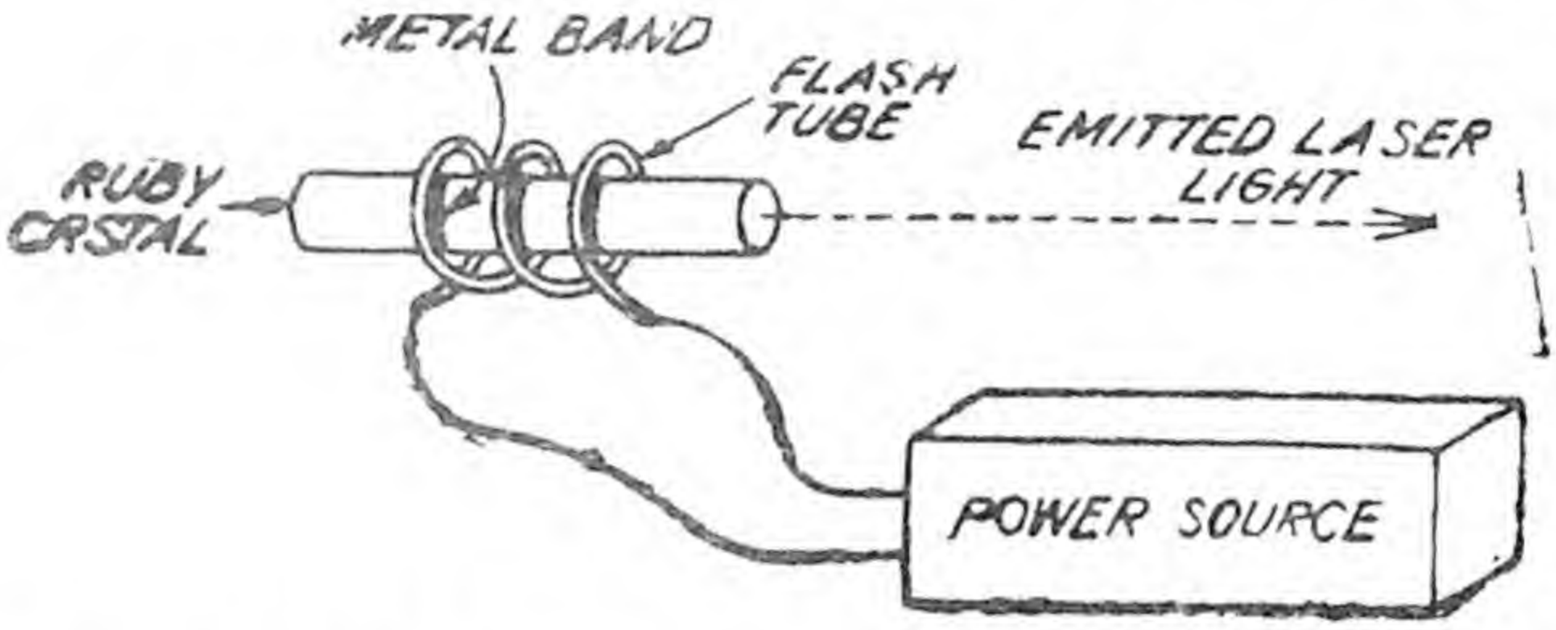


Fig. 10. Essentials of a crystal laser pumped by pulses of light from a photon discharge tube.

crystal. It is capable of generating very high power for very short intervals of time. Arrangement is shown in figure 10 and explained as follows :

Operation :

(*i*) Active material is a small cylinder of pink synthetic ruby, about 0·5 cm. in diameter and few centimeters long. Ruby contained 0·05% of chromimum, imparting a pink look to the crystal.

(*ii*) It is pumped by powerful bursts of light from a flash tube.

(*iii*) Chromium atoms are excited and raised to *H* energy state. Population is reversed. From higher energy state they do not return to ground state in a single jump but instead they require two steps to return to their ground state. First return is from *H* state to metastable state *M* which is a shorter jump and energy emitted in this transmission is passed to the crystal lattice as heat. This energy is thus not radiated in the form of photons. Refer to fig. 11.

(*iv*) The chromium atoms returned to *M* level can remain in this state for several milliseconds—the so called metastable state.

(*v*) The accumulation of coming excited atoms at the *M* level increases the population at *M* level and then transmission occurs from *M* to *L* level emitting out the photons. The emission is significant. Stimulated emission starts. Photons travelling parallel to the axis of the tube (crystal) will start a cascade of photon

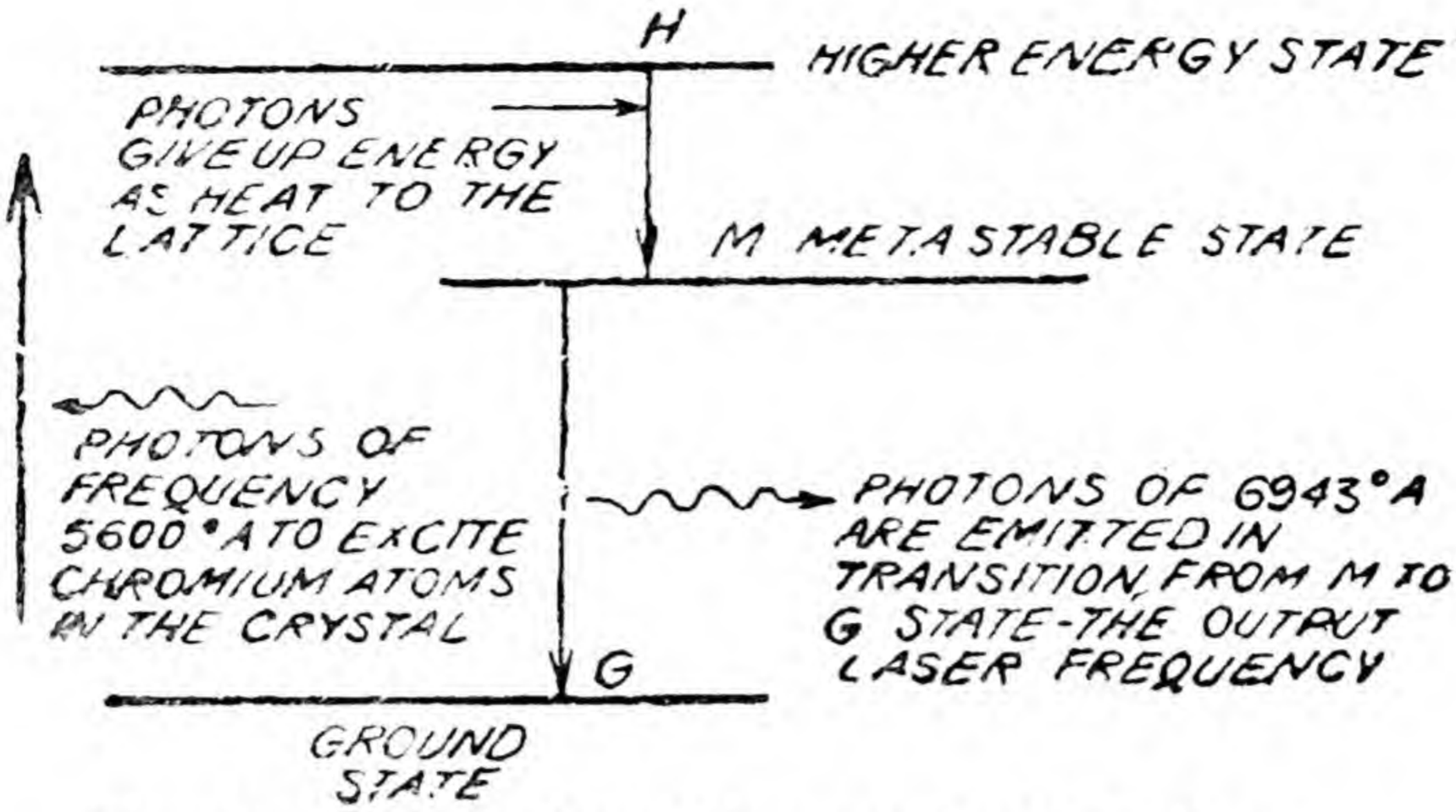


Fig. 11. Energy level diagram for Ruby laser.

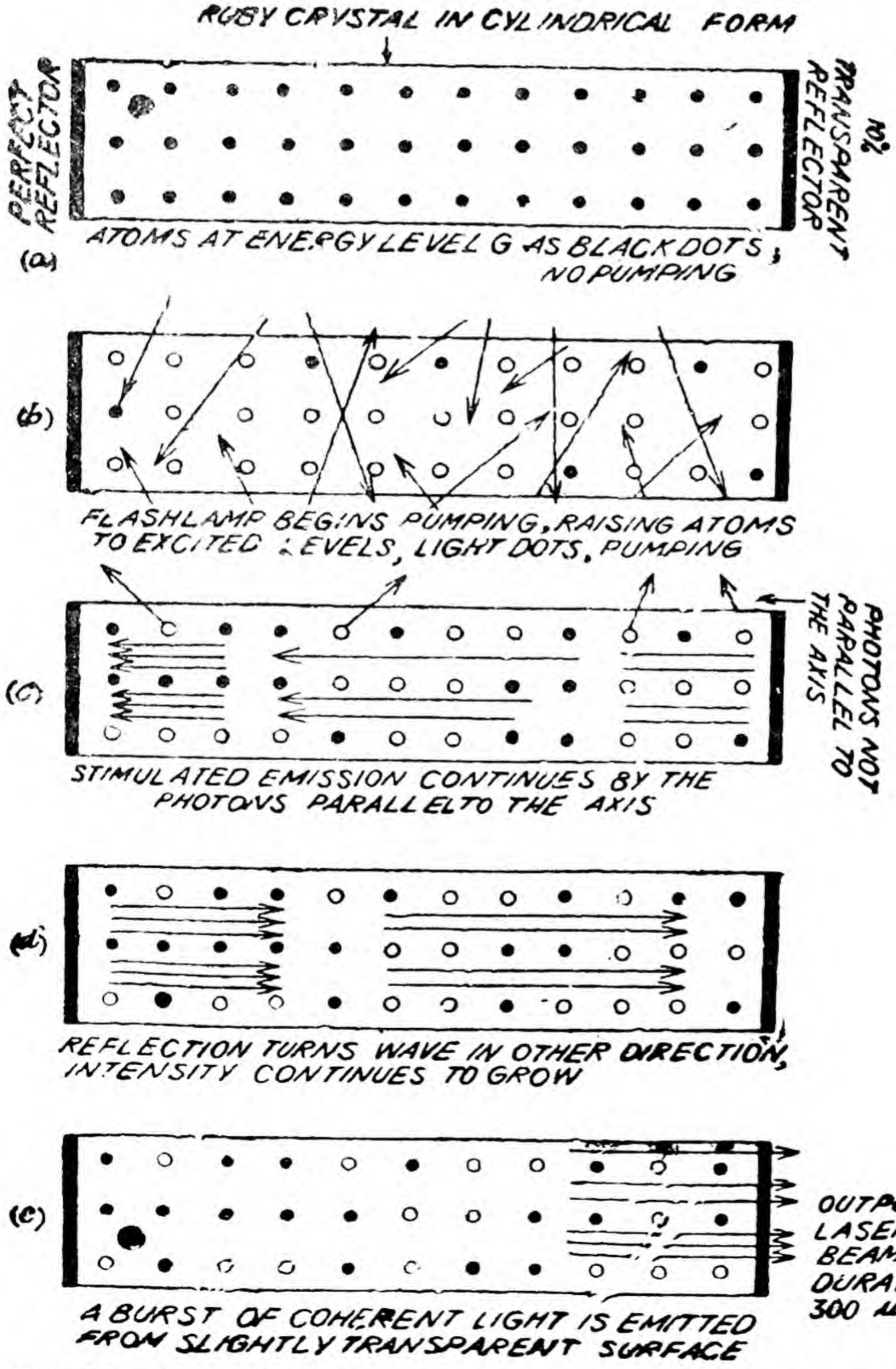


Fig. 12. Laser beam production by stimulated chain reaction.

emission while the photons travelling in any direction other than this will pass out of ruby. The process is explained in fig. 12.

(*vi*) Photon beam parallel to the axis of the crystal grows in strength and some of it bursts through the partial reflector and serves as output laser beam.

(*vii*) The frequency of stimulating wave corresponds to the energy difference between *G* and *H* states (5600 *A*°) whereas stimulated wave (output photon beam) corresponds to the energy difference between *M* and *G* states (6943 *A*°).

Output : The output beam—a red laser flash- have principal wavelength of 6943 *A*°, equal to about $4{\cdot}32 \times 10^{14}$ c/s It is well within the visible spectrum. The duration of output flash is about 300 microseconds, but it is very intense - of about 10,000 watts. This high power intensity arises due to the coherence of output beam, since all photons travelling parallel to the axis have space, time and directional coherence.

Shifting the laser frequency ; Raman Laser : E J. Woodburg discovered that by passing a beam of ruby laser light through a vessel containing benzene, additional wavelengths on either side of 6943 *A*° wavelength of ruby laser were produced.

34·8. GAS LASER :

The main drawback of Ruby laser is that the output beam is not continuous though very intense. For laboratory purposes, we require continuous laser beam though may be less intense. In 1961, Ali Javan, W.R. Bennett Jr. and D.R. Herriott reported a continuous laser.

The arrangement is shown in fig. 13.

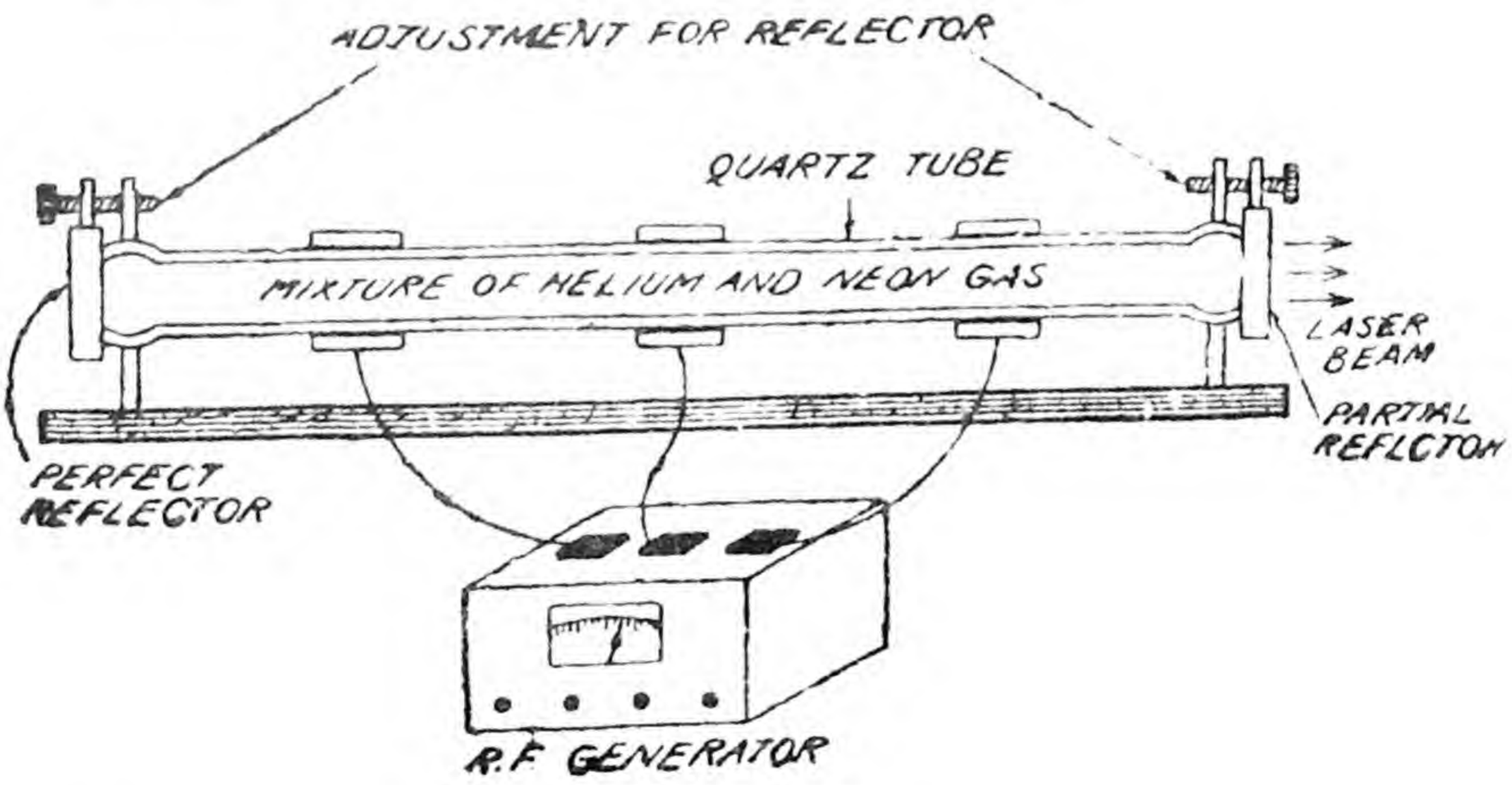


Fig. 13. Gas Laser.

Operation :

(*i*) A long narrow quartz tube contains a gas mixture, composed of a majority of helium and a minority of neon atoms. At

one end of the tube is a perfect reflector, while at the other end is a partial reflector.

(*ii*) Mixture is pumped by injecting electromagnetic energy into the tube through the metal bands. The pumped wave has a frequency of about 4.10^7 c/s.

(*iii*) After excitation, stimulated emission starts as explained in art 34·7 and we get output laser beam of about 40 watts. The frequency of output beam in gas lasers commercially solid is about $2{\cdot}6.10^{14}$ c/s. (in infra red region) or $4{\cdot}7.10^{14}$ c/s. (in visible red region) We find that in this laser, in contrast to all lasers or masers described earlier used pumped frequency is *lower* than the output beam frequency.

34 9. APPLICATIONS OF LASER :

(1) They have got wide medical applications. They have been used successfully in the treatment of detached retinas. In treatment of both human and animal cancers a preliminary success has also been obtained.

(2) They have got wide industrial and chemical applications. They can blast holes in diamonds and hard steels. They can initiate or hasten certain chemical reactions which would proceed sluggishly or not at all in the absence of suitable photons.

(3) Due to the narrow bandwidth, high intensity and high frequency, the laser beams have become a means of communication between earth and moon or other satellities.

(4) Their study is also oriented for the millitary purposes. It is predicted that they would become the legendry 'death ray' that could annihilate anything, near or far. Thus laser is being developed to serve as a war weapon.

(5) The use of lasers in computers is under solid investigation particularly by *IBM* corporation They are trying to transmit an entire memory bank from one computer to another by the use of laser beam.

EXERCISES

1. Explain the process of stimulated emission. Draw a neat diagram to represent the component of solid state maser. Explain the operation.
2. What are different uses to which laser beams are put ? Explain the operation of a gas laser with the essential components. How stimulated emission takes place with the exchange of energy between Helium and Neon atoms.

EXPERIMENTS

Expt. 1. To design a.c. and d.c. V.T.V.M.

Circuit for design is given in the figure 1.

Procedure :

(i) **For the measurement of D.C. voltage :** Connect d.c. voltage to *AA* terminals. Then with d.c. voltage off, select a range (say *5V*) with the help of switch S_1 and adjust meter scale (μA reading) at zero by varying P_1. Now apply full voltage corresponding to a given range and adjust meter scale reading (μA scale) to full scale with the help of P_2. Then give *1V*, *2V*, *3V* etc upto *5V* (with the help of voltmeter attached with d.c. voltage device) and note μA reading for each applied voltage value. Plot a graph in voltage and μA reading.

Similarly for other ranges procedure can be repeated.

(iii) **For the measurement of A.C. voltage :** Connect fig. 2 apparatus at *AA* terminals of fig. 1 and proceed as explained in d.c. voltage measurement.

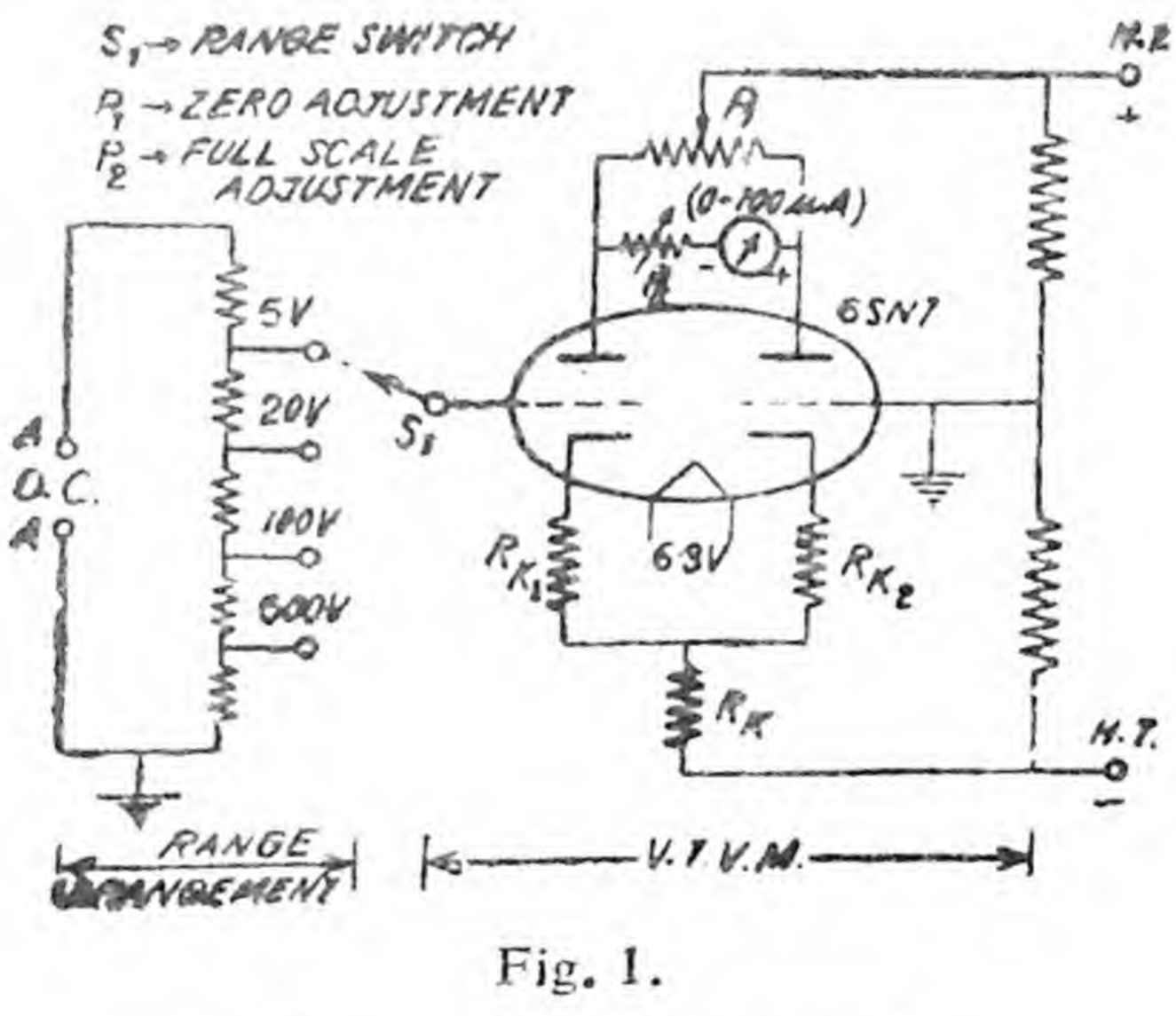


Fig. 1.

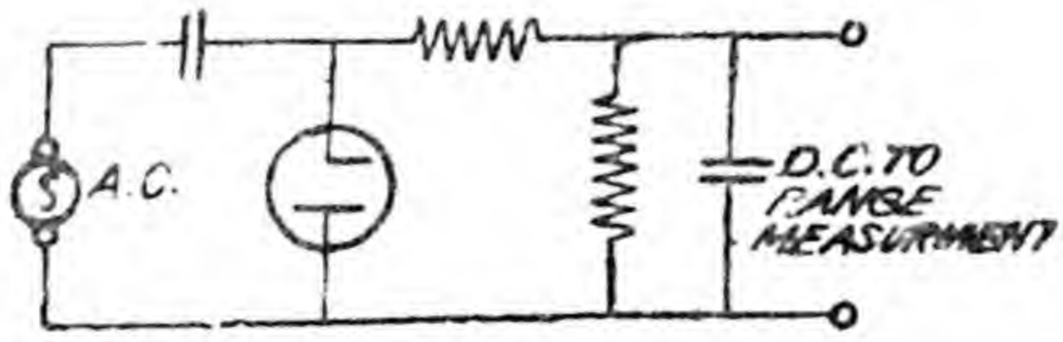
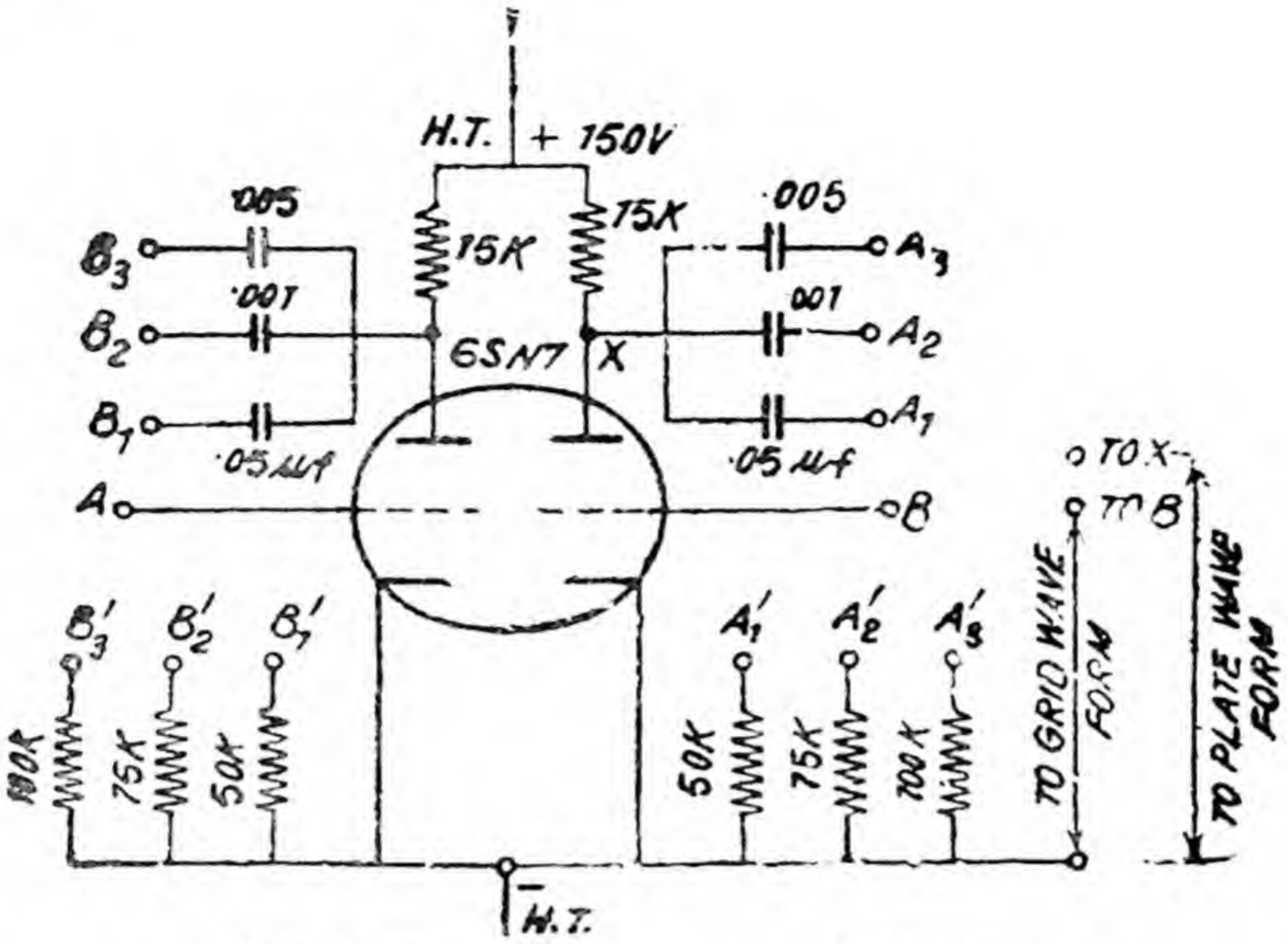


Fig. 2. Arrangement to be attached at *AA* terminals of range measurement if A.C. is to be measured.

Expt. 2. To design a multivibrator and depict the grid and plate waveforms.



R values	*C* values
50*K*	·05μf
75*K*	·001μf
100*K*	·005μf

Connect

1st reading AA_1A_1' & BB_1B_1'
2nd reading AA_1A_2' & BB_1B_2'
3rd reading AA_1A_3' & BB_1B_3'
4th reading AA_2A_1' & BB_2B_1
.. ..
9th reading AA_3A_3' & BB_3B_3'

(Circuit diagram for design)

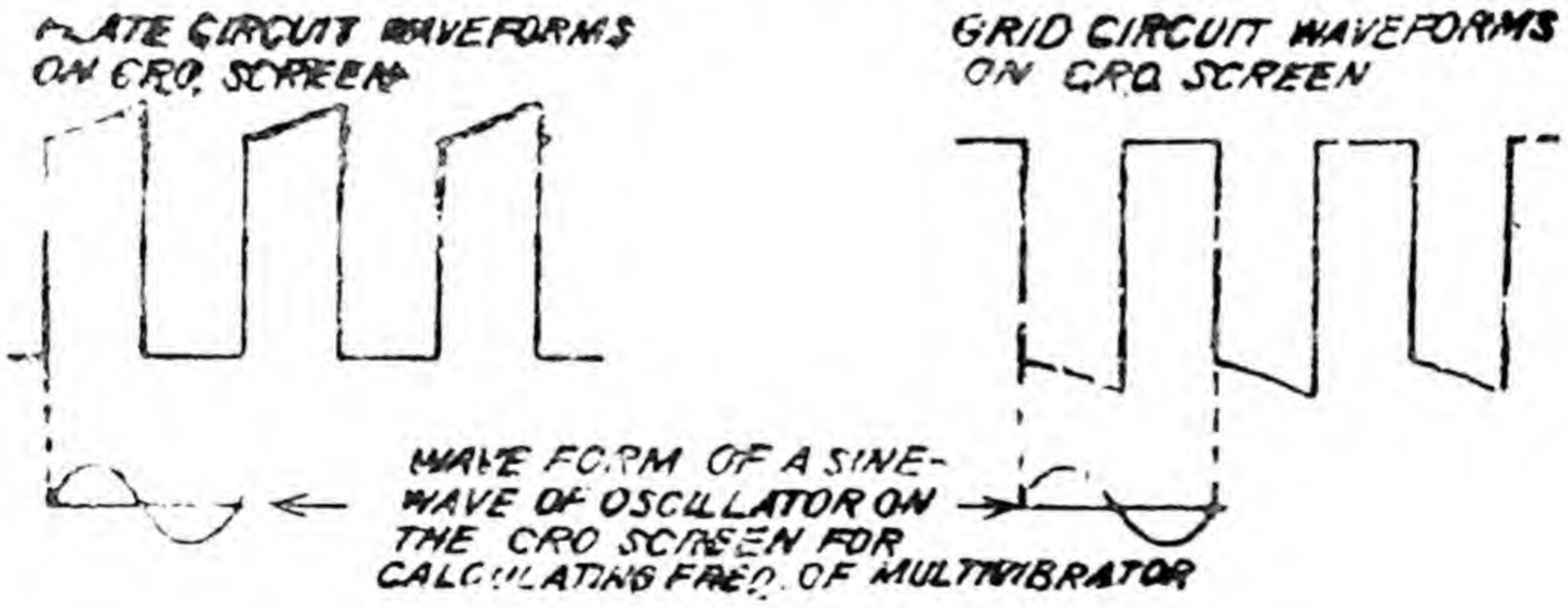


Fig. 3.

Procedure : Connect output terminals of multivibrator to *XY* plates of *CRO*. Obtain steady pattern of plate or grid waveform on the *CRO* screen with the help of frequency controls. Sketch the pattern on the paper. Repeat it with every reading.

Then inorder to calculate multivibrator frequency with every reading, connect a.f. oscillator to *XX* terminals of *CRO*. Vary oscillator frequency tc same number of wavelengths as that of multivibrator (without disturbing frequency controls of *CRO* which were already set for steady square wave pattern). Note the frequency of oscillator (that will also be the frequency of multivibrator). Say it is f, then

$$f = \frac{1}{2NRC},$$

where N is the factor which for most of the tube varies from 1 to 4. Calculate N for every reading.

R	C in μf	*frequency*	N
50K, 75K, 100K	·05		
		...	...
50K, 75K, 100K		...	...
50K, 75K, 100K	·005	...	...

Expt. 3. To design a wavemeter for certain range of frequencies (say 1 to 5 mc/s).

In circuit shown, L_2 is a calibrated variable inductor and C is calibrated variable capacitor (say 0 – 500 p.f.). Suppose minimum value at one end of C is C_1 (say 25 p.f.) and maximum value at other end is C_2. Then

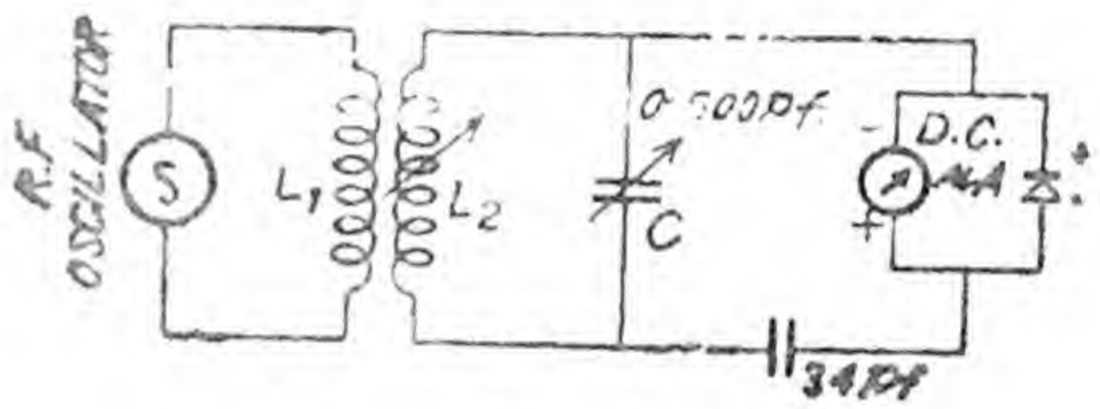


Fig. 4.

$$f_1 \text{ (say 1 mc/s)} = \frac{1}{2\pi}\sqrt{\left(\frac{1}{LC_2}\right)}$$

$$f_2 \text{ (say 5 mc/s)} = \frac{1}{2\pi}\sqrt{\left(\frac{1}{LC_1}\right)}$$

so that

$$L = \frac{1}{4\pi^2}\left(\frac{1}{f_2^2 - f_1^2}\right)\left(\frac{C_2 - C_1}{C_2 C_1}\right) \text{ henry}$$

in which f is in c/s and C is in farads.

Procedure : (*i*) Set L_2 on the above calculated value of inductance.

(*ii*) Set C at its minimum value, C_1 and vary oscillator frequency to obtain maximum indication (resonance) in μA. In fact this frequency should be f_2 but, due to stray inductance, stray capacitance and wire capacitances, it may be other than f_2, say, it is f_2'. Similarly, set C to its maximum value C_2 and vary oscillator frequency to obtain maximum indication in μA. This oscillator frequency should be f_1 but it may be different, say f_1'.

Thus wavemeter would work for frequencies lying between f_1' and f_2'.

(*iii*) Now to calibrate the designed wavemeter in frequency, give input at various oscillator frequencies and obtain maximum

indication in μA for every frequency. Write down these frequencies at the dial of variable capacitor.

Expt. 4. Measurement of L, C, R, and Q of a coil at high frequencies.

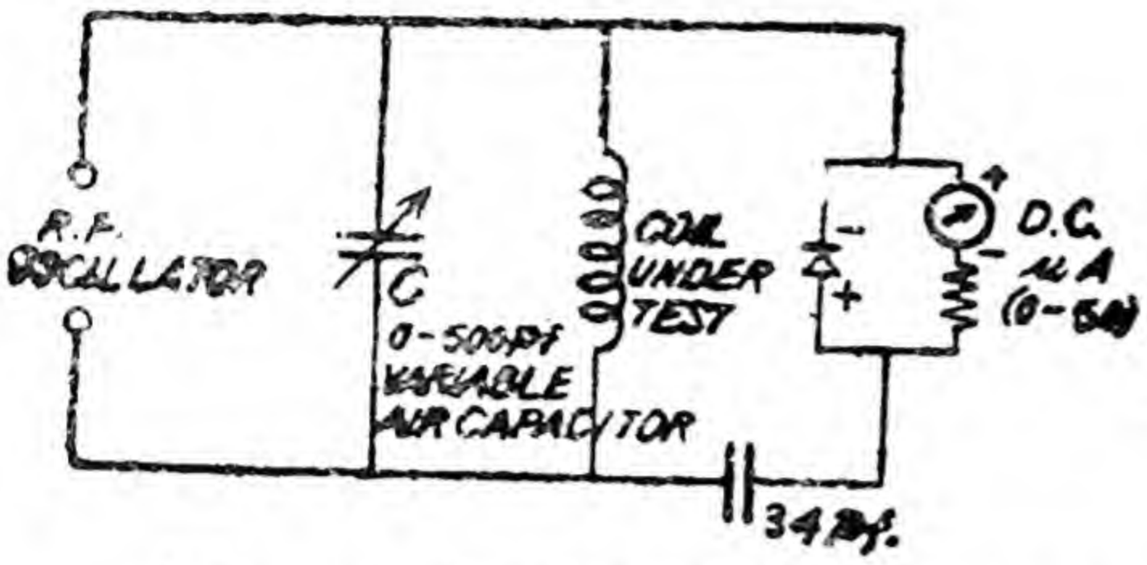


Fig. 5. Resonance measuring arrangement

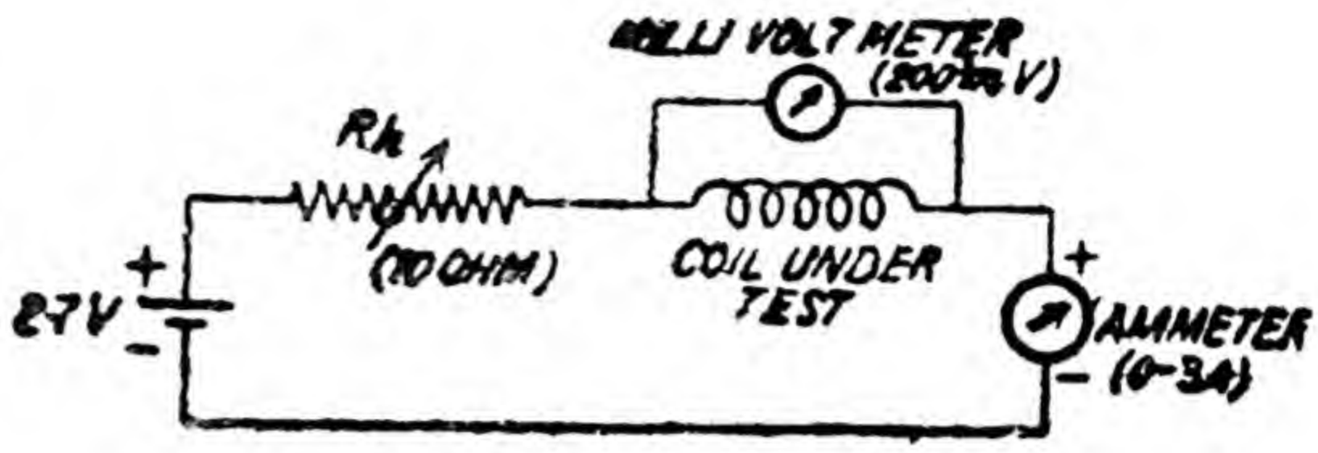


Fig. 6. Measurement of d.c. resistance.

Theory and Procedure :

(*i*) First of all, d.c. resistance of the coil may be calculated. For this measurement, circuit is shown in fig. 6. With the help of rheostat, obtain various reading of voltage and current. Then

$$\text{d.c. resistance} = \frac{\text{volts}}{\text{amperes}} = \ldots\ldots \text{ohms.}$$

(*ii*) Variable air capacitor is calibrated. Keep some value of capacitance in C. Let this be C_r, then vary oscillator frequency to get maximum indication in μA (resonance condition). Note this oscillator frequency. Let this be f_r. Now vary capacitance on either side of C_r value so that deflection in μA is reduced to 70·7% of its value at resonance (suppose maximum deflection at resonance is $20\mu A$ then vary capacitance above and belew C_r such that deflection is $\frac{70{\cdot}7}{100} \times 20 \simeq 14{\cdot}1$). Let these values of capacitance be C_1 and C_2. Repeat the whole procedure with various values of C_r and corresponding values of f_r, C_1 and C_2. For each set

$$Q = \frac{2C_r}{C_2 - C_1}, \qquad R = \frac{1}{4\pi f_r} \frac{C_2 - C_1}{C_1 C_2} \text{ ohms.}$$

$$L = \frac{\lambda_r^2}{4\pi^2 v^2 (C_r + C_0 + C_v)} \text{ Henry}$$

where v is velocity of light, C_0 and C_v include extra capacitances due to wire to other parts of the circuit. λ_r is $\frac{3 \cdot 10^8}{f_r}$ meters.

In order to calculate $(C_0 + C_v)$ plot a graph between λ_r^2 and C_r. Intercept on C_r axis will give $(C_0 + C_v)$. Draw graph between f_r and Q values. Formulae used are deduced below :

C_1 and C_2 are the capacities at half power points at which current falls to 70·7% of its maximum value.

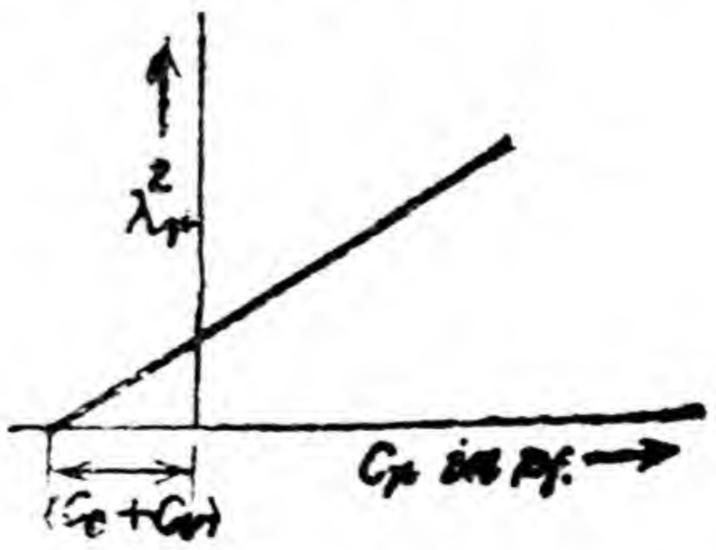


Fig. 7.

Thus
$$\frac{1}{\omega_r C_1} - \omega_r L = R$$

and
$$\frac{1}{\omega_r C_2} - \omega_r L = -R$$

so that
$$2R = \frac{C_2 - C_1}{C_1 C_2} \cdot \frac{1}{\omega_r}$$
$$R = \frac{1}{4\pi f} \frac{C_2 - C_1}{C_1 C_2}$$

Let $C_1 C_2 = C_r^2$, then
$$\frac{C_2 - C_1}{2C_r} = \omega_r C_r R = \frac{1}{Q}$$

or
$$Q = \frac{2C_r}{C_2 - C_1}$$

Further at resonance
$$\omega_r^2 L(C_r + C_0 + C_v) = 1$$
$$L = \frac{\lambda_r^2}{4\pi^2 v^2 (C_r + C_0 + C_v)}.$$

INDEX

Alignment of receivers, 1534
Amplifiers,
anode follower, 506
audio frequency power, 528
cathode follower, 498
class A, 425, 523, 534
class B, 426, 539
class C, 581
direct coupled, 608
distortion in, 432
feed-back, 485
grounded grid, 504
inductance capacitance coupled, 463
Miller effect, 428
push-pull 534
radio frequency amplifiers, 551
resistance capacitance coupled, 778, 447
speech, 621
transformer coupled, 465, 783
transistor, 725
waveform distortion, 438
Amplitude distortion, 433
Amplitude modulation, 1386
Analog computation, 977
Antenna,
action, 1206
arrays, 1248
broad-side array, 1250
continuous array, 1254
current distribution, 1227
current element, 1214
directive gain, 1243
end fire, 1251
grounded, 1230
horn, 1286
loop, 1274
measurements, 1300
non-resonant, 1228
parabolic, 1280
receiving, 1314
yagi, 1268
Appleton Hartree formula, 1340
Armstron phase modulator, 1501
Astable multivibrator, 847, 980
Attenuator, 100
Automatic frequency control, 1548
Automatic volume control, 1468
Backward-wave oscillator, 712
Balanced modulator, 1406
Band elimination filter, 90
Band-pass filter, 86
Barkhausen condition, 627
Base diffusion, 940
Basic voltage amplifier, 409
Beam power tube, 121
BCD decade counter, 1044
Bistable multivibrator, 855, 865
Blocking oscillator, 867
Boolean algebra, 995
Bootstrap circuit, 872, 929
Bridge amplifier, 983
Broad band antenna principles, 130
Bunching in reflex klystron, 681
Carcinotron, 712
Cascaded tuned amplifier, 570
Cathode drift compensation, 613
Cathode follower, 498
RC coupled, 447
Cathode ray tube, 125
Cavity magnetron, 696
Chapman theory, 1373
Child-Langmuir Law, 318
cylindrical electrodes, 321
plane electrodes, 318
Choke input filter, 378
Clamping circuit, 908, 903
Clipping circuit, 900, 905
Clock pulse, 1034
Coaxial line, 1180

Colpitt's oscillator, 639, 836
Compound feed back, 493
Coupled circuit, 24
Critical frequency, 1339
Crystal oscillator, 651, 838
Current feed back, 490
  time base, 932
Current gain in transistor, 246
Delayed AVC, 1469
D type flip flop, 1041
D-layer of ionosphere, 1371
Decade counter, 880, 1044
De-emphasis circuit, 1550
Detection,
  of a m. waves, 1445
  of f.m. waves, 1475
Dirac, 280
Differentiating circuit, 890, 976
Diplexer, 1584
Diode detector, 1451
Discriminator. 1475
Distortion in amplifiers, 432
Double resnator klystron, 685
Double tuned transistor amplifier, 803
Drain follower, 819
Drift in d.. amplifier, 612
Duct propagation, 1330
Dynatronoscillator, 642
Electron coupled oscillator, 650
Electron emission, 297
Electronic regulator, 386
E-layer of ionosphere, 1370
Electron motion in time varying
  field, 659
Electron transit ce limitations, 663
  712
Electr
Filters for power supplies, 368
Folded dipole, 1306
Frequency conversion, 1519
Frequency distortion, 436
Frequency modulation, 1412
  demodulation, 1475
Frequency stability, 649
Full adder, 1015
Grid bias modulation, 1396
Gas laser, 1617
Gas tube, 154
Gates ; NOT, AND, OR. 1000, 996, 998
Gate amplifier FET, 820
Grid leak detection, 1450
Grid modulation, 1404
Grounded antennas, 1230
Grounded emitter circuit, 727
  base circuit, 735
  collector circuit, 742
Gunn e ect, 2o7
Half adder, 1014
Hartley oscillator, 637, 834
Height, effective, of antenna, 1219
High frequency compensation, 461
High pass filter, 85
High pass RC circuit, 890
Horn antenna, 1286
Hum, 1519
Hybrid parameters, 285
Ignitron, 165
Image frequency, 1533
Impatt, 273
Impedance matching, 48
Impedance of antennas, 12o7
Impurity semi-conduc
Inductance capacit
  plifier, 463
Input admittan
  mpedar

Acc·No 2832 2823352

Isolation diffusion, 939, 944
JFET. 259
JK flip flop, 1038
Junction diode, 224
Junction transistor, 242
Klystron, 674
Laser, 1614
Layers, ionospheric, 1372
Linear modulation methods, 1390
Limiter circuit, 1483
Logic families, 1025
—DTL, 1027
—ECL, 1030
—RTL, 1025
—TTL, 1029
Logarithmic amplifier, 979
Loftin whits D.C. amplifier, 608
Low frequency compensation, 459
Low pass filter, 84
m-derived filter, 93
π-section, 96
T-section, 84
Low pass RC circuit, 892
Magnetron oscillator, 689
cavity, 696
split anode, 692
Matrixing network, 1592
Maser, 1609
Master slave JK flip flop, 1040
Masking, 943
Maximum usable frequency, 1351
Miller effect, 428
integrating sweep, 924, 874
1384
1386
Noise in amplifiers, 590
Norton theorem, 67
Operational amplifiers, 953
inverting, 957
noninverting, 959
Orthicon, 1572
Oscillators, 625
backward wave, 712
colpitt's, 639
dynatron, 642
crystal controlled, 651
electron coupled, 650
Hartley, 637
klystron, 674
magnetron, 689
multivibrator, 654
R. C., 645
tuned grid, 634
tuned plate, 631
Wien bridge, 648, 841
Oscillators (Transistor)
Colpitt, 836
Hartley, 834
tuned collector, 831
tuned emitter, 833
Parabolic reflector, 1280
Parallel resonant circuit, 15
Pentagrid converter, 1529
Pentode, 117
Phase distortion, 437
Phase modulation, 1421
Phase shift oscillator, 645, 839
Photoelectric emi
Photoem
Photor
Phot

Jeshiva Rabbi Shehziele Kan Rongeiey